Rolf Steyer

Probability and Causality

Conditional and Average Total Effects

Version: November 15, 2024

University of Jena

Preface

What can we do to reduce global warming? How can we prevent another global financial crisis? How to fight AIDS? What is the effect of being infected by a virus on dying within a fixed amount of time? And which political measures can save lives in a specific virus pandemic? These and similar questions ask about causal effects of political and social interventions, of medical or psychological treatments, or of expositions to potentially harmful or beneficial expositions, for example, to a virus or to asbestos dust. Obviously, interventions based on the wrong causal theories and hypotheses will cost the lives of thousands, huge amounts of money that could be spent more appropriately, and fundamentally change our lives. Even if our daily problems beyond these issues are less dramatic, they are of the same nature. Just think about your own actions that you have to chose in your responsibilities as a student, scientist, teacher, physician, psychologist, politician, or as a parent! Whatever you do has direct, indirect, and total effects, and these effects might be different if you take one action instead of another one. Furthermore, on the individual level, these effects differ for different persons. The effects of being infected by a virus may be much more severe for persons above 80 years of age if they already have severe health problems as compared to younger persons or those who do not have such problems. It is these kind of thoughts that make me believe that there is no other issue in the methodology of empirical sciences that deserves and needs more attention and effort than causality. And because the dependencies we are investigating are of a nondeterministic nature, we need a *probabilistic theory of causality*. In other words, we need to understand *probability* and *causality*.

Why Another Book on Causality?

The simple reason is that I wanted to understand the concepts of causality that are and should be used in our theories in science and in practical live. Existing theories and their answers never satisfied my own standards of what science should be. For me 'understanding' includes to be able to define all relevant concepts in terms of mathematics. None of the main stream theories, neither Rubin's potential outcome approach nor Pearl's graphical modeling approach satisfy this criterion. Although these pioneers contributed tremendously to our understanding and the popularity of the issue of causality in the empirical sciences, they lack mathematical rigor, thus opening the ground for misunderstandings and endless and all too often fruitless discussions.

This book presents a *mathematical theory* of causality. All terms are well-defined in terms of measure and probability theory, and all relevant propositions have a mathematical proof. This does not exclude that I made some mistakes. However, it is now possible to find and correct them, where they should have occurred. And, of course, the theory is still far from being complete.

What This Book is About

Empirical causal research involves several inferences and interpretations. Among these are:

(a) Statistical inference, that is, the inference from a data sample to parameters characterizing the distributions of random variables.
(b) Causal inference, that is, the inference from parameters characterizing the distributions of random variables to causal effects and/or dependencies.
(c) The substantive interpretation or meaning of the putative cause.
(d) The substantive interpretation or meaning of the outcome variable.
(e) The specification of the random experiment considered.

This book does not deal with all these points. We will neither discuss the mathematics of statistical inference nor the content issues of construct validity or external validity (Campbell & Stanley, 1963; Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002) involved in points (c) to (e). Instead we will focus on the second point: causal inference, that is, the inference from true parameters (i. e., not their estimates) to causal effects. Parameters characterizing the distributions of random variables such as the conditional expectation values of an outcome variable in two treatment conditions have, per se, no causal interpretation. The inference from true parameters to causal effects is what the probabilistic theory of causal effects and this book are about. As will be shown, causal effects are also parameters that characterize the joint distributions of the random variables considered in a random experiment. However, their definitions are less obvious than 'ordinary' conditional expectation values and their differences. And, sometimes causal effects are identical to differences between ordinary conditional expectation values and sometimes they are not. In other words, sometimes we can have a positive difference between conditional expectation values that seemingly indicates a positive treatment effect, whereas the causal effect is negative, and vice versa.

Basic Idea

In order to get a first impression of what this means, let us briefly formulate the basic idea that can most easily be explained if the putative (or presumed) cause is a treatment or intervention variable. Suppose an individual, or in more general terms, an observational unit (this can also be a country), could be treated by condition 1 or it could be treated by condition 0, *everything else being invariant*. If there is a difference in the outcome considered (some measure of success of the treatment), then this difference is due to the difference in the two treatment conditions. This idea goes back at least to Mill (1843/1865).

Multiple Determinacy

The problem with this first version of the basic idea is that most outcomes are *multiply determined*, that is, they are not only influenced by the treatment variable, but by many other variables as well. In the field of agricultural research, for example, the *yield* (*outcome*) of a *variety* does not only depend on the variety (*treatment*) itself, but it also depends on the quality of the *plot* (*observational unit*), such as the average hours of sunshine on the plot per day, the amount of water reaching the plot, and the number of microbes in the plot, and so on. Although Mill's idea sounds perfect, it is not immediately clear which implications it has for practice, because the number of other causes is often too large for keeping constant all of them. Furthermore, Mill's idea fails to distinguish between potential confounders and intermediate variables. Holding constant all intermediate variables as well — and not only all pretreatment variables — would imply that there is no treatment effect any more, if we assume that all treatment effects have to be transmitted by some intermediate variables.

Because of the problem of multiple determinacy, Mills conception has been complemented by Sir Ronald A. Fisher (1925/1946) and by Jerzy S. Neyman (1923/1990) in the second and third decades of the last century. Simply speaking, emphasizing and propagating the randomized experiment, Fisher replaced the *ceteris paribus* clause ('everything else invariant') by the *ceteris paribus distributionibus* clause: *all other possible causes (the 'pretreatment variables') having the same distribution.* This is what randomized assignment of units to treatment conditions, for example, based on a coin flip, secures.

A Metaphor — The Invisible Man and His Shadow

Imagine an invisible man. Although we cannot see him, suppose we know that he is there, because we can see his shadow. Furthermore, suppose we would like to measure his size. Doing that, we have two problems, a theoretical and a practical one. The *theoretical problem* is to define *size*. We have to clarify that we do not mean 'volume' or 'weight', but 'height' — without shoes, and without hat and hair. Unfortunately, actual height varies slightly in the course of a day. Hence, we define *size* to be the expectation (with respect to the uniform distribution over the 24 hours) of the momentary heights. This solves the theoretical problem; now we know what we want to measure.

However, because the man is invisible, we cannot measure his *size* directly — and this is not only because his size slightly varies over the day. The crucial problem is that we can only observe his shadow. And this is the *practical problem*: How to determine his size from his shadow? Sometimes, there is almost no shadow at all, sometimes it is huge. Some geometrical reflection yields a first simple solution: measuring the shadow when the sun has an angle of 45°. But what if it is winter and the sun does not reach this angle? Now we need more geometrical knowledge, taking into account the actual angle of the sun and the observed length of the shadow. This will yield an exact measure of the *size* of the invisible man at this time of the day as well.

The Perfect Randomized Experiment

Determining a causal effect we face the same kind of problems. First, we have to define a *causal effect*, and second, we have to find out how to determine it from empirically estimable parameters such as true means, that is, from conditional expectation values. The simple solution — corresponding to the 45° angle of the sun in the metaphor — is the perfect randomized experiment. The sample mean differences we observe in a randomized experiment only randomly deviate from the causal effect (due to random sample variation). In contrast, in quasi-experiments and observational studies, solutions to the practical problem are more sophisticated. They are also more sophisticated than in the metaphor of the invisible man, because it is not only *one* other variable (the angle) that determines the length of the shadow; instead there often are *many* other variables systematically determining the sample means as well as the true means that are estimated by these sample means. This is again the problem of multiple determinacy. Furthermore, a true effect to be estimated may even be negative although the true causal effect is positive, and vice versa. And this reversal of effects can be systematic, and not only be due to sampling error.

This book presents a solution to the theoretical and the practical problems mentioned above. Unfortunately, both solutions are not as simple and obvious as in our metaphor. Furthermore, there is not only one single kind of causal effects, even if we restrict ourselves to total causal effects and do not consider direct and indirect effects.

Total Individual and Average Causal Effects

To our knowledge, the first pioneer tackling the theoretical *and* the practical problems was Jerzy S. Neyman (1923/1990). While Fisher propagated the design technique of randomization, Neyman introduced the concepts of total individual and average causal effects, thus attempting a first solution to the theoretical problem mentioned above. (Note, however, that he used different terms for these concepts). Developing statistical methods for agricultural research, he assumed that, for each individual plot, there is an intra-individual (i. e., plot-specific) distribution of the outcome variable, say Y, under each treatment. He then defined the *individual causal effect of treatment x compared to treatment x'* to be the difference between the intra-individual (plot-specific) expectation of Y (the "true yield") given treatment ("variety") x and the intra-individual (plot-specific) expectation of Y given treatment ("variety") x'. Once the individual causal effect is defined, the *average treatment effect* of x compared to x' on Y is simply the expectation (true mean) of the corresponding individual (plot-specific) causal effects in the set (population) of observational units (plots). Similarly, several kinds of *conditional effects* can be defined, conditioning, for instance, on covariates, that is, on other causes of Y that cannot be affected by X, such as measures of the *quality of the soil* before treatment, *average hours of sunshine*, *average hours of rain*, and so on.

Total, Direct, and Indirect Effects

At about the same time as Neyman and Fisher developed their ideas, Sewall Wright (Wright, 1918, 1921, 1923, 1934, 1960a, 1960b) developed his ideas on path analysis and the concepts of total, direct, and indirect effects. While his *total effect* aims at the same idea as the causal total average effect, his *direct* and *indirect effects* were new. Simply speaking, in the context of an experiment or quasi-experiment, a direct effect of the treatment is the effect that is not transmitted through the intermediate variables; it is the conditional effect of the treatment variable holding constant the intermediate variables on one of their values. In contrast, the *indirect effect* is the difference between the total effect and the direct effect.

Fundamental Problem of Causal Inference

Whereas the basic ideas outlined above are relatively simple and straightforward, trying to put them into practice — that is, solving the practical problem mentioned above — is often difficult and needs considerable sophistication. The “fundamental problem of causal inference” (Holland, 1986) is that we cannot expose an observational unit to treatment 1 and, at the same time, to treatment 0. However, this is exactly what is necessary if we want to be sure that ‘everything else is invariant’, a clause that is also an implicit assumption in the solution proposed by Neyman. Comparing the true yield of treatment 1 to treatment 0 *within the same plot* at the same time and identical conditions is an ideal version of the *ceteris paribus* clause, which unfortunately is rarely accomplishable.

Pre-Post Designs

If we choose to first observe a unit under ‘no treatment’ and then observe it again after ‘treatment’, we may be tempted to interpret the pre-post differences as estimates of the individual causal effects of the treatment given in between. However, this interpretation might be wrong, because the unit may have developed (maturated, learned), may have suffered from critical life events, may have experienced historical change, and so on (see, e. g., Campbell & Stanley, 1963; Cook & Campbell, 1979; Shadish et al., 2002). Hence, in these *pre-post designs* or synonymously, *within-group designs*, we have to make assumptions on the nature of these possible alternative interpretations of the pre-post comparisons, for example, that they do not hold in the application considered or that they have a certain structure that can be taken into account when making causal inferences based on pre-post comparisons.

Between-Group Designs

If, instead of making comparisons within a unit, we compare different units to each other in *between-group experiments*, we certainly lose the possibility of estimating the *individual* causal effects. However, what we can hope for is that we are still able to estimate the *causal average total effect* and certain *causal conditional total effects*. But how to estimate the average of the causal individual total effects if, due to the fundamental problem of causal inference, the causal individual total effects are not estimable? Both, between-group experiments and quasi-experiments, have a set of (observational) units, at least two experimental conditions (‘treatment conditions’, ‘expositions’, ‘interventions’, etc.), and at least one outcome variable (‘response’, ‘criterion’, ‘dependent variable’) Y. In the medical sciences, the units are usually patients. In psychology the observational units are often persons, but it could be persons-in-a-situation, or groups as well. In economics it could be subjects, companies, or countries, for instance. In educational sciences the units might be school classes, schools, communities, districts, or countries. In sociology and the political sciences, the units could be persons, but also communities, countries, and so on. In this book we show how to define and also how to make inferences about the average of the causal individual total effects in such sets (and subsets) of observational units and about causal conditional total effects, conditioning on attributes of the observational units or on pretest scores, for instance.

Scope of the Theory

In order to delineate the scope of the theory, consider the following kind of *random experiment*: Draw an observational unit u (e. g., a person) out of a set of units, observe the value z of a (possibly multivariate qualitative or quantitative) covariate Z for this unit, assign the unit or observe its assignment to x, one of several experimental conditions, and record the numerical value y of the outcome variable Y. We will use U to denote the random variable representing with its value u the unit drawn. Note that many observations can be made additionally to observing U, Z, X, and Y. Although this single-unit trial is a prototype of the kind of empirical phenomena the theory is dealing with, there are other single-unit trials in which the theory can be applied as well (see ch. 2). In fact, the theory is applicable far beyond the true (i. e., the randomized experiment) and the quasi-experiment. This includes applications in which the putative causes are *not* manipulable. In this volume, we also treat the case in which the putative cause is a continuous random variable (see, e. g., the causality conditions treated in chs. 8 or 9). The theory has its limitations only if there is no clear time order of the random variables considered as putative causes or outcomes.

True Experiments and Quasi-Experiments

The single-unit trial described above is a random experiment, but not necessarily a randomized experiment. A *randomized experiment* is a special random experiment in which the drawn unit is assigned to one of the treatment conditions *via randomization*, for example, depending on the outcome of a coin flip. (In empirical research, the single-unit trials are repeated n times, where n denotes the sample size.) Referring to single-unit trials, we can distinguish the *true experiment* from the *quasi-experiment* as follows: In the *true experiment*, there are at least two treatment conditions and the assignment to one of the treatment conditions is randomized, for example, by flipping a coin. In a traditional *randomized experiment*, for instance, the treatment probabilities are chosen to be equal for all units. However, equal treatment probabilities for all units are neither essential for the definition of the true experiment nor for drawing valid causal inferences. We may as well have treatment probabilities depending on the units and/or on a covariate (for more details, see, e. g., Rem. 8.59), *as long as these treatment probabilities are fixed or known by the researcher*. Note, however, that in designs, in which different units have different treatment probabilities, standard techniques of data analysis such as t-tests or analysis of variance do not test the hypothesis about a causal effect any more.

For between-group designs, the *quasi-experiment* may be defined such that there are at least two treatment conditions; however, in contrast to the true experiment, the treatment probabilities are unknown. Nevertheless, valid causal inferences can be drawn in quasi-experiments *provided that we can rely on certain assumptions* (see the causality conditions treated in Part III of this book. In specific applications these assumptions might be wrong. If they are actually wrong, causal inferences can be completely wrong as well.

Who Should Study This Book?

The Methodologist

In the first place, I would like to address the *methodologist*, that is, the expert in empirical research methodology, especially in the social, economic, behavioral, cognitive, medical, agricultural, and biological sciences. This book provides answers to some of the most important and fundamental questions of these empirical sciences: What do we mean by terms like 'X affects Y', 'X has an effect on Y', 'X influences Y', 'X leads to Y', and so on used in our informal theories and hypotheses? How can we translate these terms into a precise language (i. e., probability theory) that is compatible with the statistical analysis of empirical data? How to design an empirical study and how to analyze the resulting data if we want to probe our theories and learn from such data about the causal dependencies postulated in our theories and hypotheses? And last but not least: How to evaluate interventions, treatments, or expositions to (possibly detrimental) environments, and learn about how which effects they have for which kind of subjects or observational-units, and under which circumstances?

The Statistician

Many statisticians believe that causality is beyond the horizon of their profession. Causality might be a matter of empirical researchers and philosophers, they say, but not their own. They think that it cannot be treated mathematically and therefore a statistician should refrain from causal interpretations. As a consequence, they ignore the issue of causality. This book proves that these beliefs are prejudices. The theory of causal effects, as presented here, is a branch of probability theory, which itself, at least since Kolmogorov (1933/1977), is a part of pure mathematics — although with an enormous potential for applications in many empirical sciences and even beyond. The main purpose of this book is to translate the informal concepts about causal effects shared by many methodologists and applied statisticians into well-defined terms of mathematical probability theory. The principle is not to use any term that itself is not defined in other mathematic terms, and the result is a purely mathematical theory of causal effects. Of course, this will make it harder to read this book for the methodologist and those not yet trained in probability theory. However, the reward is a much deeper understanding of what is essential and a much better grasp of the nature of our theories about the real world.

Of course, undefined terms are still used in this book, but only in the examples, in the interpretations, and in the motivations of the definitions. The theory itself is pure mathematics, just in the same way as Kolmogorov's probability theory presented in 1933, which explicated the mathematical, measure-theoretical structure of probabilistic concepts. Substantive meaning results, for example, if we interpret the core components of the formal structure in a specific random experiment considered. And this is also true for the theory of causal effects presented in this book.

The Empirical Scientist

The empirical scientist in the fields mentioned above has at least three good reasons to study this book. The *first* is that some crucial parts of his theories and hypotheses are explicated, at least when it comes to considering a concrete experiment or study. The ambiguity in causal language such as 'X affects Y', 'X has an effect on Y', 'X influences Y', 'X leads to Y' are not necessary any more. Reading this book will make it possible to replace these ambiguous terms by well-understood and well-defined terms, improving the precision of empirical research and theories.

The *second* motivation of the empirical scientist is that even if he knows his own theoretical concepts and hypotheses, he still has to know how to design experiments and studies that enable him to test them empirically.

Third, the standard ways of analyzing data offered in the textbooks of applied statistics and in the available computer programs often do not estimate and test the causal effects and dependencies we refer to in our theories. And this is not only bad for the empirical scientist but also for all those relying on the validity of his inferences and his expertise. Just think about all the harmful consequences of wrong causal theories in various empirical research fields, if they are applied to solving concrete problems!

The Experimental Scientist

There are two messages for those who do their research with experiments, a good one and a bad one. The good news is that, in a perfect randomized experiment, the causal average total treatment effect is indeed estimated when comparing sample means between two different treatment conditions. The bad news is that *we can not rely on randomized assignment of units to treatment conditions* when it comes to estimating *direct* and *indirect* effects. More specifically, in such an analysis it is usually not sufficient to consider intermediate variables, treatment and outcome variables. Instead we also have to include in our analysis *pre-treatment variables* such as a pre-test of the intermediate variable and a pre-test of the outcome variable and apply adjustment methods, very much in the same way as we have to use these techniques in quasi-experiments. Hence, if we want to study the black box between the treatment and the outcome variables, we have to adopt the techniques of causal modeling that are far beyond traditional comparisons of means and analysis of variance. (For more details see, e. g., Mayer, Thoemmes, Rose, Steyer, & West, 2014).

The Philosopher of Science

Philosophers of science study and teach the methodology of empirical sciences. In that respect, their task is very similar to that of the methodologist, perhaps only more general and less specific for a certain discipline. Therefore, it is not surprising that probabilistic causality has also been tackled by philosophers of science (see, e. g., Cartwright, 1979; Spohn, 1980; Stegmüller, 1983; Suppes, 1970). Compared to these approaches, our emphasis is more on those parts of the theory that have implications for the *design* of empirical studies and the *analysis of data* resulting from such studies.

The Students in These Fields

For reasons detailed before, I believe that the probabilistic theory causal effects and dependencies is the most rewarding topic in methodology. Although it is tough to get into it, you will get insights why all this methodology stuff was useful and what it was good for. At least this is what many of my students said at the end of their curriculum, even if they did not have the choice whether or not to take the courses on causal effects.

Research Traditions in Causal Effects

Several research traditions have been contributing to the theory of causal effects in various ways. From the *Neyman-Rubin tradition*, I adopted the idea that it is important to define various causal effects such as individual, conditional, and average total effects, even though we modified and extended these concepts in some important aspects. Defining causal effects is important for proving that certain methods of data analysis yield unbiased estimates of these effects if certain assumptions can be made. Are there conditions under which the analysis of change scores (between pre- and post-tests) and repeated-measures analysis of variance yield causal effects? Under which conditions do we estimate and test causal effects in the analysis of covariance? Which are the assumptions under which propensity score methods yield estimates of causal effects? Which are the assumptions under which an instrumental variable analysis estimates a causal effect? All these questions and their answers presuppose that we have a mathematical definition of causal effects. Simply speaking, Rubin's potential outcome variables are replaced by the true outcome variables (see ch. 5), allowing for variance in the outcome (or response) variables given treatment and an observational unit. Many important results of the theory, for example, about *strong ignorability* and about *propensity scores* remain unchanged, while other results are new, giving more insights, and open the floor for new research techniques.

From the *Campbellian tradition* (see, e. g., Campbell & Stanley, 1966; Cook & Campbell, 1979; Shadish et al., 2002) we learned that there are questions and problems beyond the theory causal effects itself that are relevant in empirical causal research, such as: How to generalize beyond the study? What does the treatment variable actually mean from a substantive point of view? What is the meaning of the outcome variable? And, perhaps the most general question: Are there alternative explanations for the effect? The vast majority of social scientists (including myself) have been educated in this research tradition to some degree. Although this training is still very useful as a general methodology framework, it lacks precision and clarity in a number of issues — and the definition of a causal effect is one of them that remains unnecessarily vague in their ideas dealing with interval validity.

From the *graphical modeling tradition* (see, e. g., Cox & Wermuth, 2004; Pearl, 2009; Spirtes, Glymour, & Scheines, 2000), we learned that conditional independence plays an important role in causal modeling. This research tradition has also been developing techniques to estimate causal effects and to search for causal models if specific assumptions can be made. The fact that randomization in a true experiment in no way guarantees the validity of causal inferences on *direct* effects has been brought up by this research tradition.

Structural equation modeling and *psychometrics* have been teaching us how to use latent variables and structural equation modeling in testing causal hypotheses. Due to a number of statistical programs such as AMOS (Arbuckle, 2006), EQS (Bentler, 1995), lavaan (Rosseel, 2012), LISREL (Jöreskog & Sörbom, 1996/2001), Mplus (Muthén & Muthén, 1998-2007), OpenMx (OpenMx, 2009), RAMONA (Browne & Mels, 1998), structural equation modeling became extremely popular in the social sciences. Although many users of these programs hope to find causal answers, it should be clearly stated that structural equation modeling — and this is true for all kinds of statistical models (including analysis of variance) — does neither automatically estimate and test causal effects, nor does it provide a satisfactory *theory* of causal effects and dependencies. Nevertheless, this research tradition contributes — just like other areas of statistics — a number of statistical techniques that can be very useful in causal modeling.

This book is aimed at embedding — and, where necessary, extending — conventional statistical procedures such as analysis of covariance, nonorthogonal analysis of variance, and latent variable modeling, but also more recent techniques based on propensity scores into a coherent theory of probabilistic causality.

Outline of Chapters

Chapter 1 is intended to motivate the theory by some examples. Specifically, it is shown that, in general, true mean differences do not have — without further assumptions — a valid causal interpretation. Furthermore, we present an example showing that, under certain conditions, traditional techniques of data analysis such as analysis of variance, systematically estimate the wrong parameters and test the wrong hypotheses if causal effects are aimed at.

In chapter 2, we describe in more detail some random experiments, that is, some of the empirical phenomena the theory is dealing with. These include, for example, a simple experiment in which we study the effect of a treatment on an outcome variable, but also more complex random experiments in which manifest or latent covariates (e. g., latent pretest variables) are measured, a treatment variable is observed, and manifest or latent outcome variables are assessed.

While the first two chapters are accessible without much knowledge in probability theory, the rest of the book is not. Therefore we recommend to study an introduction to probability and conditional expectation such as Steyer (2024) or Steyer and Nagel (2017) before reading the rest of this book and/or whenever these terms appear. These books are also often referred to using the abbreviations RS and SN, respectively.

In chapter 3, we introduce the additional mathematical concepts that are necessary for the definition of causal effects and that allow to meaningfully raise the question if a dependency of a random variable on another one describes a causal dependency. These terms include the concept of a *filtration* that is fundamental in the theory of stochastic processes. Using a filtration, we can introduce time order between events, families of events, and random variables. It will also be used to distinguish *potential confounders* from others kinds of random variables.

Chapter 4 introduces the mathematical structure that allows us to define a *putative cause variable* X, an *outcome variable* Y *of* X, and a *potential confounder of* X. In many cases, conditional expectations describing a causal dependence can be distinguished from conditional expectations that have no such causal interpretation by their relationship to the potential confounders of the putative cause variable X.

Chapter 5 starts with the concepts of *true outcome variables* and *true effect variables* that play am important role in the definition of total causal effects. Then several kinds of *conditional* and *average total effects* are introduced. As indicated above, the basic ideas are simple. However, in the subsequent chapters it will become obvious that the theory has tremendous consequences, for the design of experiments, for data analysis in experiments and quasi-experiments, and beyond.

In chapter 6, we will study how (conditional and unconditional) prima facie effects (*PFE*s) are related to the different kinds of causal effects. Among these prima facie effects are the true mean differences $E(Y|X{=}x) - E(Y|X{=}x')$ between two treatment conditions

x and x', and the corresponding mean differences $E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z)$ conditioning also on the value z of a covariate Z. For all these concepts, we introduce the notion of *unbiasedness with respect to total effects*.

In chapter 7 we intensively study the *Rosenbaum-Rubin conditions*, which deal with the relationship between true outcome variables and the putative cause variable. The weakest of these conditions is equivalent to unbiasedness, the strongest is the probabilistic version of strong ignorability (see, e.g., Rosenbaum & Rubin, 1984). Although these conditions are of theoretical interest, they can neither be tested empirically nor can they be used for covariate selection.

Chapter 8 deals with other causality conditions, the *Fisher conditions*, which imply unbiasedness if the putative cause variable is discrete. However, these conditions also hold for a continuous putative cause variable, for which unbiasedness and true outcome variables are not defined. These causality conditions can deliberately be created by the design technique of randomized assignment of a unit to one of the treatment conditions. Furthermore, these causality conditions are the first ones treated that can be tested in empirical applications.

In contrast to the Fisher conditions, which deal with the relationship between the putative cause variable and the potential confounders, the *Suppes-Reichenbach conditions* treated in chapter 9 focus the relationship between the outcome variable and the potential confounders. Just like the Fisher conditions, they are empirically testable and can therefore be used for covariate selection.

In chapter 10 we present a fourth kind of causality conditions, the *unconfoundedness conditions*. To our knowledge, they are the weakest conditions implying unbiasedness that are still empirically testable.

In chapter 11 we show that propensities can also be used instead of the original covariates to adjust the expectations of the outcomes in the treatment conditions and to compute the average causal effect. This also involves *weighting the outcome variable* with a function of the propensity scores.

How to Use This Book

This book is written such that standard mathematical probability theory is sufficient for a complete understanding, provided one takes the time that these topics require. In many parts, this is not a book one can just *read;* instead it is a book to be *studied*. This includes working on the questions and exercises provided in each chapter. We presume that the reader is familiar with — or learns while studying this book — the essentials of probability theory, including not only random variables and their distribution, but also conditional expectations and conditional independence. These essentials of probability theory are extracted in Steyer (2024) from the more complete and detailed book Steyer and Nagel (2017). Both books are also referred to very often for definitions, theorems, and other propositions used in this text. The references to Steyer (2024) are abbreviated by RS-Definition, RS-Theorem, RS-Remark, or RS-(3.3), the latter referring to an equation or a proposition in that book. Similarly, references to Steyer and Nagel (2017) are abbreviated by SN-Definition, SN-Theorem, SN-Remark, or SN-(10.32), for example.

The largest part of this book is devoted to the *theory* of causal effects. The *Causal Effects Explorer* (Nagengast, Kröhne, Bauer, & Steyer, 2007) can be used for exploring prima facie effects, conditional and average total effects given certain parameters. Furthermore, the program *EffectLiteR* (Mayer, Dietzfelbinger, Rosseel, & Steyer, 2016), can be used to estimate conditional and average total effects from empirical data in experiments and quasi-experiments. Both programs, which are available at *www.causal-effects.de*, may be used together with this book in a course on causal modeling. In fact, this is the content of my workshops on the analysis of causal conditional and average total effects, which are available both as videos-on-demand on the internet and on DVDs, again at *www.causal-effects.de*.

Acknowledgements

This book has been written with the help of several colleagues and students. Werner Nagel (FSU Jena) helped whenever I felt lost in probability spaces. Stephen G. West (Arizona State University) and Felix Thömmes (Cornell University) made detailed suggestions for improving readability of the book. Safir Yousfi and Sonja Hahn contributed and/or suggested concrete ideas, Sonja being extremely helpful also in checking some of the mathematics. Our students Franz Classe, Lisa Dietzfelbinger, Niclas Heider, Marc Heigener, Remo Kamm, Lawrence Lo, David Meder, Marita Menzel, Yuka Morikawa, Sebastian Nitsche, Fabian Schäfer, Michael Temmerman, Sebastian Weirich, Anna Zimmermann, and other students critically commented on previous versions, helped minimizing errors, or organizing the references. Uwe Altmann, Linda Gräfe, Sven Hartenstein, Ulf Kröhne, Axel Mayer, Marc Müller, Christof Nachtigall, Benjamin Nagengast, Andreas Neudecker, Ivailo Partchev, Jan Plötner, Steffi Pohl, Norman Rose, Marie-Ann Sengewald, and Andreas Wolf together with the others mentioned above provided the intellectual climate in which this book could be written. I am also grateful to the students and colleagues participating at our courses on the analysis of causal effects asking questions and making important comments. Over the years, this helped a lot to improve this book.

Jena, August 15, 2024

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1
Introductory Examples

For more than a century there have been examples in the statistical literature showing that comparing means or comparing probabilities (e. g., of success of a treatment) between a group exposed to a treatment and a comparison group (unexposed or exposed to a different treatment) does not necessarily answer our questions: 'Which treatment is better overall?' or 'Which treatment is better for which kind of person?' Differences between true means and differences between probabilities (or any other comparison between probabilities such as odds ratios, log odds ratios, or relative risk) are usually not the treatment effects we are looking for (see, e. g., Pearson, Lee, & Bramley-Moore, 1899; Yule, 1903; Simpson, 1951). They are just *effects at first sight* or "prima facie effects" (Holland, 1986).

Just like the shadow in the metaphor of the invisible man presented in the preface, prima facie effects reflect the effects of the treatment (the size of the invisible man), but also the effects of other causes (the angle of the sun). The goal of analyzing *causal* effects is to estimate the effect of the treatment alone, isolating it from other potential determinants, such sex, educational background, socio-economic status, and so on. The general idea is to define and, in applications, estimate a treatment effect that is not biased by preexisting differences between treatment groups that would also be observed after treatment if there were no treatment effect at all.

Overview

We illustrate systematic bias in determining *total* (as opposed to direct or indirect) treatment effects in quasi-experiments by two examples. The first one deals with a dichotomous outcome variable, the second with a quantitative one. Note that the problems described in these two examples cannot occur in a randomized experiment, but they are ubiquitous in nonrandomized quasi-experimental observational studies.

1.1 Example 1 — Joe and Ann With Self-Selection

In this example, the prima facie effect reverses if we switch from comparing the conditional probabilities of success between treatment and control, that is, from comparing

$$P(Y=1|X=1) \quad \text{to} \quad P(Y=1|X=0)$$

to comparing the corresponding probabilities additionally conditioning on the person variable U with values $u \in \{Joe, Ann\}$, that is, to comparing

$$P(Y=1|U=u, X=1) \quad \text{to} \quad P(Y=1|U=u, X=0).$$

Table 1.1. Joe and Ann with self-selection

Person u	$P(U=u)$	$P(X=1 \mid U=u)$	$P(Y=1 \mid U=u, X=0)$	$P(Y=1 \mid U=u, X=1)$
Joe	.5	.04	.7	.8
Ann	.5	.76	.2	.4

This kind of phenomenon, which is already known at least since Yule (1903), is called *Simpson's paradox* (Simpson, 1951), and it is still being debated (see, e. g., Hernán, Clayton, & Keiding, 2011; Wang & Rousseau, 2021).

Table 1.1 shows the parameters specifying a *random experiment* that is composed of three parts.

(1) A person is sampled from a set of two persons, Joe and Ann, with identical probabilities for each person u, that is, with probability $P(U=u) = .5$.
(2) If Joe is sampled, then he obtains treatment $(X=1)$ with probability $P(X=1 \mid U=Joe)$ $= .04$. If Ann is sampled, then she is treated with probability $P(X=1 \mid U=Ann) = .76$. (These numbers may reflect self-selection to treatment and the different inclinations of the two persons to go to treatment.)
(3) If Joe is sampled and *not treated*, then his probability $P(Y=1 \mid U=Joe, X=0)$ of success is .7. If he is sampled and *treated*, then his probability $P(Y=1 \mid U=Joe, X=1)$ of success is .8. In contrast, if Ann is sampled and *not treated*, then her probability $P(Y=1 \mid U=Ann, X=0)$ of success is .2, and if she is sampled and *treated*, then her probability $P(Y=1 \mid U=Ann, X=1)$ of success is .4.

This table describes a random experiment and it contains all information we need to compute the causal total effects of the treatment on the outcome variable Y (success), including the *causal conditional total effects given the person* and the *causal average total effect* of the treatment variable. These terms will be defined in chapter 5 and computed for this example in section 1.1.4.

Note that Table 1.1 does not describe a *randomized experiment*, in which, by definition, the treatment probabilities $P(X=1 \mid U=u)$ would be identical for all observational units u. Instead, it describes a *random experiment*, which is that kind of empirical phenomenon that we usually consider when we apply probability theory using terms such as random variables, their expectations, variances, distribution, correlations, etc. In inferential statistics it is those concepts about which we formulate our hypothesis and that we try to estimate in a sample.

In probability theory, we consider such a random experiment from the *pre facto perspective*. Hence, we do not consider data that would result from actually conducting such a random experiment. Data are only important in order to learn from observations about the laws of a random experiment. Data analysis is only a way to learn about these laws. But it these laws of the random experiment that are of primary interest. More precisely, if we know the eight probabilities displayed in Table 1.1, then we have all the information that we need to compute the causal conditional and average total effects of the treatment on the outcome (success). All it needs is to define these concepts in terms of probability theory, and this is what this book is about. Causal effects are nothing philosophical or even metaphysical. Instead, they are parameters that can be computed from the probabilities

Table 1.2. Joe and Ann with self-selection – explicit table

Possible outcomes ω_i (Unit, Treatment, Success)	$P(\{\omega_i\})$	Observables: Person variable U	Observables: Treatment variable X	Observables: Response variable Y	Conditional probabilities: $P(X=1 \mid U)$	$P(Y=1 \mid X)$	$P(Y=1 \mid X, U)$
ω_1 = (*Joe*, *no*, −)	.144	*Joe*	0	0	.04	.6	.7
ω_2 = (*Joe*, *no*, +)	.336	*Joe*	0	1	.04	.6	.7
ω_3 = (*Joe*, *yes*, −)	.004	*Joe*	1	0	.04	.42	.8
ω_4 = (*Joe*, *yes*, +)	.016	*Joe*	1	1	.04	.42	.8
ω_5 = (*Ann*, *no*, −)	.096	*Ann*	0	0	.76	.6	.2
ω_6 = (*Ann*, *no*, +)	.024	*Ann*	0	1	.76	.6	.2
ω_7 = (*Ann*, *yes*, −)	.228	*Ann*	1	0	.76	.42	.4
ω_8 = (*Ann*, *yes*, +)	.152	*Ann*	1	1	.76	.42	.4

of events and the distributions of random variables pertaining to the random experiment considered.

Table 1.2 describes the same random experiment as Table 1.1, but in a different way. The eight triples such as (*Joe*, *no*, −) or (*Ann*, *yes*, +) represent one of the eight possible outcomes $\omega_1, \dots, \omega_8$ of the random experiment that are gathered in the set Ω of possible outcomes. Remember, an event A is a subset of Ω that has a probability $P(A)$, which is assigned by the probability measure P to each element A in the set $\mathscr{A}$ of all events (see RS-ch. 1 for these elementary concepts of probability theory). The eight probabilities of the elementary events contain the same information as the eight probabilities in Table 1.1. All conditional probabilities appearing in Table 1.2, but also all probabilities and all conditional probabilities presented in Table 1.1 can be computed from these eight probabilities of the elementary events.

Table 1.2 has the virtue of explicitly showing all possible outcomes of the random experiment considered. Furthermore, it shows how the random variables U, X, and Y are defined, showing the assignments of their values to each of the eight possible outcomes of the random experiment (see RS-Def. 2.2 for the definition of a random variable). It also displays the conditional probabilities $P(Y=1 \mid X, U)$, $P(Y=1 \mid X)$, and $P(X=1 \mid U)$, which are random variables on the same probability space as the observables, that is, the random variables U, X, and Y (see RS-Def. 4.1 and RS-Rem. 4.12 for the definition of such a conditional probability).

The crucial point is that each of the conditional probabilities mentioned above also assigns a value to each of the eight possible outcomes of the random experiment. For example, the values assigned by $P(X=1 \mid U)$ to each outcome $\omega_i \in \Omega$ are the conditional probabilities $P(X=1 \mid U=u)$. More precisely,

$$P(X=1 \mid U)(\omega_i) = P(X=1 \mid U=u), \quad \text{if } \omega_i \in \{\omega \in \Omega\colon U(\omega) = u\}. \tag{1.1}$$

Table 1.3. Joint and marginal probabilities of treatment and success

	Treatment		
Success	No ($X=0$)	Yes ($X=1$)	
No ($Y=0$)	.240	.232	.472
Yes ($Y=1$)	.360	.168	.528
	.600	.400	1.000

Note. The entries in the four cells are the joint probabilities $P(X=x, Y=y)$, the other entries are the marginal probabilities $P(X=x)$ (last row) and $P(Y=y)$ (last column).

Similarly,

$$P(Y=1 \mid X)(\omega_i) = P(Y=1 \mid X=x), \quad \text{if } \omega_i \in \{\omega \in \Omega\colon X(\omega) = x\}, \tag{1.2}$$

and

$$P(Y=1 \mid X, U)(\omega_i) = P(Y=1 \mid X=x, U=u), \quad \text{if } \omega_i \in \{\omega \in \Omega\colon X(\omega) = x, U(\omega) = u\} \tag{1.3}$$

(see Table 1.2 in order to check these assignment rules).

1.1.1 Joint Probabilities $P(X=x, Y=y)$

Unfortunately, in realistic applications, we cannot estimate the probabilities of all elementary events displayed in Table 1.2 because, usually, we cannot repeat this random experiment. Once, a person is treated, we cannot undo treatment and repeat this process. This has been called the "fundamental problem of causal inference" (Holland, 1986). Oftentimes a treatment is irreversible and the time between treatment and the assessment of the outcome variable Y can take months and even years. Nevertheless, it is meaningful to consider the random experiment described by Table 1.2 including the person-specific probabilities of treatment and the probabilities of success given treatment and person. As will be shown in chapter 5, we even *have to* consider this kind of random experiment when we want to *define* causal effects. And once we do consider such random experiments, causal effects are relatively easy to define.

What can be estimated in empirical applications are the joint probabilities such as $P(X=x, Y=y)$ and the conditional probabilities such as $P(Y=1 \mid X=1)$ and $P(Y=1 \mid X=0)$, that is, the conditional probabilities of success given treatment and given control, respectively. In order to estimate $P(X=x, Y=y)$ we only have to observe the relative frequencies of the joint occurrence of treatment x and outcome y. In a simulation, we can easily repeat this random experiment n times in order to generate a data sample of size n (see Exercise 1-7). In contrast, in an empirical application, estimating $P(X=x, Y=y)$ requires to sample from a very large set of persons (observational units), and not just from the set $\{Joe, Ann\}$ of two persons. (See Splawa-Neyman, 1923/1990 for an early sampling model dealing with the problem of non-replacement).

Table 1.3 shows the joint probabilities $P(X=x, Y=y)$ of treatment and success, as well as the marginal probabilities $P(X=x)$ and $P(Y=y)$ of treatment and success, respectively.

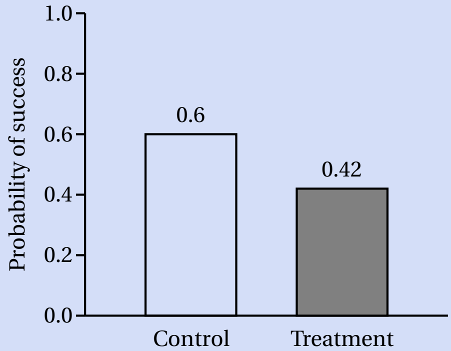

Figure 1.1. Probability of success given treatment conditions

These probabilities are easily computed from the probabilities of the elementary events displayed in the second column of Table 1.2. For example, the probability $P(X=0, Y=1)$ that the sampled person receives no treatment and is successful is the sum of the probabilities of the elementary events $\{\omega_2\} = \{(Joe, no, +)\}$ and $\{\omega_6\} = \{(Ann, no, +)\}$, that is,

$$P(X=0, Y=1) \;=\; P\big(\{(Joe,\ no,\ +)\}\big) + P\big(\{(Ann,\ no,\ +)\}\big) \;=\; .336 + .024 \;=\; .36.$$

Similarly, the probability $P(X=1, Y=1)$ that the sampled person receives treatment and is successful is the sum of the probabilities of the two elementary events $\{\omega_4\} = \{(Joe, yes, +)\}$ and $\{\omega_8\} = \{(Ann, yes, +)\}$, that is,

$$P(X=1, Y=1) \;=\; P\big(\{(Joe,\ yes,\ +)\}\big) + P\big(\{(Ann,\ yes,\ +)\}\big) \;=\; .016 + .152 \;=\; .168.$$

Table 1.3 is the theoretical analog to a contingency table that would be observed in a data sample. More precisely, if we multiply the displayed numbers by the sample size, then we receive the *expected frequencies* of the corresponding events. For example, if the sample size is 1000, then we expect 240 cases in cell $(X=0, Y=0)$ and 360 cases in cell $(X=0, Y=1)$, etc. Of course, in a data sample, the *observed frequencies* would fluctuate around these expected frequencies (see again Exercise 1-7).

1.1.2 Marginal Probabilities $P(X=x)$ and $P(Y=y)$

The marginal probabilities $P(X=x)$ and $P(Y=y)$ are also easily computed from the probabilities of the elementary events displayed in the second column of Table 1.2. For example, the probability $P(X=0)$ that the sampled person receives no treatment is the sum of the probabilities of the four elementary events $\{\omega_1\} = \{(Joe, no, -)\}$, $\{\omega_2\} = \{(Joe, no, +)\}$, $\{\omega_5\} = \{(Ann, no, -)\}$, and $\{\omega_6\} = \{(Ann, no, +)\}$, that is,

$$\begin{aligned} P(X=0) \;&=\; P\big(\{(Joe,\ no,\ -)\}\big) + P\big(\{(Joe,\ no,\ +)\}\big) + P\big(\{(Ann,\ no,\ -)\}\big) + P\big(\{(Ann,\ no,\ +)\}\big) \\ &=\; .144 + .336 + .096 + .024 \;=\; .6, \end{aligned}$$

and this implies

$$P(X=1) \; = \; 1 - P(X=0) \; = \; .4$$

Similary, the probability $P(Y=1)$ that the sampled person is successful is the sum of the probabilities of the four elementary events $\{\omega_2\} = \{(Joe, no, +)\}$, $\{\omega_4\} = \{(Joe, yes, +)\}$, $\{\omega_6\} = \{(Ann, no, +)\}$, and $\{\omega_8\} = \{(Ann, yes, +)\}$, that is,

$$\begin{aligned} P(Y=1) \; &= \; P\big(\{(Joe, no, +)\}\big) + P\big(\{(Joe, yes, +)\}\big) + P\big(\{(Ann, no, +)\}\big) + P\big(\{(Ann, yes, +)\}\big) \\ &= \; .336 + .016 + .024 + .152 \; = \; .528, \end{aligned}$$

which implies

$$P(Y=0) \; = \; 1 - P(Y=1) \; = \; .472.$$

1.1.3 Prima Facie Effect

Comparing the conditional probability $P(Y=1|X=1)$ of success given the *treatment condition* to the conditional probability $P(Y=1|X=0)$ of success given the *control condition* would lead us to the (wrong) conclusion that the *treatment is harmful*. These two conditional probabilities can be computed by

$$P(Y=1|X=1) \; = \; \frac{P(Y=1, X=1)}{P(X=1)} \; = \; \frac{.168}{.4} \; = \; .42$$

and

$$P(Y=1|X=0) \; = \; \frac{P(Y=1, X=0)}{P(X=0)} \; = \; \frac{.36}{.6} \; = \; .6,$$

respectively (see, e. g., RS-Def. 1.32 for the events $\{Y=1\}$, $\{X=0\}$, and $\{X=1\}$). Figure 1.1 displays both conditional probabilities in a bar chart.

These two conditional probabilities can be compared to each other in different ways. The simplest one is looking at the *difference* $P(Y=1|X=1) - P(Y=1|X=0)$. This is a particular case of the difference $E(Y|X=1) - E(Y|X=0)$ between two conditional expectation values, in which the outcome variable Y is dichotomous with values 0 and 1 (see RS-Rem. 3.22). Following Holland (1986), we will call this difference the (unconditional) *prima facie effect* and use the notation

$$PFE_{10} \; = \; E(Y|X=1) \; - \; E(Y|X=0) \; = \; P(Y=1|X=1) \; - \; P(Y=1|X=0).$$

Hence, in this example,

$$PFE_{10} \; = \; P(Y=1|X=1) \; - \; P(Y=1|X=0) \; = \; .42 \; - \; .6 \; = \; -.18.$$

Other possibilities of comparing the two conditional probabilities are to compute the odds ratio, its logarithm, or the risk ratio (see, e. g., SN-sect. 13.3 or chapter 4 of Rothman, Greenland, & Lash, 2008, for a detailed discussion of these and other effect parameters). No matter which of these effect parameters we choose, they all lead to the conclusion that the *treatment is harmful* (see Exercise 1-8). As shown in the following section this conclusion is utterly wrong.

Table 1.4. Joint and marginal probabilities of all three observables

Joe ($U{=}Joe$)

Success	Treatment No ($X{=}0$)	Treatment Yes ($X{=}1$)	
No ($Y{=}0$)	.144	.004	.148
Yes ($Y{=}1$)	.336	.016	.352
	.48	.02	.5

Ann ($U{=}Ann$)

Success	Treatment No ($X{=}0$)	Treatment Yes ($X{=}1$)	
No ($Y{=}0$)	.096	.228	.324
Yes ($Y{=}1$)	.024	.152	.176
	.12	.38	.5

Note. The entries in the four cells for Joe and in the four cells for Ann are the joint probabilities $P(U{=}u, X{=}x, Y{=}y)$. The other entries are the joint probabilities $P(U{=}u, X{=}x)$ (third and last row) and $P(U{=}u, Y{=}y)$ (last column), respectively, and the two marginal probabilities $P(U{=}u)$.

1.1.4 Individual Total Effects

The conclusion about the effect of the treatment is completely different if we look at the treatment effects separately for Joe and Ann. Table 1.4 shows the joint distributions of treatment, success, and the person variable U with values *Joe* and *Ann*. The probabilities of sampling Joe and of sampling Ann are identical, that is, $P(U{=}Joe) = P(U{=}Ann) = .5$. Furthermore, the joint probabilities $P(U{=}u, X{=}x, Y{=}y)$ are the probabilities of the elementary events displayed in the second column of Table 1.2. These joint probabilities are displayed again in a form analog to $(2{\times}2{\times}2)$-contingency table in Table 1.4.

As already mentioned in section 1.1.1, in empirical applications, this random experiment cannot be repeated in order to obtain a data sample. However, we can repeat it in a simulation (see Exercise 1-7). If, in such a simulation, we multiply the numbers displayed in Table 1.4 by the sample size, then we receive the *expected frequencies* of the corresponding events. For example, if the sample size is 1000, then we expect 144 cases in cell $(U{=}Joe, X{=}0, Y{=}0)$ and 336 cases in cell $(U{=}Joe, X{=}0, Y{=}1)$, etc. Of course, in data samples, the *observed frequencies* fluctuate around these expected frequencies.

Using the joint probabilities displayed in Table 1.4, the conditional probability of success for Joe in the treatment condition can be computed as follows:

$$P(Y{=}1 \,|\, X{=}1, U{=}Joe) \;=\; \frac{P(U{=}Joe, X{=}1, Y{=}1)}{P(U{=}Joe, X{=}1)} \;=\; \frac{.016}{.016 + .004} \;=\; .8$$

(see Exercise 1-9). In contrast, Joe's conditional probability of success in the control condition is

$$P(Y=1\,|\,X=0, U=Joe) \;=\; \frac{P(U=Joe, X=0, Y=1)}{P(U=Joe, X=0)} \;=\; \frac{.336}{.336+.144} \;=\; .7.$$

Hence,

$$P(Y=1\,|\,X=1, U=Joe) \;-\; P(Y=1\,|\,X=0, U=Joe) \;=\; .8 \;-\; .7 \;=\; .1,$$

which may lead us to conclude that *the treatment is beneficial for Joe*. Again, because Y is binary with values 0 and 1, this difference is a special case of the difference

$$E(Y\,|\,X=1, U=Joe) - E(Y\,|\,X=0, U=Joe),$$

which we call the *individual total (treatment) effect* of Joe, using the notation $ITE_{U;10}(Joe)$. Hence,

$$\begin{aligned} ITE_{U;10}(Joe) \;&=\; E(Y\,|\,X=1, U=Joe) \;-\; E(Y\,|\,X=0, U=Joe) \\ &=\; P(Y=1\,|\,X=1, U=Joe) \;-\; P(Y=1\,|\,X=0, U=Joe)\,. \end{aligned} \tag{1.4}$$

What about the individual total effect of Ann? Table 1.4 shows that the conditional probability of success for Ann in the treatment condition is

$$P(Y=1\,|\,X=1, U=Ann) \;=\; \frac{P(U=Ann, X=1, Y=1)}{P(U=Ann, X=1)} \;=\; \frac{.152}{.152+.228} \;=\; .4,$$

whereas it is

$$P(Y=1\,|\,X=0, U=Ann) \;=\; \frac{P(U=Ann, X=0, Y=1)}{P(U=Ann, X=0)} \;=\; \frac{.024}{.024+.096} \;=\; .2$$

in the control condition. Figure 1.2 shows these conditional probabilities in a bar chart. Considering the individual total effect

$$\begin{aligned} ITE_{U;10}(Ann) \;&=\; P(Y=1\,|\,X=1, U=Ann) \;-\; P(Y=1\,|\,X=0, U=Ann) \\ &=\; .4 \;-\; .2 \;=\; .2 \end{aligned} \tag{1.5}$$

of Ann may lead us to conclude that *the treatment is also beneficial for Ann*.

Hence, it seems that the treatment is *beneficial* for Joe and for Ann. This seems to contradict our finding ignoring the person variable. Just considering the prima facie effect

$$PFE_{10} \;=\; E(Y|X=1) - E(Y|X=0) \;=\; P(Y=1|X=1) - P(Y=1|X=0) \;=\; -.18,$$

ignoring the person variable U, the treatment seems to be *harmful*.

1.1.5 Prima Facie Effect Versus Expectation of the Individual Total Effects

In contrast to our intuition, the *prima facie effect* $E(Y|X=1) - E(Y|X=0)$ is *neither* the *simple average nor any weighted average of the corresponding individual total effects*

$$ITE_{U;10}(u) \;=\; E(Y|X=1, U=u) \;-\; E(Y|X=0, U=u)\,. \tag{1.6}$$

This is studied in more detail in the sequel.

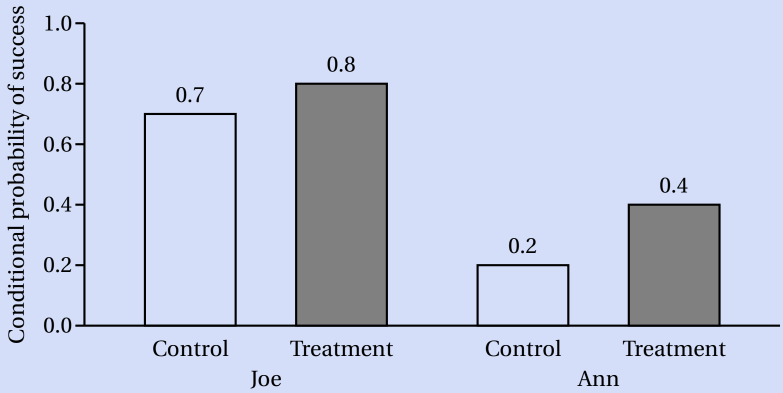

Figure 1.2. Conditional probabilities of success given treatment and person

Prima Facie Effect

The conditional probability $P(Y{=}1\,|\,X{=}0)$ of success given control is the sum of the corresponding probabilities $P(Y{=}1\,|\,X{=}0, U{=}\mathit{Joe})$ and $P(Y{=}1\,|\,X{=}0, U{=}\mathit{Ann})$, *weighted by the conditional probabilities* $P(U{=}\mathit{Joe}\,|\,X{=}0)$ and $P(U{=}\mathit{Ann}|\,X{=}0)$, respectively, that is,

$$\begin{aligned} P(Y{=}1\,|\,X{=}0) &= P(Y{=}1\,|\,X{=}0, U{=}\mathit{Joe})\cdot P(U{=}\mathit{Joe}\,|\,X{=}0) + \\ &\quad\; P(Y{=}1\,|\,X{=}0, U{=}\mathit{Ann})\cdot P(U{=}\mathit{Ann}|\,X{=}0) \\ &= .7\cdot\frac{.48}{.6} + .2\cdot\frac{.12}{.6} = .6 \end{aligned}$$

[see Box 3.2 (ii) and Exercise 1-10]. Because the difference between the conditional probabilities $P(U{=}\mathit{Joe}\,|\,X{=}0) = .48/.6$ and $P(U{=}\mathit{Ann}|\,X{=}0) = .12/.6$ is large, the probability of success in treatment 0 is much closer to .7 than to .2 (see the dots above $X{=}0$ in Fig. 1.3).

Similarly, the conditional probability $P(Y{=}1\,|\,X{=}1)$ of success given treatment condition $(X{=}1)$ is the sum of the two corresponding individual conditional probabilities $P(Y{=}1\,|\,X{=}1, U{=}\mathit{Joe})$ and $P(Y{=}1\,|\,X{=}1, U{=}\mathit{Ann})$, *weighted by the conditional probabilities* $P(U{=}\mathit{Joe}\,|\,X{=}1)$ and $P(U{=}\mathit{Ann}|\,X{=}1)$, respectively, that is,

$$\begin{aligned} P(Y{=}1\,|\,X{=}1) &= P(Y{=}1\,|\,X{=}1, U{=}\mathit{Joe})\cdot P(U{=}\mathit{Joe}\,|\,X{=}1) + \\ &\quad\; P(Y{=}1\,|\,X{=}1, U{=}\mathit{Ann})\cdot P(U{=}\mathit{Ann}|\,X{=}1) \\ &= .8\cdot\frac{.02}{.4} + .4\cdot\frac{.38}{.4} = .42. \end{aligned}$$

Hence, the *prima facie effect* is

$$\begin{aligned} PFE_{10} &= P(Y{=}1\,|\,X{=}1) - P(Y{=}1\,|\,X{=}0) \\ &= \sum_u P(Y{=}1\,|\,X{=}1, U{=}u)\cdot P(U{=}u\,|\,X{=}1) - \\ &\quad\; \sum_u P(Y{=}1\,|\,X{=}0, U{=}u)\cdot P(U{=}u\,|\,X{=}0) \\ &= .42 - .6. = -.18. \end{aligned} \tag{1.7}$$

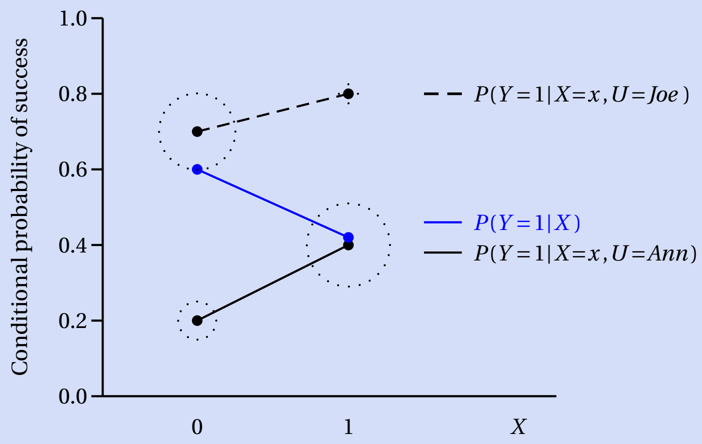

Figure 1.3. Conditional probabilities of success given treatment and person

Because the two $(X{=}1)$-conditional probabilities $P(U{=}Joe\,|\,X{=}1) = .02/.4 = .05$ and $P(U{=}Ann|\,X{=}1) = .38/.4 = .95$ are very different, the probability of success in treatment 1 is much closer to .4 than to .8 (see the dots above $X{=}1$ in Fig. 1.3). (The size of the area of the dotted circles is proportional to the conditional probabilities $P(U{=}u\,|\,X{=}x)$ that are used in the computation of the conditional expectation values $E(Y|X{=}x)$. [This kind of graphics has been adopted from Agresti, 2007)].

Average of the Individual Effects

The prima facie effect is not identical to the *expectation of the individual total effects*, which is the expectation of the random variable $ITE_{U;10}(U)$, the values of which are the two individual total effects $ITE_{U;10}(Joe)$ and $ITE_{U;10}(Ann)$ for Joe and Ann, respectively, that is,

$$\begin{aligned} E\big(ITE_{U;10}(U)\big) &= \sum_u ITE_{U;10}(u)\cdot P(U{=}u) \\ &= \sum_u P(Y{=}1|X{=}1, U{=}u)\cdot P(U{=}u) - \\ &\quad\ \sum_u P(Y{=}1|X{=}0, U{=}u)\cdot P(U{=}u). \end{aligned} \tag{1.8}$$

Because the two individual effects are $ITE_{U;10}(Joe) = .1$ and $ITE_{U;10}(Ann) = .2$,

$$E\big(ITE_{U;10}(U)\big) = .1\cdot P(U{=}Joe) + .2\cdot P(U{=}Ann) = .1\cdot\frac{1}{2} + .2\cdot\frac{1}{2} = .15\,.$$

Hence, whereas the *prima facie effect* $PFE_{10} = P(Y{=}1|X{=}1) - P(Y{=}1|X{=}0)$ is *negative*, namely −.18, the *expectation of the individual total-effect variable* is *positive*, namely .15 (see Exercise 1-15). This expectation will be called the *causal average total effect* of the treatment on the outcome variable Y, denoted ATE_{10}.

1.1.6 How to Evaluate the Treatment?

The conclusions drawn from the prima facie effect

$$PFE_{10} \;=\; P(Y{=}1\,|\,X{=}1) - P(Y{=}1\,|\,X{=}0)$$

and from the individual effects

$$ITE_{U;10}(u) \;=\; P(Y{=}1\,|\,X{=}1, U{=}u) - P(Y{=}1\,|\,X{=}0, U{=}u)$$

are contradictory. Which of these comparisons should we trust? Is the treatment harmful as $P(Y{=}1|X{=}1) - P(Y{=}1|X{=}0) = -.18$ suggests? Or is it beneficial as suggested by the two positive differences $P(Y{=}1|X{=}1, U{=}u) - P(Y{=}1|X{=}0, U{=}u)$? Which of these comparisons are meaningful for evaluating the causal total effect of the treatment on the success variable Y? Before we come back to these questions, we consider another example.

1.2 Example 2 — Experiment With Two Nonorthogonal Factors

In this section, we treat an example with three treatment conditions, representing two treatments and a control, for instance. Furthermore, there are a discrete covariate with three values, representing, for example, educational status, and a quantitative outcome variable, indicating the degree of success, for instance.[1]

Table 1.5 shows the table of a *random experiment*. Again this random experiment is composed of three parts.

(a) A person is sampled from a set of eight persons with identical probabilities for each person u, that is, with probability $P(U{=}u) = 1/8$.

(b) If Tom is sampled, then he obtains treatment 1 with probability $P(X{=}1\,|\,U{=}Tom) = 10/60$ and treatment 2 with probability $P(X{=}2\,|\,U{=}Tom) = 3/60$. The corresponding probabilities are also displayed for the other seven persons. The probabilities of getting treatment 0 can be computed from the displayed probabilities for treatment 1 and 2. For example, for Tom it is $P(X{=}0\,|\,U{=}Tom) = 1 - (10/60 + 3/60)$. And again, all these conditional probabilities may reflect self-selection to one of the treatments and the different inclinations of the persons to go to those treatments.

(c) After receiving treatment, a value of the outcome variable Y (success) is assessed. These values cannot be displayed in the table, because we assume that Y is continuous. Instead, the table displays the $(U{=}u, X{=}x)$-conditional expectation values of Y. If Tom is sampled and *not treated*, then his expectation value $E(Y\,|\,U{=}Tom, X{=}0)$

[1] In this example, we consider a (3×3)-factorial design with crossed, non-orthogonal factors. The analysis of such designs has been puzzling many statisticians (see, e. g., Aitkin, 1978; Appelbaum & Cramer, 1974; Carlson & Timm, 1974; Gosslee & Lucas, 1965; Jennings & Green, 1984; Keren & Lewis, 1976; Kramer, 1955; Langsrud, 2003; Nelder & Lane, 1995; Overall & Spiegel, 1969, 1973b, 1973a; Overall, Spiegel, & Cohen, 1975; Williams, 1972). In fact, none of the statistical packages such as SAS, SysStat, or SPSS with their Type I, II, III or IV sums of squares provide correct estimates and tests of the average effects (or main effects) for such a design unless the second factor has a uniform distribution, with equal probabilities for all values of the second factor. In this case Type III analysis yields correct results, at least, if the second factor is assumed to be fixed. However, in most applications in the social sciences, the second factor is not fixed but stochastic with a distribution of this factor (a qualitative random variable) varying between samples. Mayer and Thoemmes (2019) show how to conduct a correct analysis including average total effects (see also Exercise 1-17).

Table 1.5. Random experiment with two nonorthogonal factors

Person u	Educational status z	$P(U=u)$	$P(X=1\,|\,U=u)$	$P(X=2\,|\,U=u)$	$E(Y\,|\,X=0, U=u)$	$E(Y\,|\,X=1, U=u)$	$E(Y\,|\,X=2, U=u)$
Tom	*low*	1/8	10/60	3/60	120	100	80
Tim	*low*	1/8	18/60	9/60	120	100	80
Joe	*med*	1/8	26/60	17/60	90	90	70
Jim	*med*	1/8	26/60	17/60	100	100	80
Ann	*med*	1/8	26/60	17/60	120	100	100
Eva	*med*	1/8	26/60	17/60	130	110	110
Sue	*hi*	1/8	12/60	44/60	60	100	140
Mia	*hi*	1/8	16/60	36/60	60	100	140

of Y is 120. If he is sampled and receives treatment 1, then his expectation value $E(Y\,|\,U=Tom, X=1)$ is 100, and if he gets treatment 2, then his expectation value $E(Y\,|\,U=Tom, X=2)$ is 80. The corresponding conditional expectation values are also displayed in Table 1.5 for the other seven persons.

This table also contains the values of a qualitative covariate Z, which indicates an attribute of the person, his or her *educational status*. Because it is an attribute of the person, there is no extra sampling of Z. The value z of Z is fixed as soon as the person is actually sampled. In chapter 2, we will also deal with random experiments in which we *do have* an extra sampling process of one or several covariates. This will always be the case if, given the person, his or her value on Z is not fixed. A typical example is Z being a fallible pretest, for instance, a psychological test (say, of *live satisfaction*) that is not perfectly reliable, so that there is measurement error.

For this example to be realistic, we have to assume that there is still variation of the outcome variable Y in each combination of person and treatment condition. This conditional variance may be due to (a) *measurement error*, but also to (b) *mediator effects*, that is, to effects of variables and events that are in between X and the outcome variable Y in the process considered. Because Y is continuous and subject to measurement error and mediator effects, a full table similar to Table 1.2 with all possible outcomes is not feasible. It will not exist if Y is actually continuous, which would be true if we would assume, for example, that Y has a normal distribution given the combination of a person u and a treatment condition x.

Nevertheless, it is still possible to present a table that is analog to Table 1.1. In that table, the conditional probabilities $P(Y=1\,|\,X=x, U=u)$ are identical to the conditional expectation values $E(Y\,|\,X=x, U=u)$ if Y is binary with values 0 and 1 (see RS-Rem. 3.22).

Table 1.6. Conditional expectation values of the outcome variable Y given treatment

Treatment condition	$E(Y\,|\,X=x)$	$P(X=x)$
$X=0$ (Control)	111.25	1/3
$X=1$ (Treatment 1)	100	1/3
$X=2$ (Treatment 2)	114.25	1/3
$E(Y)$	108.5	

1.2.1 Prima Facie Effects

The conditional expectation values of the outcome variable Y given one of the three treatment conditions x are displayed in Table 1.6. The ratios in the last column are the treatment probabilities $P(X=x)$, which are 1/3 for all three values x of X. Note that this is *not* a randomized design as will become obvious if we look at the joint probabilities of X and Z (see Table 1.7). Furthermore, considering the conditional expectation values, and not the sample means, should make clear that we are not discussing *statistical inference* (i. e., inference from sample statistics to true parameters), but *causal inference*, that is, inference from the conditional expectation values such as $E(Y\,|\,X=x)$ or $E(Y\,|\,X=x, Z=z)$ to causal effects.

If our evaluation of the treatment effects were based on the differences between the conditional expectation values $E(Y\,|\,X=x)$ of Y in the three treatment conditions x, then we would conclude that there is a negative effect of treatment 1 compared to control, namely,

$$E(Y\,|\,X=1) - E(Y\,|\,X=0) = 100 - 111.25 = -11.25,$$

and a positive effect of treatment 2 compared to control, namely,

$$E(Y\,|\,X=2) - E(Y\,|\,X=0) = 114.25 - 111.25 = 3$$

(see Exercise 1-16).

1.2.2 $(Z=z)$-Conditional Prima Facie Effects

A second attempt to evaluate the 'effects' of the *treatment* is to look at the differences between the conditional expectation values of Y in the three treatment conditions *given one of the three values of Z*: *low*, *med*, and *hi*. These $(Z=z)$-conditional effects are also called *simple effects* in the literature on analysis of variance.

Table 1.7 displays the conditional expectation values of the outcome variable Y in the nine cells of the (3×3)-design. The ratios in parentheses are the probabilities that the pairs (x, z) of values of X and Z are observed. Hence, this table contains the conditional expectation values (true cell means) $E(Y\,|\,X=x, Z=z)$ of the outcome variable Y, and the joint probabilities $P(X=x, Z=z)$ determining the true joint distribution of X and Z.[2]

[2] In this context, 'true' just indicates that we are not referring to sample means or relative frequencies in a sample. Instead these are the true means around which sample means would vary.

Table 1.7. Conditional expectation values $E(Y|X{=}x, Z{=}z)$ given treatment and status

Treatment	Status						
	low $(Z{=}0)$		*med* $(Z{=}1)$		*hi* $(Z{=}2)$		
$X{=}0$	120	(20/120)	110	(17/120)	60	(3/120)	(40/120)
$X{=}1$	100	(7/120)	100	(26/120)	100	(7/120)	(40/120)
$X{=}2$	80	(3/120)	90	(17/120)	140	(20/120)	(40/120)
		(30/120)		(60/120)		(30/120)	

Note. Probabilities $P(X{=}x, Z{=}z)$, $P(Z{=}z)$, and $P(X{=}x)$ in parentheses.

In the *low status condition* $(Z = 0)$, there are large negative effects, both of treatment 1 and of treatment 2 compared to the control:

$$PFE_{Z;10}(0) = E(Y|X{=}1, Z{=}0) - E(Y|X{=}0, Z{=}0) = 100 - 120 = -20$$

and

$$PFE_{Z;20}(0) = E(Y|X{=}2, Z{=}0) - E(Y|X{=}0, Z{=}0) = 80 - 120 = -40.$$

In the *medium status condition* $(Z = 1)$, there are also negative effects of treatment 1 and of treatment 2 compared to the control:

$$PFE_{Z;10}(1) = E(Y|X{=}1, Z{=}1) - E(Y|X{=}0, Z{=}1) = 100 - 110 = -10$$

and

$$PFE_{Z;20}(1) = E(Y|X{=}2, Z{=}1) - E(Y|X{=}0, Z{=}1) = 90 - 110 = -20.$$

Finally, in the *high status condition* $(Z = 2)$, the effects of treatment 1 and treatment 2 are both positive:

$$PFE_{Z;10}(2) := E(Y|X{=}1, Z{=}2) - E(Y|X{=}0, Z{=}2) = 100 - 60 = 40$$

and

$$PFE_{Z;20}(2) := E(Y|X{=}2, Z{=}2) - E(Y|X{=}0, Z{=}2) = 140 - 60 = 80.$$

Based on these comparisons, we can conclude that the 'effects' of the treatments depend on the status of the subjects: the differences between the expectations of Y are negative for subjects with low and medium status, and they are positive for the subjects with high status.

1.2.3 Average of the $(Z{=}z)$-Conditional Prima Facie Effects

Now we consider the *average* of the $(Z{=}z)$-conditional prima facie effects, where Z is again the qualitative covariate *status*.[3] Because we already looked at the corresponding $(Z{=}z)$-conditional prima facie effects (see section 1.2.2), we just have to compute their averages,

[3] Note that we assume that Z is a random variable. In contrast, in analysis of variance it is assumed that Z is a fixed factor with a fixed number of observations for each value z of Z, that is, these numbers of observations are assumed to be invariant across different samples. In many empirical applications, this assumption is not realistic, but it does not invalidate the statistical conclusions as long as the parameters of interest do not involve the distribution of Z. However, a hypothesis about the average total effect *does* involve the distribution of Z if we the term 'average' is specified as an expectation value, and this is the reason why programs on analysis of variance usually are not able to correctly estimate and test hypotheses about average total effects. For more details see again Mayer and Thoemmes (2019).

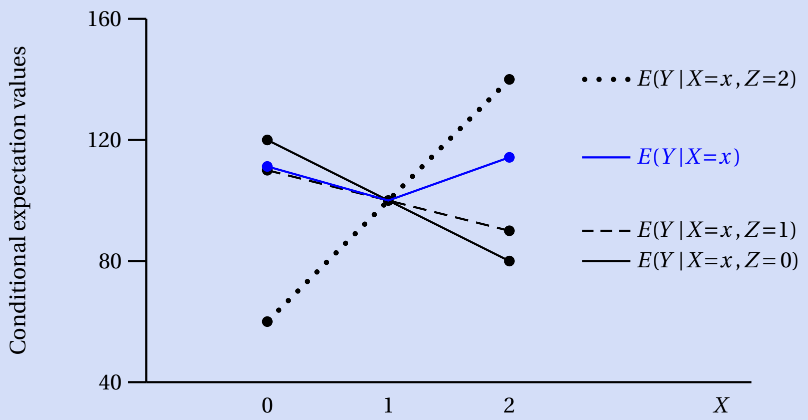

Figure 1.4. Conditional expectation values of Y given treatment and status

more precisely, the expectations of these conditional effects over the distribution of *status*:

$$E\big(PFE_{Z;10}(Z)\big) \;=\; \sum_z PFE_{Z;10}(z)\cdot P(Z=z) \;=\; (-20)\cdot\frac{1}{4}+(-10)\cdot\frac{1}{2}+40\cdot\frac{1}{4} \;=\; 0. \tag{1.9}$$

Hence, the average of the $(Z=z)$-conditional prima facie effects of treatment 1 compared to the control is 0.

Comparing treatment 2 to control yields the average of the $(Z=z)$-conditional prima facie effects:

$$E\big(PFE_{Z;20}(Z)\big) \;=\; \sum_z PFE_{Z;20}(z)\cdot P(Z=z) \;=\; (-40)\cdot\frac{1}{4}+(-20)\cdot\frac{1}{2}+80\cdot\frac{1}{4} \;=\; 0. \tag{1.10}$$

According to this result, the average effect of the $(Z=z)$-conditional prima facie effects of treatment 2 compared to the control is 0 as well.

1.2.4 Individual Effects

In this fictive example we can also look at the *individual effects* of treatment 1 compared to control and treatment 2 compared to control. These two effects can be read from Table 1.5 for each person. For example, for Tom the individual effect of treatment 1 compared to control is

$$ITE_{U;10}(Tom) \;=\; E(Y|X=1, Tom) - E(Y|X=0, Tom) \;=\; 100-120 \;=\; -20,$$

and his individual effect of treatment 2 compared to control is

$$ITE_{U;20}(Tom) \;=\; E(Y|X=1, Tom) - E(Y|X=0, Tom) \;=\; 80-120 \;=\; -40.$$

Correspondingly, for Joe the individual effect of treatment 1 compared to control is

$$ITE_{U;10}(Joe) \;=\; E(Y \,|\, X{=}1, Joe) - E(Y \,|\, X{=}0, Joe) \;=\; 90 - 90 \;=\; 0,$$

and his individual effect of treatment 2 compared to control is

$$ITE_{U;20}(Joe) \;=\; E(Y \,|\, X{=}1, Joe) - E(Y \,|\, X{=}0, Joe) \;=\; 70 - 90 \;=\; -20.$$

For the reasons mentioned in section 1.1.1, unlike the $(Z{=}z)$-conditional prima facie effects treated in section 1.2.2, the individual effects usually cannot be estimated in empirical applications. Nevertheless, they will play a crucial role in the definition of causal effects.

Of course, individual effects are more informative than their average if we want to know which treatment is the best for which individual. Nevertheless, we might ask: What are the *total individual treatment effects on average*? And, which are the $(Z{=}z)$*-conditional effects*, that is, the total individual treatment effects on average, given the value z of Z? Furthermore, if the total individual effects cannot be estimated in empirical applications under realistic assumptions, is it possible to estimate at least the average of the total individual treatment effects and/or the total individual treatment effects on average given the value z of Z? And if yes, under which conditions?

1.2.5 Average of the Individual Effects

Note that we have two averages of the individual effects in this example; we can compare treatment 1 to control and treatment 2 to control. Because we already looked at the corresponding individual effects, we just have to compute their averages, that is, the expectations of these conditional effects over the distribution of the person variable U, that is,

$$E\big(ITE_{U;10}(U)\big) \;=\; \sum_u ITE_{U;10}(u) \cdot P(U{=}u)$$

$$= (100-120)\cdot\frac{1}{8} + (100-120)\cdot\frac{1}{8} + (90-90)\cdot\frac{1}{8} + \ldots + (100-60)\cdot\frac{1}{8} \;=\; 0.$$

Hence, the average total effect of treatment 1 compared to the control is 0. Comparing treatment 2 to control yields

$$E\big(ITE_{U;20}(U)\big) \;=\; \sum_u ITE_{U;20}(u) \cdot P(U{=}u)$$

$$= (80-120)\cdot\frac{1}{8} + (80-120)\cdot\frac{1}{8} + (70-90)\cdot\frac{1}{8} + \ldots + (140-60)\cdot\frac{1}{8} \;=\; 0.$$

According to this result, the average total effect of treatment 2 compared to the control is 0 as well.

1.2.6 $(Z{=}z)$-Conditional Total Effects

Again, because we already know the individual effects, we just have to compute their averages given the value z of the covariate, or more precisely, the $(Z{=}z)$-conditional expectation values of the corresponding individual effects. We exemplify the computations for the value *med* of the covariate (second factor) Z. Comparing treatment 1 to control yields

$$E\big(ITE_{U;10}(U) \,\big|\, Z{=}med\big) \;=\; \sum_u ITE_{U;10}(u) \cdot P(U{=}u \,|\, Z{=}med)$$

$$= (90-90)\cdot\frac{1}{4} + (100-100)\cdot\frac{1}{4} + (100-120)\cdot\frac{1}{4} + (110-130)\cdot\frac{1}{4} \;=\; -10.$$

Hence, the $(Z{=}med)$-conditional total effect of treatment 1 compared to control is -10. Comparing treatment 2 to control we obtain

$$E\big(ITE_{U;20}(U)\,\big|\,Z{=}med\big) \;=\; \sum_u ITE_{U;20}(u)\cdot P(U{=}u\,|\,Z{=}med)$$
$$= (70-90)\cdot\frac{1}{4} + (80-100)\cdot\frac{1}{4} + (100-100)\cdot\frac{1}{4} + (110-110)\cdot\frac{1}{4} \;=\; -10.$$

Hence, the $(Z{=}med)$-conditional total effect of treatment 2 compared to control is -10 as well.

The computations for the other two values of Z are analogous. For $Z{=}low$ we obtain $E\big(ITE_{U;10}(U)\,\big|\,Z{=}low\big) = -20$ and $E\big(ITE_{U;20}(U)\,\big|\,Z{=}low\big) = -40$. In contrast, for $Z{=}hi$ we obtain $E\big(ITE_{U;10}(U)\,\big|\,Z{=}hi\big) = -40$ and $E\big(ITE_{U;20}(U)\,\big|\,Z{=}hi\big) = -80$.

1.2.7 How to Evaluate the Treatment?

To summarize, we discussed several ways that may, at first sight, be used to evaluate the treatment effects in empirical applications: *First*, we may compare the differences between the conditional expectation values $E(Y\,|\,X{=}x)$ of the outcome variable in the three treatment conditions $x \in \{0,1,2\}$. *Second*, we may consider the corresponding differences between the conditional expectation values $E(Y\,|\,X{=}x, Z{=}z)$ *given each of the three values* $z \in \{low, med, hi\}$ *of status*. *Third*, we may compare the expectations of these differences between the $(X{=}x, Z{=}z)$-conditional expectation values over the distribution of Z. All these comparisons yield different results. Which of them are meaningful for the evaluation of the treatment effects? All three of them, or only two, just one, or none at all? And which are the conditions under which they are meaningful?

Furthermore, we also presented three parameters based on individual effects that may be used to evaluate the treatment effects: *First*, the *individual total effect* of a treatment compared to a control. These effects are hard to estimate in empirical applications, unless we introduce very restrictive assumptions. *Second*, the expectation of these individual effects, which in this example, might also be called *average total effects*. This kind of effect is less informative than the individual total effects, but it is a summary parameter that informs us if the treatment is beneficial, ineffective, or even harmful on average. *Third*, the conditional expectation values of the individual total effects, which in this example, might also be called the $(Z{=}z)$-conditional total treatment effects. They inform us with a single number for each value z if the treatment is beneficial, ineffective, or even harmful on average for those individuals with value z on the covariate Z. Box 1.1 provides a summary of these effects.

In this example, the averages of the $(Z{=}z)$-conditional prima facie effects are identical to the averages of the individual total effects. Is this just a coincidence? Or is this due to systematic conditions that hold in this example? If yes, which are these conditions?

1.3 Summary and Conclusion

In this chapter, we treated two examples. In the first one, a dichotomous treatment variable X has a negative (prima facie) effect $P(Y{=}1\,|\,X{=}1) - P(Y{=}1\,|\,X{=}0)$ on a *dichotomous outcome variable* Y ('success'), although the corresponding individual treatment effects

$$P(Y=1|X=1,U=u)-P(Y=1|X=0,U=u)$$

are positive. Taking the expectation of these two individual effects also yields a positive effect.

In the second example, there are nonzero differences $E(Y|X=1) - E(Y|X=0)$ and $E(Y|X=2) - E(Y|X=0)$, where Y is a *quantitative outcome variable*, and nonzero conditional 'effects' $E(Y|X=1,Z=z) - E(Y|X=0,Z=z)$ and $E(Y|X=2,Z=z) - E(Y|X=0,Z=z)$ for the different values z of *status*. The expectations of these $(Z=z)$-conditional prima facie effects (comparing treatment 1 to 0 and comparing treatment 2 to 0) over the three status conditions are zero.

The Problem

Box 1.1 displays the various total effects that have been computed and discussed in the two examples. Because the conclusions drawn from each of these putative effects are contradictory, which of these should we trust? In the first example: Is the treatment harmful — as the difference $P(Y=1|X=1) - P(Y=1|X=0)$ suggests? Or is it beneficial as suggested by the individual effects $P(Y=1|X=1,U=u) - P(Y=1|X=0,U=u)$? In the second example: Do the prima facie effects $E(Y|X=1) - E(Y|X=0)$ have a meaningful causal interpretation? Or do the $(Z=z)$-conditional prima facie effects $E(Y|X=1,Z=z) - E(Y|X=0,Z=z)$ have a meaningful causal interpretation? And, does this apply also to their expectation?

In the first example, we demonstrated that we cannot expect that the difference

$$P(Y=1|X=1)-P(Y=1|X=0)$$

is the average (expectation) of the corresponding person-specific differences

$$P(Y=1|X=1,U=u)-P(Y=1|X=0,U=u)\,.$$

Similarly, in the second example, we showed that we can not expect that the difference

$$E(Y|X=1)-E(Y|X=0)$$

is the average (expectation) of the corresponding differences

$$E(Y|X=1,Z=z)-E(Y|X=0,Z=z)$$

given a value z of status. And, how do we know that these $(Z=z)$-conditional effects are meaningful for the evaluation of the treatment? As noted before, these questions are not related to *statistical* inference; they are not raised at the sample level, but on the level of true conditional expectation values!

Hence our examples show that conditional expectation values and their differences, the prima facie effects, can be totally misleading in evaluating the effects of a treatment variable X on an outcome variable Y. This conclusion can also be extended to conditional probabilities, to correlations and to all other parameters describing relationships and dependencies between random variables. They all are like the shadow in the metaphor of the invisible man (see the preface).

If this is true, is the whole idea of *learning from experience* — the core of empirical sciences — wrong? Our answer is 'No'. However, we have to be more explicit in what we

Box 1.1 Various effects treated in this chapter

$PFE_{xx'}$	*Prima facie effect* of treatment x compared to treatment x'. It is defined by $$PFE_{xx'} := E(Y \mid X{=}x) - E(Y \mid X{=}x').$$
$PFE_{Z;xx'}(z)$	$(Z{=}z)$-*conditional prima facie effect* of treatment x compared to treatment x'. It is defined by $$PFE_{Z;xx'}(z) := E(Y \mid X{=}x, Z{=}z) - E(Y \mid X{=}x', Z{=}z).$$
$E\big(PFE_{Z;xx'}(Z)\big)$	*Expectation of the* $(Z{=}z)$-*conditional prima facie effects* of treatment x compared to treatment x'. If Z is discrete, then it is computed by $$E\big(PFE_{Z;xx'}(Z)\big) := \sum_z PFE_{Z;xx'}(z) \cdot P(Z{=}z).$$
$ITE_{U;xx'}(u)$	*Individual effect* of treatment x compared to treatment x'. It is defined by $$ITE_{U;xx'}(u) := E(Y \mid X{=}x, U{=}u) - E(Y \mid X{=}x', U{=}u).$$
$E\big(ITE_{U;xx'}(U)\big)$	*Expectation of the individual effects* of treatment x compared to treatment x'. It is computed by $$E\big(ITE_{U;xx'}(U)\big) := \sum_u ITE_{U;xx'}(u) \cdot P(U{=}u).$$
$E\big(ITE_{U;xx'}(U) \,\big\vert\, Z{=}z\big)$	$(Z{=}z)$-*conditional expectation value of the individual effects* of treatment x compared to treatment x'. It is computed by $$E\big(ITE_{U;xx'}(U) \,\big\vert\, Z{=}z\big) := \sum_u ITE_{U;xx'}(u) \cdot P(U{=}u \mid Z{=}z).$$

mean by terms like 'X affects Y', 'X has an effect on Y', 'X influences Y', 'X leads to Y', and so on used in our theories and hypotheses. How can these terms be uniquely defined in a language that is compatible with statistical analyses of empirical data? How to design an empirical study and how to look at the resulting data if we want to probe our theories and learn about the causal effects postulated in these theories and hypotheses?

In the chapters to come we will show that these parameters are meaningful *under certain conditions*, just like the shadow of the invisible man can be meaningful under certain conditions in order to measure his height. In the metaphor a crucial condition is the 45° degree angle of the sun. Do we also have such a crucial condition for causal inference? We know that a reversal of total effects does not occur in randomized experiments, that is, in experiments in which observational units (in the social and behavioral sciences, usually the subjects or individuals) are randomly assigned to one of at least two treatment conditions. In the randomized experiment comparing conditional expectation values *is* informative about total causal treatment effects. But why? What is so special in the randomized experiment? Which are the mathematical conditions that we create in a randomized experiment? Are there also conditions that can be utilized in quasi-experimental evaluation studies? How can we estimate causal effects in quasi-experimental observational studies? Conclusive answers to these questions can be hoped for only within a theory of causal effects.

Relevance of the Problem

Obviously, these questions are of fundamental importance for the methodology of empirical sciences and for the empirical sciences themselves. The answers to these questions have consequences for the design and analysis of experiments, quasi-experiments, and other studies aiming at estimating the effects of *treatments*, *interventions*, or *expositions*. No *prevention study* can meaningfully be conducted and analyzed without knowing the concepts of causal effects and how they can be estimated from empirical data. Similarly, without a clear concept of causal effects we are not able to learn from our data about the effects of a certain (possibly harmful) environment on our health, or about the effects of certain behaviors such as smoking or drug abuse. Again, this is similar to the problem of measuring the invisible man's size via the length of his shadow: only with a clear concept of *size*, some basic knowledge in geometry, and the additional information such as the angle of the sun at the time of measurement are we able to determine his size of the man from the length of his shadow.

Research Traditions

Of course, raising these questions and attempting answers is not new. Immense knowledge and wisdom about experiments and quasi-experiments has been collected in the Campbellian tradition of experiments and quasi-experiments (see, e. g., Campbell & Stanley, 1963; Cook & Campbell, 1979; Shadish et al., 2002). In the last decades, a more formal approach has been developed supplementing the Campbellian theory and terminology in important aspects: the theory of causal effects in the Neyman-Rubin tradition (see, e. g., Splawa-Neyman, 1923/1990; Rubin, 1974, 2005). Many papers and books indicate the growing influence of this theory (see, e. g., Greenland, 2000, 2004; Höfler, 2005; Rosenbaum, 2002a; Rubin, 2006; Winship & Morgan, 1999; Morgan & Winship, 2007) and remarkable efforts have already been made to integrate it into the Campbellian framework (West, Biesanz, & Pitts, 2000). Furthermore, these questions have also been dealt with in the graphical modeling tradition (see, e. g., Pearl, 2009; Spirtes et al., 2000) as well as in biometrics, econometrics, psychometrics, epidemiology, and other fields dealing with the methodology of empirical research.

Outlook

In this volume, we present the theory of causal total effects in terms of mathematical probability theory. We show that a number of questions that have been debated controversially and inconclusively can now be given a clear-cut answer. What kinds of causal effects can meaningfully be defined? Which design techniques allow for unbiased estimation of causal effects? How to analyze nonorthogonal ANOVA designs (cf., e. g., Aitkin, 1978; Appelbaum & Cramer, 1974; Gosslee & Lucas, 1965; Maxwell & Delaney, 2004; Overall et al., 1975)? How to analyze non-equivalent control-group designs (cf., e. g., Reichardt, 1979)? Should we compare pre-post differences between treatment groups (cf., e. g., Lord, 1967; Senn, 2006; van Breukelen, 2006; Wainer, 1991)? Should we use analysis of covariance to adjust for differences between treatment and control that already existed prior to treatment (cf., e. g., Maxwell & Delaney, 2004; Cohen, Cohen, West, & Aiken, 2003)? Should we use propensity score methods instead of the more traditional procedures mentioned above (cf., e. g., Rosenbaum & Rubin, 1983b)? How do we deal with non-compliance to treatment assignment (cf., e. g., Cheng & Small, 2006; Dunn et al., 2003; Jo, 2002a, 2002b, 2002c; Jo, Asparouhov, Muthén, Ialongo, & Brown, 2008; J. Robins & Rotnitzky, 2004; J. M. Robins, 1998)?

We do not treat the statistical sampling models with their distributional assumptions, their implications for parameter estimation, and the evaluation (or tests) of hypotheses about these parameters. However, we will discuss the virtues and problems of general strategies of data analysis such as the analysis of difference scores, analysis of covariance, its generalizations, and analysis based on propensity scores.

1.4 Exercises

▷ **Exercise 1-1** Why do we need the concept of a causal treatment effect?

▷ **Exercise 1-2** What is the relationship between the unconditional prima facie effect PFE_{10} and the expectations $E(Y|X=0)$ and $E(Y|X=1)$ of the outcome variable Y in the two treatment conditions?

▷ **Exercise 1-3** Compute the probabilities $P(X=x, Y=y)$ presented in Table 1.3 from the probabilities $P(U=u, X=x, Y=y)$ presented in Table 1.4.

▷ **Exercise 1-4** Which are the kinds of prima facie effects treated in this chapter?

▷ **Exercise 1-5** What is the difference between statistical inference and causal inference?

▷ **Exercise 1-6** Why are the conditional expectation values $E(Y|X=x)$ in treatment conditions x also the $(X=x)$-conditional probabilities for the event $\{Y=1\}$ in the first example treated in this chapter?

▷ **Exercise 1-7** Download *Kbook Table 1.1.sav* from *www.causal-effects.de*. This data set has been generated from Table 1.1 for a sample of size $N = 10,000$. Compute the contingency table corresponding to Table 1.3 and the associated estimates of the conditional probabilities $P(Y=1|X=0)$ and $P(Y=1|X=1)$.

▷ **Exercise 1-8** Use $P(Y=1|X=1) = .42$ and $P(Y=1|X=0) = .6$ computed in section 1.1.3 in order to compute the corresponding odds ratio, its logarithm, and the risk ratio, according to the definitions of these parameters presented in SN-sect. 13.3.

▷ **Exercise 1-9** Compute the conditional probability $P(Y=1|X=1, U=Joe)$ from Table 1.4.

▷ **Exercise 1-10** Compute the probability $P(Y=1|X=0)$ from the corresponding conditional probabilities $P(Y=1|X=0, U=u)$.

▷ **Exercise 1-11** What (i. e., how big) are the unconditional prima facie effects of the treatments, that is, the prima facie effects $E(Y|X=1) - E(Y|X=0)$ and $E(Y|X=2) - E(Y|X=0)$ in the second example of this chapter?

▷ **Exercise 1-12** What are the conditional prima facie effects of the treatments, that is, the prima facie effects $E(Y|X=1, Z=z) - E(Y|X=0, Z=z)$ and $E(Y|X=2, Z=z) - E(Y|X=0, Z=z)$ in the second example of this chapter?

▷ **Exercise 1-13** What are the averages of the conditional prima facie effects

$$E(Y|X=1, Z=z) - E(Y|X=0, Z=z) \qquad \text{and} \qquad E(Y|X=2, Z=z) - E(Y|X=0, Z=z)$$

in the second example of this chapter?

▷ **Exercise 1-14** Compute the conditional probability $P(U=\mathit{Tom} \mid X=0)$ from the parameters presented in Tables 1.5 and 1.6.

▷ **Exercise 1-15** Open the *Causal Effects Xplorer* with table *K-book table 1.1.tab*. Change the conditional probabilities $P(X=1|U=u)$ of receiving treatment 1 for Joe and Ann to $2/5$. Then compare the two individual treatment effects of Joe and Ann and their average to the prima facie effect $E(Y|X=1) - E(Y|X=0)$.

▷ **Exercise 1-16** Open the *Causal Effects Xplorer* with table *K-book Table 1.1.tab* displaying the conditional probabilities $P(U=u|X=x)$. Then use RS-Box 3.2 (ii) in order to compute the three conditional expectation values $E(Y|X=x)$ displayed in Table 1.6 from the parameters presented in Table 1.5.

▷ **Exercise 1-17** Download *Kbook Table 1.5.sav* from *www.causal-effects.de*. This data set has been generated with the Causal Effects Xplorer from Table 1.5 for a sample of size $N = 10,000$ with error variance 10 given each person.

(a) Compute the cell means and the relative frequencies of observations in each of the nine cells of the (3×3)-table.

(b) Use each of the procedures offered by your statistical program package to analyze the data including a test of the main effects of the treatment factor (most programs offer Typ I, II and III sums of squares for such an analysis).

(c) Compare the results of these analyses to the parameters presented in Table 1.7.

Solutions

▷ **Solution 1-1** We need the concept of a causal treatment effect, because the two examples show that differences between conditional expectation values are meaningless for the evaluation of the effects of a treatment, unless we can show how the differences between these conditional expectation values are related to the causal treatment effects. Obviously, without a definition of causal treatment effects this is not possible. Estimating causal treatment effects is crucial for answering questions such as 'Does the treatment help our patients with respect to the outcome variable considered?'

▷ **Solution 1-2** The unconditional prima facie effect PFE_{10} is defined as the difference between the two conditional expectation values $E(Y|X=1)$ and $E(Y|X=0)$.

▷ **Solution 1-3** This can easily be verified by adding the probabilities for the observations of the pairs (x,z) of X and Z over males and females. This yields $.144 + .096 = .24$, $.004 + .228 = .232$, $.336 + .024 = .36$ and $.016 + .152 = .168$.

▷ **Solution 1-4** The kinds of prima facie effects treated in this chapter are: the *unconditional prima facie effect*, the *conditional prima facie effect* given the value z of a covariate Z, and the *average of the* $(Z=z)$*-conditional prima facie effects*. The unconditional prima facie effect of treatment 1 compared to treatment 0 is the difference $PFE_{10} = E(Y|X{=}1) - E(Y|X{=}0)$ between the conditional expectation values of an outcome variable Y given the two treatment conditions. The $(Z{=}z)$-conditional prima facie effect is the difference $PFE_{Z;10}(z) = E(Y|X{=}1, Z{=}z) - E(Y|X{=}0, Z{=}z)$ between the $(X{=}1, Z{=}z)$-conditional expectation value and the $(X{=}0, Z{=}z)$-conditional expectation value of the outcome variable Y. The average prima facie effect is the expectation of the $(Z{=}z)$-conditional prima facie effects over the distribution of Z [see Eqs. (1.9) and (1.10)].

▷ **Solution 1-5** In *statistical* inference we estimate and test hypotheses about parameters characterizing the (joint or marginal) distributions of random variables from sample data. In *causal* inference we interpret some of these parameters as causal effects, provided that certain conditions are satisfied that allow for such a causal interpretation.

▷ **Solution 1-6** $E(Y|X{=}x) = P(Y{=}1|X{=}x)$, because, in this example, Y is dichotomous with values 0 and 1. In this case, the term $P(Y{=}1|X{=}x)$ is defined by $E(Y|X{=}x)$ (see RS-Rem. 3.22).

▷ **Solution 1-7** No solution provided. Just compare your results to the true parameters presented in Table 1.3 and to the conditional probabilities $P(Y{=}1|X{=}0)$ and $P(Y{=}1|X{=}1)$ presented in section 1.1.3.

▷ **Solution 1-8** The *odds ratio* is

$$\frac{P(Y{=}1|X{=}1)}{1-P(Y{=}1|X{=}1)} \bigg/ \frac{P(Y{=}1|X{=}0)}{1-P(Y{=}1|X{=}0)} \approx .483.$$

Because this number is smaller than 1 it indicates that there is a negative effect of the treatment. The natural logarithm of the odds ratio is the *log odds ratio*, which is

$$\ln\left[\frac{P(Y{=}1|X{=}1)}{1-P(Y{=}1|X{=}1)} \bigg/ \frac{P(Y{=}1|X{=}0)}{1-P(Y{=}1|X{=}0)}\right] \approx -0.728.$$

This number is smaller than 0 indicating that there is a negative effect of the treatment. The log odds ratio is identical to the *logistic regression coefficient* λ_1 in the equation

$$P(Y{=}1|X) = \frac{\exp(\lambda_0 + \lambda_1 \cdot X)}{1+\exp(\lambda_0 + \lambda_1 \cdot X)}.$$

Another closely related parameter is the *risk ratio*

$$\frac{P(Y{=}1|X{=}1)}{P(Y{=}1|X{=}0)} = .7.$$

Because this number is smaller than 1, it indicates that there is a negative effect of the treatment. Hence, no matter which of these parameters we use, we would always come to the same (wrong) conclusion that the treatment is detrimental for our patients.

▷ **Solution 1-9** Using the joint probabilities presented in Table 1.4, the definition of the conditional probability yields

$$P(Y{=}1|X{=}1, U{=}\mathit{Joe}) = \frac{P(X{=}1, Y{=}1, U{=}\mathit{Joe})}{P(X{=}1, U{=}\mathit{Joe})} = \frac{.016}{.016+.004} = .8.$$

▷ **Solution 1-10** First of all, note that the theorem of total probability (see RS-Th. 1.38), can also be applied to conditional probabilities (see RS-Th. 1.42). In this exercise, it is applied to the $(X{=}0)$-conditional probabilities $P(Y{=}1|X{=}0) = P^{X=0}(Y{=}1)$. Hence, according to this theorem,

$$\begin{aligned} P(Y{=}1|X{=}0) &= P(Y{=}1|X{=}0, U{=}\mathit{Joe}) \cdot P(U{=}\mathit{Joe}\,|X{=}0) + \\ &\quad P(Y{=}1|X{=}0, U{=}\mathit{Ann}) \cdot P(U{=}\mathit{Ann}\,|X{=}0). \end{aligned}$$

The probabilities $P(Y{=}1|X{=}0, U{=}\mathit{Joe}) = .7$ and $P(Y{=}1|X{=}0, U{=}\mathit{Ann}) = .2$ are computed analogously to Exercise 1-9 and the other two probabilities occurring in this formula are $P(U{=}\mathit{Joe}\,|X{=}0) = .48/.6$ and $P(U{=}\mathit{Ann}|X{=}0) = .12/.6$ (see Table 1.4). Hence,

$$P(Y{=}1|X{=}0) = \frac{.7 \cdot .48}{.48+.12} + \frac{.2 \cdot .12}{.48+.12} = .6.$$

▷ **Solution 1-11** The prima facie effects $E(Y|X=1)-E(Y|X=0)$ and $E(Y|X=2)-E(Y|X=0)$ can be computed from Table 1.6 as follows:

$$PFE_{10} \;=\; E(Y|X=1)-E(Y|X=0) \;=\; 100.00-111.25 \;=\; -11.25$$

and

$$PFE_{20} \;=\; E(Y|X=2)-E(Y|X=0) \;=\; 114.25-111.25 \;=\; 3.00.$$

▷ **Solution 1-12** The conditional prima facie effects

$$E(Y|X=1, Z=z)-E(Y|X=0, Z=z) \quad \text{and} \quad E(Y|X=2, Z=z)-E(Y|X=0, Z=z)$$

can be computed from Table 1.7. For *low status* ($Z=low$), they are:

$$PFE_{Z;10}(low) \;=\; E(Y|X=1, Z=low)-E(Y|X=0, Z=low) \;=\; 100-120 \;=\; -20$$

$$PFE_{Z;20}(low) \;=\; E(Y|X=2, Z=low)-E(Y|X=0, Z=low) \;=\; 80-120 \;=\; -40.$$

For *medium status* ($Z=med$), they are:

$$PFE_{Z;10}(med) \;=\; E(Y|X=1, Z=med)-E(Y|X=0, Z=med) \;=\; 100-110 \;=\; -10$$

$$PFE_{Z;20}(med) \;=\; E(Y|X=2, Z=med)-E(Y|X=0, Z=med) \;=\; 90-110 \;=\; -20.$$

Finally, for *high status* ($Z=hi$), the conditional prima facie effects are:

$$PFE_{Z;10}(hi) \;=\; E(Y|X=1, Z=hi)-E(Y|X=0, Z=hi) \;=\; 100-60 \;=\; 40$$

$$PFE_{Z;20}(hi) \;=\; E(Y|X=2, Z=hi)-E(Y|X=0, Z=hi) \;=\; 140-60 \;=\; 80.$$

▷ **Solution 1-13** Using the results of the last exercise, the average of the ($Z=z$)-conditional prima facie effects can be computed from the conditional effects as follows:

$$\begin{aligned}
E\big(PFE_{Z;10}(Z)\big) \;&=\; PFE_{Z;10}(low)\cdot P(Z=low)+PFE_{Z;10}(med)\cdot P(Z=med)+PFE_{Z;10}(hi)\cdot P(Z=hi)\\
&=\; (-20)\cdot\frac{1}{4}+(-10)\cdot\frac{1}{2}+40\cdot\frac{1}{4} \;=\; 0.
\end{aligned}$$

$$\begin{aligned}
E\big(PFE_{Z;20}(Z)\big) \;&=\; PFE_{Z;20}(low)\cdot P(Z=low)+PFE_{Z;20}(med)\cdot P(Z=med)+PFE_{Z;20}(hi)\cdot P(Z=hi)\\
&=\; (-40)\cdot\frac{1}{4}+(-20)\cdot\frac{1}{2}+80\cdot\frac{1}{4} \;=\; 0.
\end{aligned}$$

▷ **Solution 1-14**

$$\begin{aligned}
P(U=Tom\,|\,X=0) \;&=\; \frac{P(U=Tom\,, X=0)}{P(X=0)} \;=\; \frac{P(X=0\,|\,U=Tom)\cdot P(U=Tom)}{P(X=0)}\\
&=\; \frac{(47/60)\cdot(1/8)}{1/3} \;=\; \frac{47}{160}.
\end{aligned}$$

▷ **Solution 1-15** With this change, the prima facie effect changes to $E(Y|X=1)-E(Y|X=0)=.6-.45=.15$, which is the average of the two individual total effects, which still are .10 for Joe and .20 for Ann. Note that identical treatment probabilities $P(X=1\,|\,U=u)$ for all persons u is what we create by randomly assigning a person to treatment 1 in a randomized experiment.

▷ **Solution 1-16** Rewriting RS-Box 3.2 (ii) for our example,

$$E(Y|X=x) \;=\; \sum_{u} E(Y|X=x, U=u)\cdot P(U=u\,|\,X=x).$$

Hence,

$$\begin{aligned} E(Y \mid X=0) &= \sum_u E(Y \mid X=0, U=u) \cdot P(U=u \mid X=0) \\ &= 120 \cdot \frac{47}{160} + 120 \cdot \frac{33}{160} + \ldots + 60 \cdot \frac{8}{160} = 111.25, \end{aligned}$$

$$\begin{aligned} E(Y \mid X=1) &= \sum_u E(Y \mid X=1, U=u) \cdot P(U=u \mid X=1) \\ &= 100 \cdot \frac{10}{160} + 100 \cdot \frac{18}{160} + \ldots + 100 \cdot \frac{16}{160} = 100, \end{aligned}$$

and

$$\begin{aligned} E(Y \mid X=2) &= \sum_u E(Y \mid X=2, U=u) \cdot P(U=u \mid X=2) \\ &= 80 \cdot \frac{3}{160} + 80 \cdot \frac{9}{160} + \ldots + 140 \cdot \frac{36}{160} = 114.25. \end{aligned}$$

▷ **Solution 1-17** No solution provided. Just compare your results to the parameters presented in Table 1.7.

Chapter 2
Some Typical Kinds of Random Experiments

In chapter 1 we have seen that comparing conditional expectation values of an outcome variable between treatment groups can be completely misleading if used for the evaluation of treatment effects. In this chapter we continue preparing the stage for the theory of causal total effects and dependencies, describing the kind of empirical phenomena it refers to: single-unit trials of experiments or quasi-experiments, but also single-unit trials of observational studies in which causal total effects and dependencies can be investigated. First examples of such a single-unit trial have already been treated in sections 1.1 and 1.2 (see in particular Tables 1.2 and 1.5).

A *single-unit trial* is a specific random experiment. Note the distinction between a *random experiment* and a *randomized experiment*. Stochastic dependencies between events and between random variables always refer to a random experiment, but not necessarily to a *randomized experiment* in which the assignment of a subject to one of the treatment conditions is determined exclusively by a random procedure such as flipping a coin. In contrast, a *random experiment* is the concrete empirical phenomenon to which stochastic dependencies between events and random variables (described by conditional distributions, probabilities, correlations, and conditional expectations) refer to.

The single-unit trial *is not the sample* dealt with in statistical models. In a sample, we consider repeating the single-unit trial many times in one way or another. This is necessary if we want to deal with estimation of parameters and tests of hypotheses about these parameters, some of which might be causal effects. The single-unit trial does *not allow* treating problems of parameter estimation or hypothesis testing. However, it is sufficient for defining causal effects and studying how to identify them, that is, investigating under which conditions and how they can be computed from empirically estimable parameters.

A single-unit trial is also what we refer to in hypotheses and theories of the empirical sciences. In many text books on applied statistics the dazzling term ‘population’ is used instead, obfuscating what we are actually talking about when we use probabilistic terms such as expectation, variance, covariance, correlation, regression, etc. Furthermore, single-unit trials are what is of interest in practical work. How does the treatment of a patient affect the outcome of this patient if compared to another possible treatment? What is the treatment effect for a male, and what is its effect for a female? Which variables explain inter-individual differences in individual causal total effects? All these practical questions are raised using concepts referring to a single-unit trial.

Overview

We start with the single-unit trial of simple experiments and then treat increasingly more complex ones introducing additional design features. Specifically, we will introduce the single-unit trials of experiments and quasi-experiments with fallible covariates, a multifactorial design with more than one treatment, multilevel experiments and quasi-experiments, and experiments and quasi-experiments with latent covariates and/or outcome variables.

We also discuss different kinds of random variables that will play a crucial role in the chapters to come. Among these random variables are the observational-unit variable or person variable, manifest and latent covariates, treatment variables, as well as manifest and latent outcome variables. In this chapter, we confine ourselves to an informal description of single-unit trials and the random variables involved, preparing the stage for their mathematical representations in the subsequent chapters.

2.1 Simple Experiments

As a first class of random experiments we consider the single-unit trials of *simple experiments and quasi-experiments.* Such single-unit trials are experiments and quasi-experiments in which *no fallible covariates* are assessed, that is, no covariates whose values consist in part of a measurement error. Such a single-unit trial consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,

(b) assigning the unit or observing its assignment to one of several experimental conditions (represented by the value x of the treatment variable X),

(c) recording the value y of the outcome (or response) variable Y.

Figure 2.1 displays a tree representation of the set of possible outcomes of this single-unit trial. Note that this is the kind of random experiment we considered in the Joe-Ann example presented in section 1.1 and in the two-factorial design example treated in section 1.2. The random variables X (treatment), Y (success), and Z (status), the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x,Z{=}z)$, as well as the probabilities $P(X{=}x)$, $P(Z{=}z)$, $P(X{=}x,Z{=}z)$ all referred to such a single-unit trial. Of course, all these conditional expectation values and probabilities are unknown in empirical applications. Nevertheless, they are among the parameters that determine the outcome of a single-unit trial, just in the same way as the probability of *heads* determines the outcome of flipping a coin.

In order to illustrate this point, imagine flipping a deformed coin that has the shape of a Chinese wok, and suppose that in this case the probability of flipping *heads* is .8 instead of .5. Although this probability does not deterministically determine the outcome of flipping the coin, it stochastically determines the outcome.

In fact, we may consider the single-unit trial of (a) sampling a coin u from a set of coins, (b) forming ($X{=}1$) or not forming ($X{=}0$) a wok out of it, and (c) observing whether ($Y{=}1$) or not ($Y{=}0$) we flip *heads.* In this single-unit trial, the difference .8 − .5 = .3 would be the causal total effect of the treatment variable X on the outcome variable Y. Note that the probabilities .8 and .5 and their difference .3 refer to this single-unit trial, although these probabilities can only be estimated if we conduct many of these single-unit trials, that is, if we draw a data sample. However, if these probabilities were known, we could dispense with a sample (including the data that would result from drawing it), and still have a perfect theory and prediction for the outcome of such a single-unit trial (see Exercise 2-1).

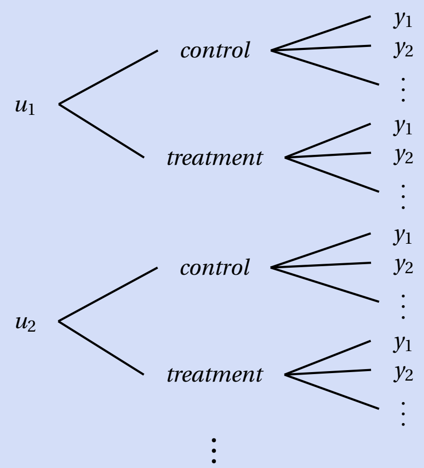

Figure 2.1. A simple experiment or quasi-experiment

Sampling a Unit

The first part of this single-unit trial consists of sampling an observational unit. In the social sciences, units often are persons, but they might be groups, school classes, schools and even countries. Usually such units change over time. Therefore, it should be emphasized that, in simple experiments and quasi-experiments, we are talking about the *units at the onset of treatment*. Later we will see that we have to distinguish between *units at the onset of treatment* and *units at the time of assessment of the outcome variable*, which might be months or even years later (for details see Steyer, Mayer, Geiser, & Cole, 2015). In a single-unit trial of simple experiments and quasi-experiments, the units can be represented by the observational-unit variable U, whose possible values u are the *units at the onset of treatment*.

Note that the unit at the onset of treatment also comprises his or her experiences a year and/or the day before treatment, as well as the psycho-bio-social situation in which he or she is at the onset of treatment. Both, the experiences and the situation, already happen *before* the onset of treatment (see again Steyer et al., 2015 for more details). Therefore, they are attributes of the observational units u. They can be treated in the same way as other attributes such as sex and educational status. However, if these attributes are actually assessed and if this assessment is fallible, then we have to distinguish between these attributes and their fallible assessments (see sect. 2.2).

Treatment Variable

In an experiment or quasi-experiment, there is always a treatment variable, which we focus as a cause[1] and usually denote it by X. In a *true experiment*, the unit drawn is assigned

[1] We use the term (putative) cause for a random variable if we consider its causal effect on an outcome variable. Note that a causal effect can also be 0.

to one of the possible treatments with a probability that is fixed by the experimenter. In contrast, in a *quasi-experiment* we just observe selection (e. g., self-selection or selection by someone else) to one of the treatment conditions. In the simplest case there are at least two treatment conditions, for example, *treatment* and *control*. These treatment conditions are the possible values of the treatment variable X. For simplicity, we use the values $0, 1, \dots, J$ to represent $J+1$ treatment conditions. Furthermore, unless stated otherwise, we presume that treatment *assignment* and actual *exposure* to treatment are equivalent, that is, we assume that there is perfect compliance.

Selection of a unit into one of the treatment conditions x may happen with unknown probabilities. This is the case, for example, in self-selection or assignment by an unknown physician. In this case we talk about a quasi-experiment. However, assignment can also be done with known probabilities that are equal for different units such as in the simple randomized experiment or with known probabilities that may be unequal for different units such as in the *conditionally randomized experiment*. In this case, these treatment probabilities may also depend on a covariate Z representing pre-treatment attributes of the units. As mentioned above, *conditional and unconditional randomized assignment, distinguish the true experiment from the quasi-experiment*, in which the assignment probabilities are unknown. (See Remarks 8.58 and 8.59 for more details on randomization and conditional randomization.)

Potential Confounders and Covariates

In simple experiments and quasi-experiments, the focus is usually on total treatment effects on an outcome variable. Hence, if we are interested in the treatment variable as a cause, then each attribute of the observational units is a *potential confounder*. Examples are *sex*, *race*, *educational status*, and *socio-economic status*. Once the unit is drawn, its *sex*, *race*, *educational status*, and *socio-economic status* are fixed. This means that there is no additional sampling process associated with assessing these potential confounders. This is also the reason why they do not appear in points (a) to (c) describing the single-unit trial.

A potential confounder is also called a *covariate* if it is actually assessed and used together with X in a conditional expectation or a conditional distribution. Note that a potential confounder can also be unobserved, and in this case we usually do not call it a covariate.

Because potential confounders represent attributes of the unit *at the onset of treatment* they can never be affected by the treatment. However, there can be (stochastic) dependencies between the treatment variable and potential confounders. In the Joe-Ann example treated in sect. 1.1, for instance, there is a stochastic dependence between the treatment variable X and the person variable U. Similarly, in the second example presented in section 1.2 there is a stochastic dependence between Z (*status*) and the treatment variable X.

Multidimensional Potential Confounders and Covariates

Potential confounders – and therefore also covariates – may be uni- or multi-dimensional, qualitative (such as $Z_1 :=$*sex* and $Z_2 :=$ *educational status*) or quantitative (such as $Z_3 :=$ *height* and $Z_4 :=$*body mass index*). If it is a multivariate variable made up of several uni-dimensional variables, it may consist of qualitative *and* quantitative potential confounders such as $Z_5 = (Z_1, Z_4)$.

Specific Potential Confounders

Note that the observational-unit variable U and the U-conditional treatment probability $P(X{=}x \mid U)$ (see RS-Def. 4.4) are potential confounders as well. The values of $P(X{=}x \mid U)$ are the conditional probabilities $P(X{=}x \mid U{=}u)$, which are attributes of the persons u [see Eq. (1.1)]. Similarly, the Z-conditional treatment probability $P(X{=}x \mid Z)$ is also a potential confounder provided that Z is a covariate [see Def. 4.11 (iv) and Rem. 4.16]. Furthermore, the *assignment* to treatment x with values '*yes*' and '*no*' is also a potential confounder if *assignment to treatment* and *exposure to treatment* (again with values '*yes*' and '*no*') are not identical and *exposure to treatment* is focused as a (putative) cause. This distinction is useful in experiments with non-compliance (see, e. g., Jo, 2002a, 2002b, 2002c; Jo et al., 2008).

Unobserved Potential Confounders

Even if we consider a multivariate potential confounder Z consisting of several univariate potential confounders, there are always unobserved variables that are prior or simultaneous to treatment. Such variables are called *unobserved potential confounders*. Sometimes they are also called *hidden confounders* (cf., e. g., Rosenbaum, 2002a). Of course such an unobserved potential confounder may bias the conditional expectation values of the outcome variable just in the same way as an observed covariate. Whether or not the conditional expectation values of the outcome variable in the treatment conditions are unbiased such that their differences represent causal total effects does not only depend on the relationship between the observed variables such as X, Y, and the observed (possible multivariate) covariate, say Z, but also on the relationship of these variables to the unobserved potential confounders. Potential confounders exert their maleficent effects irrespective of whether or not we observe them.

Outcome Variable

Of course, the outcome variable Y refers to a time at which the treatment might have had its impact. Hence, treatment variables are always prior to the outcome variable. In principle, we may also observe several outcome variables, for example, in order to study how the effects of a treatment grow or decline over time or to study treatment effects that are not confined to a single outcome variable. All random variables mentioned above refer to a concrete single-unit trial and they have a joint distribution. Each combination of unit, treatment condition, and score of the outcome variable may be an observed result of such a single-unit trial. This implies that the variables U, Z, X, and Y, as well as unobserved potential confounders, say W, have a joint distribution (see RS-Def. 2.38 and SN-section 5.3). Once we specified the random experiment to be studied, this joint distribution is fixed, even though it might be known only in parts or even be unknown altogether.

Causal Effects and Causal Dependencies

There are already several kinds of causal effects that can be considered in the single-unit trial of a simple experiment or quasi-experiment. For simplicity, suppose the treatment has just two values, say *treatment* and *control*. *First*, there is the *causal average total effect* of treatment (compared to control) on the outcome variable Y. *Second*, there are the *causal conditional total treatment effects* on Y, where we may condition on any function of the observational-unit variable U. If, for example, $Z :=$ *sex* with values m for *male* and f for *female*, then we may consider the causal $(Z = m)$-conditional total treatment effect on Y, that is, the causal average total treatment effect for males, and the causal $(Z = f)$-conditional total treatment effect on Y, that is, the causal average total treatment effect for females. Similarly, if $Z :=$ *socio-economical status*, we may consider the causal conditional total treatment effects on Y for each status group, etc. *Third*, although difficult and often impossible to estimate, we may also consider the causal *individual total effect* of *treatment* compared to *control* on Y.

By definition, within a *simple* experiment and quasi-experiment we cannot consider any *direct* treatment effects with respect to one or more specified potential mediators, that is, the effects of the treatment on the outcome variable that *are not* transmitted through specified potential mediators. However, the causal total treatment effects discussed above are, of course, transmitted through potential mediators, irrespective of whether or not we observe (or are aware of) these potential mediators.

2.2 Experiments With Fallible Covariates

Another class of random experiments are single-unit trials of experiments and quasi-experiments in which we assess a fallible covariate. In this case, the fallible covariate does *not* represent a (deterministic) attribute of the observational units. The single-unit trial of such an experiment of quasi-experiment consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,
(b) assessing the values $z_1, \ldots, z_k$ of the covariates (pre-treatment variables) $Z_1, \ldots, Z_k$, $k \geq 1$.
(c) assigning the unit or observing its assignment to one of several experimental conditions (represented by the value x of the treatment variable X),
(d) recording the value y of the outcome variable Y.

The crucial distinction between a simple (quasi-) experiment and a (quasi-) experiment *with fallible covariates* is that there is variability of at least one of the covariates *given the observational unit u* (see Fig. 2.2). In this case, we may distinguish between the *latent covariate*, say ξ, representing the attribute to be assessed and its *fallible measures*, some manifest variables that can actually be observed. (For the theory of latent variables see Steyer et al., 2015). Also note that sometimes it is crucial to adjust the effect of X on Y by conditioning on the *latent variable* ξ in order to fully adjust for the bias of the prima-facie effect of X on Y. In these cases, only adjusting for the manifest variables that measure the latent covariate ξ does not completely remove bias (see, e. g., Sengewald, Steiner, & Pohl, 2019).

Furthermore, this distinction also implies that the *unit* whose attributes are measured *at the time when the potential confounder is assessed* is not identical any more to the *unit*

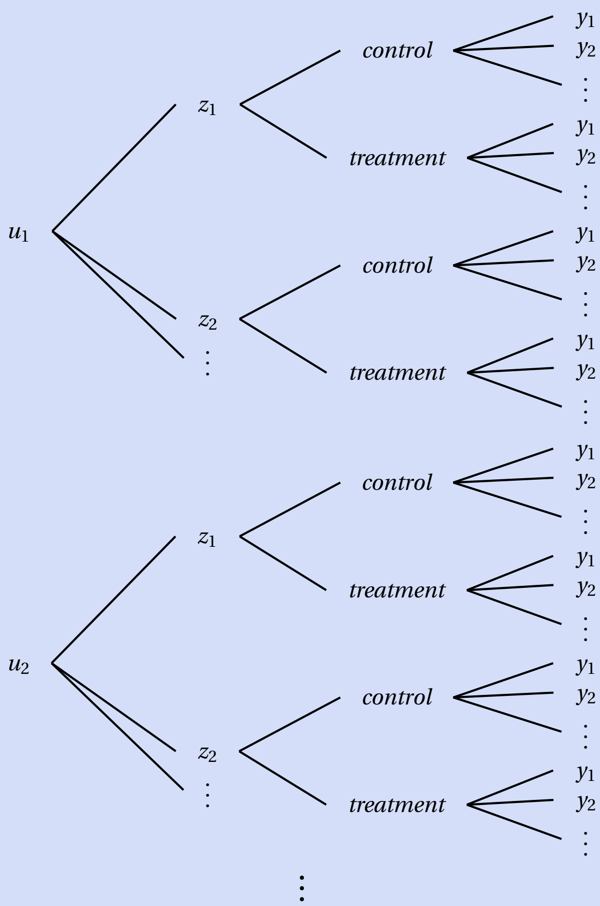

Figure 2.2. Experiment or quasi-experiment with a fallible covariate

at the onset of treatment (see section 2.1). The covariate might be assessed some months before the treatment is given — enough time and plenty of possibilities for the unit to change in various ways, for example, due to maturation, learning, critical life events, and other experiences that are not fixed yet at the time of assessing the covariate. As a consequence, a variable, say W, representing such intermediate events or experiences may also affect the outcome variable Y and the treatment variable. Hence, such intermediate variables are also potential confounders. This is one of the reasons why we need to define causal effects in a more general way than in the Neyman-Rubin tradition (see ch. 5).

Note that assessing a fallible covariate does not only change the interpretation of the observational-unit variable U (now its values are the units of the time of assessment of the manifest covariates), but it also changes the random experiment, and with it, the empirical phenomenon we are considering. Assessing a fallible covariate often involves that the sampled person fills in a questionnaire or takes a test. Assessing, prior to treatment, a fallible covariate such as a test of an ability, an attitude, or a personality trait, may change the observational units and their attributes, as well as the effects of the treatment on a specified outcome variable, which usually is related to such pre-treatment variables. This has already been discussed by Campbell and Stanley (1963), who also recommended designs for studying how pre-treatment assessment modifies the effects of the treatment variable on the outcome variable.

Potential Confounders and Covariates

Which are the potential confounders of the treatment variable X in such a single-unit trial? First of all, it is each attribute of the units at the time of the assessment of the observed covariates. This does not only include variables such as *sex*, *race*, and *educational status*, but also a latent covariate, say ξ, (which might be multi-dimensional). Furthermore, aside from the manifest covariates, each variable W representing an intermediate event or experience of the unit (occurring in between the assessment of the observed covariates and the onset of the treatment), as well as any attribute of the *unit at the onset of treatment* is a potential confounder as well, irrespective of whether or not these potential confounders are observed.

Note that a latent covariate ξ may be considered a cause of its fallible measures $Z_1, \dots, Z_k$ and of the outcome variable Y. This is not in conflict with the theory that the treatment variable X is a cause of Y as well. In this kind of single-unit trial, we have several causes and several outcome variables, and a cause itself can be considered as an outcome variable. For example, it would be possible to consider the treatment variable X to be causally dependent on the manifest or latent covariates. In other words, we may also raise the question if the conditional treatment probabilities $P(X{=}1 \mid Z_1, \dots, Z_k)$ or $P(X{=}1 \mid \xi)$ describe causal dependencies. This makes clear that the terms 'potential confounder' and 'covariate' can only be defined with respect to a focused cause.

2.3 Two-Factorial Experiments

As a third class of random experiments we consider *two-factorial experiments*. The single-unit trial of such a two-factorial experiment or quasi-experiment consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,

(b) assigning the unit or observing its assignment to one of several experimental conditions that are defined by the pair (x, z) of levels of two treatment variables X and Z, respectively.

(c) recording the value y of the outcome variable Y.

Sampling a Unit

Because we presume that no fallible potential confounders such as 'severity of symptoms', 'motivation for treatment', etc. are assessed before treatment, sampling an observational unit means that we are sampling a *unit at the onset of treatment*.

Treatment Variables

As a simple example, let us consider an experiment in which we study the effects — including the joint effects — of two treatment factors, say *individual therapy* represented by X (with values '*yes*' and '*no*') and *group therapy* represented by Z (with values '*yes*' and '*no*').

In such a two-factorial experiment, we may consider *group therapy* as a covariate and *individual therapy* to be the cause in order to ask for the conditional and average total effects of individual therapy given *group therapy* and given *no group therapy*. In contrast, we may also consider *individual therapy* to be a covariate and *group therapy* to be the focused treatment variable. Finally, we may also consider the two-dimensional variable (X, Z) as the cause. Which option is chosen depends on the causal effects we are interested in (see below).

Outcome Variable

Again, the outcome variable Y refers to a time at which the treatment might have exerted the effects to be estimated. Hence, both treatment variables are prior to the outcome variable considered. And again, we may also observe several outcome variables, for example, in order to study how effects of a treatment grow or decline over time or to study effects that are not confined to a single outcome variable.

Causal Effects

There are several causal effects we might look at. If X and Z have only two values, then we may be interested in the following effects on the outcome variable Y:

(a_1) the conditional total effect of 'individual therapy' as compared to 'no individual therapy' given that the unit treated also receives 'group therapy',
(b_1) the corresponding conditional total effect given that the unit does *not* receive 'group therapy', and
(c_1) the average of these conditional total effects of 'individual therapy' as compared to 'no individual therapy', averaging over the two values of Z (group therapy).

Vice versa, we might also be interested in the following effects on the outcome variable Y:

(a_2) the conditional total effect of 'group therapy' as compared to 'no group therapy' given that the unit treated also receives 'individual therapy',
(b_2) the corresponding conditional total effect given that the unit does *not* receive 'individual therapy', and
(c_2) the average of these conditional total effects of 'group therapy' as compared to 'no group therapy', averaging over the two values of X.

Furthermore, there are other causal effects on Y we might study, namely

(a_3) the total effect of receiving 'individual therapy' *and* 'group therapy' as compared to receiving none of the two treatments.
(b_3) the total effect of receiving 'individual therapy' *and* 'no group therapy' as compared to receiving 'group therapy' *and* 'no individual therapy'.

All these effects may answer meaningful causal questions. In fact there are even more causal effects than those listed above. For example, we could compare each of the four combinations of the two treatments to an average of the other treatments. Furthermore, many additional causal effects can be considered if we condition on other covariates such as *sex* or *educational status*.

Potential Confounders and Covariates

If we focus on the effect of X (individual therapy), then we consider Z (group therapy) as a covariate of X. In contrast, we treat X as a covariate of Z if we study the effects of Z (group therapy). Furthermore, in both cases, each attribute of the unit at the onset of treatment (such as *sex* or *educational status*) could be considered as covariates as well. Assessing these covariates does not appear in points (a) to (c) of the random experiment, because these covariates are (deterministic) functions of the observational-unit variable. Therefore, there is no additional sampling process associated with their assessment.

This is also true for other potential confounders, for example, variables characterizing the situation in which the unit is at the onset of treatment, the number of *hours slept* last night, or *day time* at which the unit receives its treatment. Even variables that characterize early experiences in the childhood of the unit such as a *broken home* or *mother's child care behavior* are potential confounders in this single-unit trial. They are there and exert their effects even if they are not assessed.

Note again that assessment of these potential confounders in a questionnaire filled in by the person constitutes a new random experiment that may differ in important ways from a random experiment in which the unit has no such task (see sect. 2.2). In psychology, an assessment often is a treatment of its own.

2.4 Multilevel Experiments

In multilevel experiments and quasi-experiments we also study the effect of a treatment on an outcome variable. However, in such a design the observational units are nested within higher hierarchical units referred to as *clusters*. Examples include experiments, in which students are nested within classrooms, patients are nested within clinics, and inhabitants are nested in cities and neighborhoods. Multilevel designs can be classified as designs with treatment assignment at the unit-level or at the cluster-level. Furthermore, multilevel designs differ with respect to the assignment of units to clusters. There are designs with pre-existing clusters and there are designs with assignment of units to clusters. All these designs involve different single-unit trials.

A single-unit trial with *pre-existing clusters* consists of:

(a) sampling a cluster c (e.g., a school class, a neighborhood, or a hospital) from a set of clusters,
(b) sampling an observational unit u (e.g., a person) from a set of units within the cluster,
(c) assigning the unit or the cluster (depending on the design) or observing their assignment to one of several experimental conditions (represented by the value x of the treatment variable X),
(d) recording the value y of the outcome variable Y.

In contrast, a single-unit trial *with assignment of units to clusters* consists of:

(a) sampling an observational unit u (e. g., a person) from a set of units,
(b) assigning the unit or observing its assignment to one of several clusters (represented by the value c of the cluster variable C),
(c) assigning the unit or the cluster (depending on the design) or observing their assignment to one of several experimental conditions (represented by the value x of the treatment variable X),
(d) recording the value y of the outcome variable Y.

In the experiment with pre-existing clusters, each unit can only appear in one cluster, whereas in the experiment with assignment of units to a cluster, each unit can be assigned to any cluster. Note again, that we are considering single-units trials from the pre-factual perspective, not from the post-factual or 'counter-factual' perspective (see the remarks following the description of the random experiment presented in Table 1.1). Hence, in experiments with assignment of units to a cluster, the cluster variable can bias the dependency of the outcome variable on the treatment variable *on the level of the observational unit*. In this aspect this design resembles the multifactorial design described in section 2.3.

Potential Confounders and Covariates

Which are the potential confounders in multilevel designs if the treatment variable X is considered as the cause? The answer depends on the type of design considered: In designs with assignment of units to clusters, attributes of the observational unit such as *sex*, *race*, or *educational status*, are potential confounders of X. Other potential confounders are attributes of the cluster such as *school type*, *hospital ownership*, or *school-level socio-economic status* or *school-level intelligence*. The last two kinds of potential confounders would be defined as conditional expectations of the corresponding potential confounders at the unit-level given the cluster variable.

In these designs, clusters may not only be considered as potential confounders, but also as treatments, because some of the effects observed later on may depend on the composition of the group to which a particular unit, say Joe, is assigned. Receiving group therapy together with beautiful Ann in the same group might make a great difference as compared to getting it together with awful Joe. In designs in which clusters as a whole are assigned to treatment conditions, only attributes of the cluster can influence the assignment. Hence, in data analysis we would focus on controlling for the potential confounders on the cluster level (see, e. g., Nagengast, 2009, for more details).

2.5 Experiments With Latent Outcome Variables

We may also consider single-unit trials of experiments with a *latent outcome variable*. The basic goal of such experiments is to investigate the effect of the treatment variable X on a *latent* outcome variable, say η. This is of interest, for example, where a quantitative outcome variable can only be measured by qualitative observations such as solving or not solving certain items indicating the (latent) ability. However, it can also be of interest if the manifest measures are *linearly* related to the latent variable such as in models of classical test theory (see, e. g., Steyer, 2001) or in models of latent state-trait theory (see, e. g., Steyer et al., 2015). If, for example, there are three manifest variables Y_1, Y_2, and Y_3 measuring a single latent variable η, then we may ask if there is just one single effect of the treatment on the latent outcome variable η – which transmits these effects to the manifest variables Y_1, Y_2, and Y_3 – instead of three separate effects of X on each variable Y_i. Hence, the latent variable may also be considered to be a mediator variable. Showing that all effects of X on the variables Y_i are indirect, that is, mediated by η is one of the research efforts that aims at establishing construct validity of the latent variable η.

In the simplest case with a single latent variable, we consider the following single-unit trial:

(a) Sampling a person u_0 out of a set of persons,

(b) assigning the unit or observing its assignment to one of several experimental conditions (represented by the value x of the treatment variable X),

(c) recording the values $y_1, \dots, y_m$ of the manifest outcome variables $Y_1, \dots, Y_m$.

In this single-unit trial, a value u_1 of the observational-unit variable U_1 represents the observational *unit at the onset of treatment*, while the latent outcome variable η represents some attribute of the same unit at the time point at which the outcome of the treatment is assessed. In such a design we have to distinguish between u_1, a unit at time 1, and u_2, the same unit at time 2 (see Steyer et al., 2015) for more details). Obviously, time 2 is *after* treatment and *prior* to the observation of the manifest outcome variables Y_i, at least as long as we preclude change in the latent variable during the process of assessing the manifest outcome variables. If this cannot be precluded, we would have to consider the time sequence in assessing the manifest outcome variables (e. g., of the items to be solved or answered) as well.

Potential Confounders and Covariates

Which are the potential confounders in such a single-unit trial? Again, the answer depends on the cause considered. If it is the treatment variable X, then each attribute of the *unit at the onset of treatment* is a potential confounder (with respect to X). Obviously, this again includes variables such as *sex*, *race*, and *educational status*. Note that in this kind of experiments, the set of potential confounders of X is the same irrespective of the choice of the outcome variable. Remember, we may not only consider the *latent* outcome variable η but also the *manifest* outcome variables Y_i, for example, in order to study whether or not the effects of X on these manifest outcome variables are perfectly transmitted (or mediated) through the latent variable η.

Choosing the latent outcome variable η as a cause of the manifest outcomes variables Y_i brings additional potential confounders into play, for instance, all those variables that are *in between* treatment and the assessment of η. If, for example, we consider an experiment studying the effects of different teaching methods, these additional potential confounders are critical life events (such as father or mother leaving the family), or additional lessons taken after treatment and before outcome assessment, for instance.

Box 2.1 Glossary of new concepts

Note that all terms mentioned in this box are still of an informal nature. Their mathematical specification starts in chapter 3.

Random experiment	The kind of empirical phenomenon to which events, random variables, and their dependencies refer.
Randomized experiment	An random experiment in which the experimenter fixes the treatment probabilities for each observational unit.
Single-unit trial	A particular kind of random experiment that consists of sampling a single unit from a set of observational units and observing the values of one or more random variables related to this unit.
Cause	A random variable. Its effect on an outcome variable is considered.
Outcome variable	A random variable. Its dependency on a cause is considered.
Potential confounder	If we confine the discussion to total causal effects, then it is a random variable that is prior or simultaneous to the cause considered. It might be correlated with the cause and the outcome variable.
Covariate	A potential confounder that is considered together with X in a conditional expectation or a conditional distribution.
Fallible covariate	A covariate that is assessed with measurement error.
Latent covariate	A covariate that is not directly observed. Instead it is defined using some parameters of the joint distribution of a set of manifest random variables.
Intermediate variable	A variable that might mediate (transmit) the effect of the cause on the outcome variable. The cause is always prior to an potential mediator and an potential mediator is always prior to the outcome variable. An potential mediator is not *necessarily* affected by the cause and it does not *necessarily* have an effect on the outcome variable.
Mediator	An intermediate variable on which X has a causal effect and which itself has a causal effect on the outcome variable Y.

2.6 Summary and Conclusion

In this chapter we described a number of random experiments in informal terms. The purpose was to get a first idea of which kind of empirical phenomena causal theories and hypotheses refer to. We focused on single-unit trials, which are the kinds of empirical phenomena we are interested in, both in theory and practice. We emphasized that a single-unit trial is a random experiment and discussed several kinds of random variables playing a crucial role in the theory of causal effects. We also mentioned that there is a certain *time order* among these random variables, for example, saying that the potential confounders are 'prior' or 'simultaneous' to the treatment variable, which itself is 'prior' to the outcome variable. Furthermore, for each single-unit trial and each cause in such a single-unit trial, we discussed the *potential confounders* involved. We emphasized that each cause considered in such a single-unit trial has its own set of potential confounders.

Other Single-Unit Trials

The single-unit trials discussed in this chapter are just a small selection of single-unit trials in which causal effects and causality of stochastic dependencies are of interest. We might also consider single-unit trials with latent covariates *and* latent outcome variables *and* manifest and/or latent potential mediators, but also single-unit trials with multiple mediation. Furthermore, we could also consider single-unit trials of growth curve models (see, e. g., Biesanz, Deeb-Sossa, Aubrecht, Bollen, & Curran, 2004; Bollen & Curran, 2006; Meredith & Tisak, 1990; Singer & Willett, 2003; Tisak & Tisak, 2000), latent change models (see, e. g., McArdle, 2001; Steyer, Eid, & Schwenkmezger, 1997; Steyer, 2005), or cross-lagged panel models (see, e. g., Kenny, 1975; Rogosa, 1980; Watkins, Lei, & Canivez, 2007; Wolf, Chandler, & Spies, 1981). Causality is also an issue in uni- and multivariate time-series analysis as well as in stochastic processes with continuous time. However, in this book our examples will usually deal with experiments and quasi-experiments, including latent covariates and outcome variables.

Outlook

In chapter 3 we will study additional mathematic concepts that allow us to meaningfully talk about time order between events and random variables. This will be used in chapter 4, in which we introduce the notions of a *causality space* and the concept of a *potential confounder*. This will provide the mathematical framework and language in which causal effects can meaningfully be defined (see ch. 5).

2.7 Exercises

▹ **Exercise 2-1** Imagine that the probabilities of a crash for a flight with Airline A is ten times smaller than with Airline B. Which airline would you choose?

▹ **Exercise 2-2** Why does the theory of causal effects refer to single-unit trials?

▹ **Exercise 2-3** Why is it important to know which random experiment we are talking about?

▹ **Exercise 2-4** Which type of random experiment did we refer to in the two examples described in chapter 1?

▹ **Exercise 2-5** Why is it important to emphasize that, in simple experiments and quasi-experiments (see section 2.1), the observational-unit variable U represents the observational units *at the onset of treatment*?

▹ **Exercise 2-6** What is the basic idea of a potential confounder of a cause?

▹ **Exercise 2-7** Which kinds of causal effects can be considered in the simple experiment or quasi-experiment in which no *fallible* potential confounder and no potential mediator is assessed?

Solutions

▹ **Solution 2-1** If your answer is A, then you implicitly apply these probabilities to the random experiment of flying *once* with A or B, even if these probabilities have been estimated in a sample. This example serves to emphasize that, not only in theory but also in practice, we are mainly interested in a single-unit trial, not in a sample consisting of many such single-unit trials, and in particular not in what applies to sample size going to infinity. (This is how many applied statisticians try to specify the term 'population'.)

▹ **Solution 2-2** Within such a single-unit trial, the various concepts of causal effects can be defined and we can study how to identify these causal effects from the parameters describing the joint distribution of the random variables considered. In such a single-unit trial, there usually is a clear time order which helps (but is not sufficient) to disentangle the possible causal relationships between the random variables considered.

▹ **Solution 2-3** Different random experiments are different empirical phenomena. Although the names of the variables in different random experiments might be the same, the variables themselves are different entities, implying that the dependencies and effects between these variables might differ between different random experiments.

▹ **Solution 2-4** The type of random experiment we refer to in these examples is the single-unit trial of simple experiments and quasi-experiments described in section 2.1, because there is no extra sampling of a covariate. Instead, the value of this covariate is fixed as soon as the person is sampled. That is, the covariate is an attribute of the person.

▹ **Solution 2-5** In the social sciences, units are often persons, and persons can change over time. If, in a simple experiment or quasi-experiment, a value u of U represents the observational unit sampled *at the onset of treatment*, each potential confounder is a function of U. If, in contrast, U represents the *observational unit at the assessment of a fallible covariate* (see sect. 2.2), which is some time prior to the onset of treatment, then there can be other potential confounders in between assessment of the fallible potential confounder and the onset of treatment. We have to consider these additional potential confounders both in the definition of causal effects and in data analysis.

▹ **Solution 2-6** A potential confounder of a cause is a random variable that is prior or simultaneous to the cause, at least as long as we only consider total effects. (If we also consider direct effects, then a potential confounder can also be posterior to a cause.)

▹ **Solution 2-7** If the treatment has just two values, say *treatment* and *control*, then there are different kinds of causal effects of the treatment variable on the outcome variable Y, such as the average total treatment effect, the conditional total treatment effects given a value z of a covariate Z, and the *individual total effect* of X on Y given an observational unit u. Aside from these treatment effects, we may also consider the causal effects of a potential confounder Z on the treatment variable X, but also on the outcome variable Y.

Part II

Basic Concepts of the Theory of Causal Total Effects

Chapter 3

Time Order

In chapter 1 we studied some examples showing that the conditional expectation values $E(Y|X{=}x)$ of an outcome variable Y and their differences $E(Y|X{=}x) - E(Y|X{=}x')$, the *prima facie effects*, can be seriously misleading in evaluating the causal effect of a (treatment) variable X on an (outcome or response) variable Y. Hence, conditional expectation values cannot be used offhandedly to define the causal effects in which we are interested when we want to evaluate a treatment, an intervention, or an exposition. For the purpose of such an evaluation, the concept of a conditional expectation value $E(Y|X{=}x)$ has two deficits. The *first* one is that the terms $E(Y|X{=}x)$ and $E(Y|X{=}x')$ do not necessarily describe the kind of dependency in which we are interested for the evaluation of a treatment. This deficit is related to (causal) bias, which will be treated in chapter 6. The *second* deficit is that it does not guarantee that X *is prior to* Y in time, which is indispensable for a difference $E(Y|X{=}x) - E(Y|X{=}x')$ to describe a causal effect of a value x of X compared to another value x' of X.

Overview

In the present chapter, we focus on time order of sets of events and of random variables. We introduce the concepts of a *filtration* and the relations *prior to*, *simultaneous to*, and *prior or simultaneous to* with respect to a filtration. Note that the definitions of these relations do not involve a probability measure. Instead, they can be introduced for *measurable set systems* and *measurable maps*. In the framework of a probability space, these relations represent time order of *sets of events* and *random variables*, respectively. For brevity, we refrain from explicitly treating the corresponding relations among *measurable sets* (and with it, among *events*).

Requirements

Reading this chapter requires that the reader is familiar with the contents of the first two chapters of Steyer (2024). The first of these chapters deals with the concepts of *probability* and *conditional probability* of events, including the necessary mathematical framework such as a *probability space* $(\Omega, \mathscr{A}, P)$ consisting of a set Ω of possible outcomes, a σ-algebra $\mathscr{A}$ of events, and a probability measure P on $\mathscr{A}$. The second chapter introduces the concepts of a *random variable* as a special measurable map and its *distribution* as a special image measure. These chapters will be referred to as RS-chapter 1 and RS-chapter 2. The same kind of shortcut is used when referring to other parts of that book, such as sections, definitions, theorems, remarks, or equations, for instance.

3.1 Filtration

The definition of an event in probability theory does not presume that there is a time order between events, sets of events, and random variables. However, in many applications of probability theory such a time order is important, in particalur if causal interpretations of dependencies are intended. In both examples presented in chapter 1, for instance, it is crucial for the evaluation of the treatment that the treatment variable is *prior* to the outcome variable Y. Such a time order can be defined with respect to a *filtration*, a fundamental concept of the theory of stochastic processes (see, e. g., Klenke, 2020, Def. 9.9).

Definition 3.1 [Filtration]
Let $(\Omega, \mathscr{A})$ be a measurable space, $T \subset \mathbb{R}$, and $T \neq \emptyset$. A family $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ of σ-algebras $\mathscr{F}_t \subset \mathscr{A}$ is called a filtration in $\mathscr{A}$, if $\mathscr{F}_s \subset \mathscr{F}_t$, for all $s, t \in T$ with $s \leq t$.

Note that this concept is defined in the context of a measurable space $(\Omega, \mathscr{A})$ (see RS-Def. 1.4), which just consists of a set Ω and a σ-algebra $\mathscr{A}$. No probability measure is involved, and this also applies to the prior-to relations that will be introduced in section 3.2. Hence, throughout this chapter, we do not presume that there is a probability measure P on $(\Omega, \mathscr{A})$. Nevertheless, our examples refer to random experiments, which *are* represented by a probability space $(\Omega, \mathscr{A}, P)$.

RST: Vielleicht auch filtered measurable space einführen.

Example 3.2 [Joe and Ann With Self-Selection] In the random experiment presented in Table 1.2, all elements $\omega_1, \ldots, \omega_8$ of the *set of possible outcomes* are listed in the first column of this table. The set of these possible outcomes is the Cartesian product

$$\begin{aligned} \Omega &= \Omega_U \times \Omega_X \times \Omega_Y \\ &= \{\omega_1, \ldots, \omega_8\} \\ &= \big\{(Joe, no, -), \ldots, (Ann, yes, +)\big\}, \end{aligned} \tag{3.1}$$

where $\Omega_U = \{Joe, Ann\}$, $\Omega_X = \{no, yes\}$, and $\Omega_Y = \{-, +\}$. The *σ-algebra on* Ω is specified by $\mathscr{A} = \mathscr{P}(\Omega)$, that is, $\mathscr{A}$ is chosen to be the power set of Ω, which is the set of all subsets of Ω, consisting of $2^8 = 256$ elements. Finally, the *probability measure* $P\colon \Omega \to \mathscr{A}$ is specified by the assignment of the probabilities $P(\{\omega_i\})$ to the eight elements of Ω. These probabilities are shown in the second column of Table 1.2. All other 248 probabilities $P(A)$, $A \in \mathscr{A}$, can be computed from the probabilities $P(\{\omega_i\})$, $i = 1, \ldots, 8$, because, except for the empty set, they are unions of the elementary events $\{\omega_i\}$ [see RS-Box 1.1 (x)].

In this example, the *person variable*

$$U\colon \Omega \to \Omega_U \tag{3.2}$$

has the co-domain $\Omega_U = \{Joe, Ann\}$, the *treatment variable*

$$X\colon \Omega \to \Omega'_X \tag{3.3}$$

has the co-domain $\Omega'_X = \{0, 1\}$, and the *outcome variable*

$$Y\colon \Omega \to \Omega'_Y \tag{3.4}$$

has the co-domain $\Omega'_Y = \{0,1\}$. Table 1.2 shows the assignment of values of these random variables to each element of Ω. Furthermore, we choose the σ-algebras $\mathscr{A}_U = \mathscr{P}(\Omega_U)$ and $\mathscr{A}'_X = \mathscr{A}'_Y = \mathscr{P}(\{0,1\})$ to be the power sets of Ω_U and of $\Omega'_X = \Omega'_Y$, respectively. Hence, the value space of U is $(\Omega_U, \mathscr{A}_U)$, and $(\Omega'_X, \mathscr{A}'_X) = (\Omega'_Y, \mathscr{A}'_Y)$ is the value space of X and Y (see RS-Def. 2.2).

Now, we specify the filtration $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$, $T = \{1,2,3\}$, in $\mathscr{A}$ by

$$\mathscr{F}_1 = \sigma(U), \qquad \mathscr{F}_2 = \sigma(U,X), \qquad \mathscr{F}_3 = \sigma(U,X,Y) \tag{3.5}$$

(see RS-Def. 2.12). The first of these three σ-algebras is the *σ-algebra generated by U*,

$$\begin{aligned} \sigma(U) &= \{U^{-1}(A') \colon A' \in \mathscr{A}_U\} \\ &= \{U^{-1}(\{Joe\}),\, U^{-1}(\{Ann\}),\, U^{-1}(\Omega_U),\, U^{-1}(\varnothing)\}. \end{aligned} \tag{3.6}$$

This σ-algebra consists of four events, the *inverse images* $U^{-1}(A') = \{\omega \in \Omega \colon U(\omega) \in A'\}$ of the elements A' of

$$\mathscr{A}_U = \mathscr{P}(\Omega_U) = \{\{Joe\}, \{Ann\}, \Omega_U, \varnothing\}$$

under the map U. These inverse images are the sets

$$\begin{aligned} U^{-1}(\{Joe\}) &= \{(Joe,no,-),(Joe,no,+),(Joe,yes,-),(Joe,yes,+)\} \\ &= \{\omega_1,\omega_2,\omega_3,\omega_4\} \end{aligned} \tag{3.7}$$

$$\begin{aligned} U^{-1}(\{Ann\}) &= \{(Ann,no,-),(Ann,no,+),(Ann,yes,-),(Ann,yes,+)\} \\ &= \{\omega_5,\omega_6,\omega_7,\omega_8\} \end{aligned} \tag{3.8}$$

$$U^{-1}(\Omega_U) = \Omega \tag{3.9}$$

$$U^{-1}(\varnothing) = \varnothing \tag{3.10}$$

(see Exercises 3-1 and 3-2). Hence, these sets are the four events that the person variable U takes on

- the value *Joe*, which is the event that *Joe is drawn*,
- the value *Ann*, which is the event that *Ann is drawn*,
- a value in the set $\Omega_U = \{Joe, Ann\}$, which is the event that *Joe or Ann is drawn*,
- a value in the empty set, which is the event that *no one is drawn*.

The *σ-algebra generated by (the bivariate random variable)* (U,X) (see RS-sect. 2.1.4) is

$$\sigma(U,X) = \{(U,X)^{-1}(A') \colon A' \in \mathscr{A}_U \otimes \mathscr{A}'_X\}. \tag{3.11}$$

This σ-algebra consists of $2^4 = 16$ elements, the inverse images $(U,X)^{-1}(A')$ of the elements A' of the *product σ-algebra*

$$\begin{aligned} \mathscr{A}_U \otimes \mathscr{A}'_X &= \sigma(A_1 \times A_2 \colon A_1 \in \mathscr{A}_U, A_2 \in \mathscr{A}'_X) \\ &= \Big\{\{(Joe,0)\}, \{(Joe,1)\}, \{(Ann,0)\}, \{(Ann,1)\}, \\ &\qquad \{(Joe,0),(Joe,1)\}, \{(Ann,0),(Ann,1)\}, \\ &\qquad \{(Joe,0),(Ann,0)\}, \{(Joe,1),(Ann,1)\}, \end{aligned} \tag{3.12}$$

$$\big\{(Joe,0),(Ann,1)\big\}, \big\{(Joe,1),(Ann,0)\big\},$$
$$\big\{(Joe,0),(Joe,1),(Ann,0)\big\}, \big\{(Joe,0),(Joe,1),(Ann,1)\big\},$$
$$\big\{(Joe,0),(Ann,0),(Ann,1)\big\}, \big\{(Joe,1),(Ann,0),(Ann,1)\big\},$$
$$\Omega_U \times \Omega'_X, \emptyset\Big\}$$

(see RS-Def. 1.15). In this example, this set is identical to the power set of the Cartesian product $\Omega_U \times \Omega'_X$. Hence, gathering the inverse images of all elements of $\mathscr{A}_U \otimes \mathscr{A}'_X$ [see Eq. (3.11)] yields

$$\begin{aligned}\sigma(U,X) = \big\{&\{\omega_1,\omega_2\}, \{\omega_3,\omega_4\}, \{\omega_5,\omega_6\}, \{\omega_7,\omega_8\},\\ &\{\omega_1,\omega_2,\omega_3,\omega_4\}, \{\omega_5,\omega_6,\omega_7,\omega_8\},\\ &\{\omega_1,\omega_2,\omega_5,\omega_6\}, \{\omega_1,\omega_2,\omega_7,\omega_8\},\\ &\{\omega_3,\omega_4,\omega_5,\omega_6\}, \{\omega_3,\omega_4,\omega_7,\omega_8\},\\ &\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}, \{\omega_1,\omega_2,\omega_3,\omega_4,\omega_7,\omega_8\},\\ &\{\omega_1,\omega_2,\omega_5,\omega_6,\omega_7,\omega_8\}, \{\omega_3,\omega_4,\omega_5,\omega_6,\omega_7,\omega_8\}, \Omega, \emptyset\big\}\end{aligned} \tag{3.13}$$

(see Exercise 3-4). Comparing $\sigma(U)$ to $\sigma(U,X)$ [see Eqs. (3.7) to (3.10)] shows that all elements of $\sigma(U)$ are also elements of $\sigma(U,X)$. Hence, $\sigma(U) \subset \sigma(U,X)$.

Finally, the σ-algebra $\sigma(U,X,Y)$ generated by (U,X,Y) is identical to the power set of Ω. It consists of $2^8 = 256$ elements. Because $\sigma(U) \subset \sigma(U,X)$ and the power set consists of *all subsets of* Ω, we can conclude $\sigma(U) \subset \sigma(U,X) \subset \sigma(U,X,Y)$. Hence, $\mathscr{F}_T$ defined by Equations (3.5) is a filtration in $\mathscr{A}$. ◁

3.2 Prior-To Relations

Now we introduce time order. More precisely, we define the *prior-to relation* among *measurable set systems* (*sets of events*), and *measurable maps* (*random variables*). For these definitions, it suffices to refer to a *measurable space* $(\Omega,\mathscr{A})$. In the framework of a probability space $(\Omega,\mathscr{A},P)$, measurable set systems (i.e., subsets of $\mathscr{A}$) are *sets of events* (see RS-Def. 1.4) and measurable maps are *random variables* (see RS-Def. 2.2).

Reading the following definition, note that $\exists$ means 'there is' and $\wedge$ symbolizes the conjunction of two propositions, that is, the logical 'and'. Furthermore, remember that the σ-algebra $\sigma(X)$ generated by a measurable map X on $(\Omega,\mathscr{A})$ is a set system satisfying, among other things, $\sigma(X) \subset \mathscr{A}$ (see RS-Def. 2.12).

Definition 3.3 [Prior-to Relations]
Let $(\Omega,\mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t\in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C},\mathscr{D} \subset \mathscr{A}$.

(i) We say $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$ (and $\mathscr{D}$ posterior in $\mathscr{F}_T$ to $\mathscr{C}$), denoted $\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D}$, if
  (a) $\exists\, s \in T$: $\mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$, and
  (b) $\exists\, t \in T$: $\mathscr{D} \subset \mathscr{F}_t$.

(ii) Let X and Y be measurable maps on $(\Omega,\mathscr{A})$. Then we say X is prior in $\mathscr{F}_T$ to Y (and Y posterior in $\mathscr{F}_T$ to X), denoted $X \underset{\mathscr{F}_T}{\prec} Y$, if $\sigma(X)$ is prior in $\mathscr{F}_T$ to $\sigma(Y)$.

Remark 3.4 [Prior-to Relation of Measurables Sets and of Events] Let $(\Omega, \mathscr{A})$ be a measurable space and $\{C\}$, $\{D\}$ be set systems containing the sets $C, D \in \mathscr{A}$ as their only element. In the context of a probability space $(\Omega, \mathscr{A}, P)$, the sets C and D represent events. We say that *C is prior in $\mathscr{F}_T$ to D* (or *D posterior in $\mathscr{F}_T$ to C*) and denote it by $C \underset{\mathscr{F}_T}{\prec} D$, if $\{C\}$ is prior in $\mathscr{F}_T$ to $\{D\}$. Using this definition, all propositions about the prior-to relation of measurable set systems can be applied to the prior-to relation of measurable sets. However, for brevity, we refrain from explicitly spelling out more details about the prior-to relation of measurable sets or events. ◁

Example 3.5 [Joe and Ann With Self-Selection] In Example 3.2, we specified the probability space $(\Omega, \mathscr{A}, P)$, the random variables U, X, Y, and the filtration $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ in $\mathscr{A}$, $T = \{1, 2, 3\}$, by

$$\mathscr{F}_1 = \sigma(U), \qquad \mathscr{F}_2 = \sigma(U, X), \qquad \mathscr{F}_3 = \sigma(U, X, Y),$$

where $\sigma(U)$, $\sigma(U, X)$, and $\sigma(U, X, Y)$ are the σ-algebras generated by the random variables U, (U, X), and (U, X, Y), respectively.

According to Definition 3.3 (i), the *set system* $\sigma(U)$ *is prior in* $\mathscr{F}_T$ *to the set system* $\sigma(X)$ because $\sigma(U) \subset \mathscr{F}_1$, $\sigma(X) \not\subset \mathscr{F}_1$ [see Def. 3.3 (i) (a)], but $\sigma(X) \subset \mathscr{F}_2$ [see Def. 3.3 (i) (b)]. Note that

$$\mathscr{F}_2 = \sigma(U, X) = \sigma\big(\sigma(U) \cup \sigma(X)\big)$$

[see RS-Eq. (2.16)], which implies $\sigma(X) \subset \mathscr{F}_2$. Similarly, $\sigma(U)$ is prior in $\mathscr{F}_T$ to $\sigma(Y)$ because $\sigma(U) \subset \mathscr{F}_1$, $\sigma(Y) \not\subset \mathscr{F}_1$, but $\sigma(Y) \subset \mathscr{F}_3$. Again note that

$$\mathscr{F}_3 = \sigma(U, X, Y) = \sigma\big(\sigma(U) \cup \sigma(X) \cup \sigma(Y)\big)$$

[see again RS-Eq. (2.16)], which implies $\sigma(Y) \subset \mathscr{F}_3$. Finally, $\sigma(X)$ is prior in $\mathscr{F}_T$ to $\sigma(Y)$ because $\sigma(X) \subset \mathscr{F}_2$ and $\sigma(Y) \not\subset \mathscr{F}_2$, but $\sigma(Y) \subset \mathscr{F}_3$ (see Exercise 3-5).

Remember, that random variables are measurable maps and that the prior-to relation of measurable maps is defined via their generated σ-algebras . Now that we clarified the prior-to relation of the set systems $\sigma(U)$, $\sigma(X)$, and $\sigma(Y)$, we can also conclude that U is prior in $\mathscr{F}_T$ to X and Y, and that X is prior in $\mathscr{F}_T$ to Y [see Def. 3.3 (ii)]. ◁

3.2.1 Properties of the Prior-to Relation of Measurable Set Systems

Now we study some properties of the prior-to relations. First of all, let us ascertain that the prior-to relation of measurable set systems is asymmetric and transitive. Reading this theorem, note that $\neg$ denotes the *negation* of a proposition and $\Rightarrow$ the *implication* between two propositions.

Theorem 3.6 [Asymmetry and Transitivity]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D}, \mathscr{E} \subset \mathscr{A}$. Then:

(i) $\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \;\Rightarrow\; \neg(\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C})$ *(asymmetry)*

(ii) $(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E}) \;\Rightarrow\; \mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E}$. *(transitivity)*

(Proof p. 70)

Hence, according to Proposition (i) of Theorem 3.6, the set system $\mathscr{C}$ being prior in $\mathscr{F}_T$ to the set system $\mathscr{D}$ implies that $\mathscr{D}$ is not prior in $\mathscr{F}_T$ to $\mathscr{C}$. This property of the prior-to relation is called *asymmetry*. Furthermore, according to Proposition (ii) of Theorem 3.6, the prior-to relation is *transitive*. That is, if $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$ that itself is prior in $\mathscr{F}_T$ to $\mathscr{E}$, then $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{E}$.

The following remark prepares another important property of the prior-to relation of measurable set systems.

Remark 3.7 [Some Propositions About Subsets] Some general properties of subsets are:

$$(A \subset B \ \wedge \ A \subset C) \quad \Leftrightarrow \quad A \subset (B \cap C) \tag{3.14}$$

$$(A \subset C \ \wedge \ B \subset C) \quad \Leftrightarrow \quad (A \cup B) \subset C \tag{3.15}$$

and

$$(A \subset B \ \wedge \ B \subset C) \quad \Rightarrow \quad A \subset C \tag{3.16}$$

(see Exercises 3-6 to 3-8). Hence, according to Proposition (3.16), if $(\Omega, \mathscr{A})$ is a measurable space and $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, then:

$$(\mathscr{C}_0 \subset \mathscr{C} \ \wedge \ \mathscr{C} \subset \mathscr{F}_t) \quad \Rightarrow \quad \mathscr{C}_0 \subset \mathscr{F}_t. \tag{3.17}$$

$\triangleleft$

Theorem 3.8 [An Implication of $\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D}$ for a Subset of $\mathscr{C}$]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$. Then

$$\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \quad \Rightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\prec} \mathscr{D}, \qquad \textit{if } \mathscr{C}_0 \subset \mathscr{C}. \tag{3.18}$$

(Proof p. 70)

Now we turn to some properties of the prior-to relation of measurable set systems that involve σ-algebras generated by set systems. Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C} \subset \mathscr{A}$. Because each $\mathscr{F}_t$, $t \in T$, is a σ-algebra,

$$\forall\, t \in T\colon \quad \mathscr{C} \subset \mathscr{F}_t \ \Leftrightarrow \ \sigma(\mathscr{C}) \subset \mathscr{F}_t \tag{3.19}$$

holds for the σ-algebra $\sigma(\mathscr{C})$ generated by $\mathscr{C}$ [see RS-Def. 1.7 and RS-Prop. (1.5)]. This proposition is used in the proof of the following theorem.

Theorem 3.9 [First Properties of the Prior-to Relation Involving σ-Algebras]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$. Then:

(i) $\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \quad \Leftrightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\prec} \sigma(\mathscr{D})$

(ii) $\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \quad \Leftrightarrow \quad \sigma(\mathscr{C}) \underset{\mathscr{F}_T}{\prec} \mathscr{D}.$

(Proof p. 70)

Hence, according to Proposition (i) of Theorem 3.9, the set system $\mathscr{C}$ being prior in $\mathscr{F}_T$ to the set system $\mathscr{D}$ is equivalent to $\mathscr{C}$ being prior in $\mathscr{F}_T$ to the σ-algebra $\sigma(\mathscr{D})$ generated by $\mathscr{D}$. And, according to Proposition (ii) of this theorem, $\mathscr{C}$ being prior in $\mathscr{F}_T$ to $\mathscr{D}$ is equivalent to $\sigma(\mathscr{C})$ being prior in $\mathscr{F}_T$ to $\mathscr{D}$.

Remark 3.10 [The σ-Algebra Generated by the Union of Two Set Systems] Now we turn to some properties of the prior-to relation of set systems involving the σ-algebra generated by the union of two set systems. This is of interest because

$$\sigma(X,Y) = \sigma\big(\sigma(X) \cup \sigma(Y)\big) \tag{3.20}$$

(see RS-Lem. 2.15). According to this equation, the union of the σ-algebras $\sigma(X)$ and $\sigma(Y)$ generated by the measurable maps X and Y on a measurable space $(\Omega, \mathscr{A})$ generates the σ-algebra generated by the bivariate measurable map (X,Y) (see RS-sect. 2.1.4).

Note that, if $\mathscr{C}$ is a σ-algebra on Ω, then

$$\sigma(X) \cup \sigma(Y) \subset \mathscr{C} \quad \Leftrightarrow \quad \sigma\big(\sigma(X) \cup \sigma(Y)\big) \subset \mathscr{C} \tag{3.21}$$

(see RS-Rem. 1.9). Therefore, and because a filtration $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ is a family of σ-algebras, *all properties of the prior-to relation of set systems involving the σ-algebra generated by the union of two set systems can be translated to corresponding properties involving a bivariate measurable map or a bivariate random variable* (see RS-sect. 2.1.4). ◁

In the following theorem we use the notation

$$\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \quad :\Leftrightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \;\wedge\; \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}. \tag{3.22}$$

Theorem 3.11 [Properties Involving the Union of Set Systems]
Let $(\Omega, \mathscr{A})$ *be a measurable space,* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$*, and* $\mathscr{C}, \mathscr{D}, \mathscr{E} \subset \mathscr{A}$*. Then:*

(i) $\big(\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \;\wedge\; \exists\, t \in T\colon \mathscr{E} \subset \mathscr{F}_t\big) \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{D} \cup \mathscr{E})$

(ii) $\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \quad \Leftrightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\precsim} \mathscr{E}.$

(Proof p. 71)

Hence, according to Proposition (i) of Theorem 3.11, if $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$ and $\mathscr{E}$ *is in the filtration* (i.e., there is a $t \in T$ with $\mathscr{E} \subset \mathscr{F}_t$), then $\mathscr{C}$ is also prior in $\mathscr{F}_T$ to $\sigma(\mathscr{D} \cup \mathscr{E})$. Furthermore, $\mathscr{C}$ and $\mathscr{D}$ being prior in $\mathscr{F}_T$ to $\mathscr{E}$ is equivalent to $\sigma(\mathscr{C} \cup \mathscr{D})$ being prior in $\mathscr{F}_T$ to $\mathscr{E}$ [see Prop. (ii)].

Remark 3.12 [A Special Case] Note that $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$ implies the existence of a $t \in T$ with $\mathscr{E} \subset \mathscr{F}_t$ [see Prop. (b) of Def. 3.3 (i)]. Hence,

$$\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}, \mathscr{E} \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{D} \cup \mathscr{E}) \tag{3.23}$$

is an immediate implication of Theorem 3.11 (i). In this proposition, the premise is a shortcut for $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \wedge \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$. Hence, if $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$ and to $\mathscr{E}$, then $\mathscr{C}$ is also prior in $\mathscr{F}_T$ to the σ-algebra generated by the union of $\mathscr{D}$ and $\mathscr{E}$. ◁

According to the following theorem, $\mathscr{C}$ being prior in $\mathscr{F}_T$ to $\mathscr{D}$ is equivalent to $\mathscr{C}$ being prior in $\mathscr{F}_T$ to the σ-algebra generated by the union of $\mathscr{C}$ and $\mathscr{D}$.

Theorem 3.13 [Another Property Involving the Union of Set Systems]
Let $(\Omega, \mathscr{A})$ *be a measurable space,* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$*, and* $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$*. Then:*

$$\mathscr{C} \underset{\mathscr{F}_T}{\lessdot} \mathscr{D} \quad \Leftrightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\lessdot} \sigma(\mathscr{C} \cup \mathscr{D}). \tag{3.24}$$

(Proof p. 71)

Example 3.14 [Joe and Ann With Self-Selection] In Example 3.2, we already specified the probability space $(\Omega, \mathscr{A}, P)$, the random variables U, X, Y, and the filtration $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ in $\mathscr{A}$, $T = \{1, 2, 3\}$, by

$$\mathscr{F}_1 = \sigma(U), \qquad \mathscr{F}_2 = \sigma(U, X), \qquad \mathscr{F}_3 = \sigma(U, X, Y).$$

Remember, $\sigma(X, Y) = \sigma\big(\sigma(X) \cup \sigma(Y)\big)$ [see Eq. (3.20)], which implies $\sigma(X), \sigma(Y) \subset \sigma(X, Y)$ [see RS-Prop. (1.9)]. Similarly, $\sigma(U, X, Y) = \sigma\big(\sigma(U) \cup \sigma(X) \cup \sigma(Y)\big)$, and this implies $\sigma(U)$, $\sigma(X, Y) \subset \sigma(U, X, Y)$.

In this example,

$$\sigma(U) \underset{\mathscr{F}_T}{\lessdot} \sigma(X), \sigma(Y), \sigma(X, Y), \sigma(U, X, Y),$$

because $\sigma(U) \subset \mathscr{F}_1$ and $\sigma(X), \sigma(Y), \sigma(X, Y), \sigma(U, X, Y) \not\subset \mathscr{F}_1$, but $\sigma(X) \subset \mathscr{F}_2$ and $\sigma(Y)$, $\sigma(X, Y)$, $\sigma(U, X, Y) \subset \mathscr{F}_3$ [see Def. 3.3 (i)]. That is, $\sigma(U)$ is prior in $\mathscr{F}_T$ to $\sigma(X)$, prior to $\sigma(Y)$, prior to $\sigma(X, Y)$, and prior to $\sigma(U, X, Y)$.

Finally,

$$\sigma(X) \underset{\mathscr{F}_T}{\lessdot} \sigma(Y), \sigma(X, Y), \sigma(U, X, Y),$$

because $\sigma(X) \subset \mathscr{F}_2$ and $\sigma(Y), \sigma(X, Y), \sigma(U, X, Y) \not\subset \mathscr{F}_2$, but $\sigma(Y), \sigma(X, Y), \sigma(U, X, Y) \subset \mathscr{F}_3$ [see again Def. 3.3 (i)]. Hence, $\sigma(X)$ is prior in $\mathscr{F}_T$ to $\sigma(Y)$, prior to $\sigma(X, Y)$, and prior to $\sigma(U, X, Y)$. ◁

3.2.2 *Properties of the Prior-to Relation of Measurable Maps*

Now we will translate the properties of the prior-to relation of measurable set systems to the prior-to relation of measurable maps. For a full understanding, the following remarks on the measurability of a composition are important.

Remark 3.15 [Composition of Two Maps] Let $X: \Omega \to \Omega'_X$ and $g: \Omega'_X \to \Omega'_g$ be maps. Then a map $W: \Omega \to \Omega'_g$ is called the *composition of X and g*, denoted $g \circ X$ or $g(X)$ if

$$W(\omega) = g\big(X(\omega)\big), \quad \forall\, \omega \in \Omega. \tag{3.25}$$

◁

The following lemma is useful whenever we deal with the composition of two measurable maps. (For a proof, see RS-Theorem 2.11, RS-Definition 2.12, RS-Remark 2.13, and Lemma 2.34).

Lemma 3.16 [Measurability of a Composition]
Let $X\colon (\Omega, \mathscr{A}) \to (\Omega'_X, \mathscr{A}'_X)$ *and* $g\colon (\Omega'_X, \mathscr{A}'_X) \to (\Omega'_g, \mathscr{A}'_g)$ *be measurable maps and assume* $W = g(X)$*. Then W is X-measurable, that is, then* $\sigma(W) \subset \sigma(X)$*.*

According to this lemma, we can conclude that W is X-measurable, whenever it is the composition of a measurable map X and another measurable map g. Note that no requirements are necessary for the three measurable spaces involved. However, this lemma only contains a *sufficient* condition of X-measurability of W.

Remark 3.17 [Measurability of a Composition of X and g] Now we add a requirement for the measurable space $(\Omega'_g, \mathscr{A}'_g)$. This yields a *necessary and sufficient* condition of X-measurability of W. Let $X\colon (\Omega, \mathscr{A}) \to (\Omega'_X, \mathscr{A}'_X)$ and $W\colon (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ be measurable maps on a measurable space $(\Omega, \mathscr{A})$. Then $\sigma(W) \subset \sigma(X)$ if and only if there is a measurable map $g\colon (\Omega'_X, \mathscr{A}'_X) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ such that $W = g(X)$ is the composition of X and g (see RS-Lem. 2.35). Hence, *if W is numerical, then it is X-measurable if and only if it can be written as the composition of X and g*. ◁

Example 3.18 [Some Simple Examples] If $W = \alpha \cdot X$, $\alpha \in \mathbb{R}$, then W is X-measurable (see Exercise 3-9). Another example is $W = 1_{X=x}$, the indicator of X taking on the value x. Note that, in Lemma 3.16 and Remark 3.17, X may also be a multidimensional map. For example, if $W = \beta_1 X_1 + \beta_2 X_2$ with $\beta_1, \beta_2 \in \mathbb{R}$, then W is (X_1, X_2)-measurable, that is, W is measurable with respect to the bivariate map $X = (X_1, X_2)$ (see Exercise 3-10). Another example is the product $W = 1_{X=x} \cdot 1_{Y=y}$ of the indicators of X taking on the value x and Y taking on the value y. In this example, W is (X, Y)-measurable and $(1_{X=x}, 1_{Y=y})$-measurable. ◁

Now we turn to properties of the prior-to relation of measurable maps; the most important ones are gathered in Box 3.1. Because random variables on a probability space $(\Omega, \mathscr{A}, P)$ are measurable maps on the measurable space $(\Omega, \mathscr{A})$, all propositions about measurable maps in this box and in the present section also hold for random variables.

Box 3.1 starts repeating the definition of the prior-to relation of measurable maps. This definition is based on the framework of a measurable space $(\Omega, \mathscr{A})$ and a filtration $\mathscr{F}_T$ in $\mathscr{A}$. Proposition (i) of Box 3.1 is the *asymmetry* property of the prior-to relation of measurable maps. According to this proposition, if X is prior in $\mathscr{F}_T$ to Y, then Y is not prior in $\mathscr{F}_T$ to X. This proposition immediately follows from Theorem 3.6 (i) because the prior-to relation of measurable maps is defined via their generated σ-algebras [see Def. 3.3 (ii)].

According to Proposition (ii) of Box 3.1, X is prior in $\mathscr{F}_T$ to Y if and only if X is prior in $\mathscr{F}_T$ to (X, Y), where (X, Y) denotes the bivariate measurable map on $(\Omega, \mathscr{A})$ that consists of X and Y (for more details see RS-sect. 2.1.4). This proposition is an immediate implication of Theorem 3.13 and Definition 3.3 (ii).

According to Propositions (iii) and (iv) of this box, if W is X-measurable, then X being prior in $\mathscr{F}_T$ to Y implies that W is also prior in $\mathscr{F}_T$ to Y and to the bivariate map (X, Y). Proposition (iii) is an immediate implication of Theorem 3.8, whereas Proposition (iv) follows from Theorem 3.13, Proposition (3.20), and Theorem 3.8 [see Exercise 3-11].

Now we turn to Propositions (v) and (vi) of Box 3.1, which involve an additional measurable map Z on the measurable space $(\Omega, \mathscr{A})$. We prove these propositions in the following theorem. In this theorem, we use the notation

Box 3.1 Properties of the prior-to relation of measurable maps

Let X and Y be measurable maps on a measurable space $(\Omega, \mathscr{A})$, let $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ be a filtration in $\mathscr{A}$, and let $\sigma(X)$ and $\sigma(Y)$ denote the σ-algebras generated by X and Y, respectively. Then we say that *X is prior in $\mathscr{F}_T$ to Y* (and *Y posterior in $\mathscr{F}_T$ to X*), denoted $X \underset{\mathscr{F}_T}{\prec} Y$, if the following two conditions hold:

(a) $\exists\, s \in T\colon \sigma(X) \subset \mathscr{F}_s \wedge \sigma(Y) \not\subset \mathscr{F}_s$

(b) $\exists\, t \in T\colon \sigma(Y) \subset \mathscr{F}_t$.

A first property is

$$X \underset{\mathscr{F}_T}{\prec} Y \quad \Rightarrow \quad \neg(Y \underset{\mathscr{F}_T}{\prec} X). \qquad \text{(asymmetry)} \tag{i}$$

Additionally, let (X, Y) denote the bivariate measurable map consisting of X and Y. Then:

$$X \underset{\mathscr{F}_T}{\prec} Y \quad \Leftrightarrow \quad X \underset{\mathscr{F}_T}{\prec} (X, Y). \tag{ii}$$

Additionally, let also W be a measurable map on $(\Omega, \mathscr{A})$. Then:

$$X \underset{\mathscr{F}_T}{\prec} Y \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\prec} Y, \qquad \text{if } \sigma(W) \subset \sigma(X) \tag{iii}$$

$$X \underset{\mathscr{F}_T}{\prec} Y \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\prec} (X, Y), \qquad \text{if } \sigma(W) \subset \sigma(X). \tag{iv}$$

Additionally, let also Z be a measurable map on $(\Omega, \mathscr{A})$, let (Y, Z) denote the bivariate map consisting of Y and Z, and $\sigma(X, Y)$ the σ-algebra generated by (X, Y). Then:

$$X, Y \underset{\mathscr{F}_T}{\prec} Z \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\prec} Z, \qquad \text{if } \sigma(W) \subset \sigma(X, Y) \tag{v}$$

$$X \underset{\mathscr{F}_T}{\prec} Y \wedge Y \underset{\mathscr{F}_T}{\prec} Z \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\prec} Z, \qquad \text{if } \sigma(W) \subset \sigma(X, Y). \tag{vi}$$

Additionally, let $\sigma(Z)$ denote the σ-algebra generated by Z. Then:

$$\big(X \underset{\mathscr{F}_T}{\prec} Y \wedge \exists\, t \in T\colon \sigma(Z) \subset \mathscr{F}_t\big) \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\prec} (Y, Z), \qquad \text{if } \sigma(W) \subset \sigma(X). \tag{vii}$$

$$X, Y \underset{\mathscr{F}_T}{\prec} Z \quad :\Leftrightarrow \quad (X \underset{\mathscr{F}_T}{\prec} Z \wedge Y \underset{\mathscr{F}_T}{\prec} Z). \tag{3.26}$$

Theorem 3.19 [Further Implications]
Let W, X, Y, Z be measurable maps on a measurable space $(\Omega, \mathscr{A})$ and $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$. Then:

(i) $X, Y \underset{\mathscr{F}_T}{\prec} Z \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\prec} Z, \qquad$ *if* $\sigma(W) \subset \sigma(X, Y)$.

(ii) $(X \underset{\mathscr{F}_T}{\prec} Y \wedge Y \underset{\mathscr{F}_T}{\prec} Z) \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\prec} Z, \qquad$ *if* $\sigma(W) \subset \sigma(X, Y)$.

(Proof p. 72)

Hence, according to Proposition (v) of Box 3.1, if X and Y are prior in $\mathscr{F}_T$ to Z, then each (X, Y)-measurable map W is prior in $\mathscr{F}_T$ to Z as well. Note that each X-measurable map W is also (X, Y)-measurable, that is,

$$\sigma(W) \subset \sigma(X) \quad \Rightarrow \quad \sigma(W) \subset \sigma(X, Y). \tag{3.27}$$

Furthermore, according to Proposition (vi) of Box 3.1, if X is prior in $\mathscr{F}_T$ to Y that itself is prior in $\mathscr{F}_T$ to Z, then each (X,Y)-measurable map W is also prior in $\mathscr{F}_T$ to Z.

Remark 3.20 [Two Special Cases] For $W = X$, Proposition (ii) of Theorem 3.19 yields

$$(X \underset{\mathscr{F}_T}{\lessapprox} Y \;\wedge\; Y \underset{\mathscr{F}_T}{\lessapprox} Z) \quad \Rightarrow \quad X \underset{\mathscr{F}_T}{\lessapprox} Z. \qquad \text{(transitivity)} \tag{3.28}$$

Hence, if X is prior in $\mathscr{F}_T$ to Y and Y is prior to Z, then X is prior in $\mathscr{F}_T$ to Z as well. This is the *transitivity* property of the prior-to relation of measurable maps.

Furthermore, for $W = (X,Y)$, Proposition (ii) of Theorem 3.19 yields

$$(X \underset{\mathscr{F}_T}{\lessapprox} Y \;\wedge\; Y \underset{\mathscr{F}_T}{\lessapprox} Z) \quad \Rightarrow \quad (X,Y) \underset{\mathscr{F}_T}{\lessapprox} Z. \tag{3.29}$$

Hence, if X is prior in $\mathscr{F}_T$ to Y and Y is prior to Z, then the bivariate measurable map (X,Y) is prior in $\mathscr{F}_T$ to Z as well. ◁

Now we turn to Proposition (vii) of Box 3.1, which is proved in the following theorem.

Theorem 3.21 [Another Implication of $X \underset{\mathscr{F}_T}{\lessapprox} Y$ for an X-Measurable Map]
Let W, X, Y, Z be measurable maps on a measurable space $(\Omega, \mathscr{A})$ and $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$. Then

$$\big(X \underset{\mathscr{F}_T}{\lessapprox} Y \;\wedge\; \exists\, t \in T\colon \sigma(Z) \subset \mathscr{F}_t\big) \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\lessapprox} (Y,Z), \qquad \textit{if } \sigma(W) \subset \sigma(X). \tag{3.30}$$

(Proof p. 72)

Hence, if X is prior in $\mathscr{F}_T$ to Y and Z is in the filtration $\mathscr{F}_T$, then each X-measurable map W is prior in $\mathscr{F}_T$ to the bivariate map (Y,Z).

Remark 3.22 [More Implications of $X \underset{\mathscr{F}_T}{\lessapprox} Y$] For $W = X$, Theorem 3.21 implies

$$\big(X \underset{\mathscr{F}_T}{\lessapprox} Y \;\wedge\; \exists\, t \in T\colon \sigma(Z) \subset \mathscr{F}_t\big) \quad \Rightarrow \quad X \underset{\mathscr{F}_T}{\lessapprox} (Y,Z). \tag{3.31}$$

Hence, if X is prior in $\mathscr{F}_T$ to Y and Z is in the filtration $\mathscr{F}_T$, then X is prior to the bivariate map (Y,Z). ◁

Example 3.23 [Joe and Ann With Self-Selection] In Example 3.14, we already showed that

$$\sigma(U) \underset{\mathscr{F}_T}{\lessapprox} \sigma(X), \sigma(Y), \sigma(X,Y), \sigma(U,X,Y).$$

and

$$\sigma(X) \underset{\mathscr{F}_T}{\lessapprox} \sigma(Y), \sigma(X,Y), \sigma(U,X,Y).$$

Therefore, according to Definition 3.3 (ii),

$$U \underset{\mathscr{F}_T}{\lessapprox} X,\; Y,\; (X,Y),\; (U,X,Y). \tag{3.32}$$

That is, U is prior in $\mathscr{F}_T$ to X, prior in $\mathscr{F}_T$ to Y, and prior in $\mathscr{F}_T$ to (the multivariate random variables) (X,Y) and (U,X,Y). Furthermore,

$$X \underset{\mathscr{F}_T}{\precsim} Y,\ (X,Y),\ (U,X,Y), \tag{3.33}$$

that is, X is prior in $\mathscr{F}_T$ to Y, prior in $\mathscr{F}_T$ to the bivariate random variable (X,Y), and prior in $\mathscr{F}_T$ to the trivariate random variable (U,X,Y).

In order to illustrate some propositions of Box 3.1, consider the indicator $1_{U=Joe}$ of the event that *Joe is sampled*. Because $1_{U=Joe}$ is U-measurable, according to Proposition (3.32) and Box 3.1 (iii),

$$1_{U=Joe} \underset{\mathscr{F}_T}{\precsim} X,\ Y,\ (X,Y),\ (U,X,Y). \tag{3.34}$$

Furthermore, according to Propositions (3.32), (3.33), and Box 3.1 (vi),

$$1_{U=Joe} \underset{\mathscr{F}_T}{\precsim} Y,\ (X,Y),\ (U,X,Y). \tag{3.35}$$

Finally, consider the product $1_{U=Joe} \cdot X$, which is the indicator of the event that *Joe is sampled and treated*. (Note that, in this example, X is an indicator, too.) This indicator is (U,X)-measurable, that is, $\sigma(1_{U=Joe} \cdot X) \subset \sigma(U,X)$. Hence, according to (3.32), (3.33), and Box 3.1 (v), the random variable $1_{U=Joe} \cdot X$ is prior in $\mathscr{F}_T$ to Y. $\lhd$

3.3 Simultaneous-to Relations

In chapter 2, we already discussed that we consider as potential confounders of X all those random variables that are *prior* or *simultaneous* to the focused putative cause variable X. As an example of a potential confounder that is simultaneous to X we mentioned a second treatment variable that is manipulated at the same time as X. As another example, consider studying the effects of X, the *amount of antibodies at time* t. In such a case, we might also consider Z, the *amount of leucocites at time* t, referring to the same time point t. Then X and Z would be simultaneous to each other.

Again, we start defining the simultaneous-to relation for measurable set systems $\mathscr{C}$ and $\mathscr{D}$, and extend it to measurable maps X and Y via their generated σ-algebras $\sigma(X)$ and $\sigma(Y)$.

Definition 3.24 [Simultaneous-to Relations]
Let $(\Omega,\mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t\in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C},\mathscr{D} \subset \mathscr{A}$.

(i) We say that $\mathscr{C}$ and $\mathscr{D}$ are simultaneous in (or, with respect to) $\mathscr{F}_T$, denoted $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$, if the following two conditions hold:

(a) $\exists\, t \in T$: $\mathscr{C} \subset \mathscr{F}_t$

(b) $\forall\, t \in T$: $\mathscr{C} \subset \mathscr{F}_t \;\Leftrightarrow\; \mathscr{D} \subset \mathscr{F}_t$.

(ii) Let X and Y be measurable maps on $(\Omega,\mathscr{A})$. Then we say that X and Y are simultaneous in $\mathscr{F}_T$, denoted $X \underset{\mathscr{F}_T}{\approx} Y$, if $\sigma(X)$ and $\sigma(Y)$ are simultaneous in $\mathscr{F}_T$.

The simultaneous-to relation always refers to a measurable space $(\Omega, \mathscr{A})$ and a filtration $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ in $\mathscr{A}$. For $\mathscr{C}$ and $\mathscr{D}$ to be simultaneous in $\mathscr{F}_T$, or, synonymously, for *$\mathscr{C}$ to be simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$*, we require that there is an element $t \in T$ such that the set system $\mathscr{C}$ is a subset of $\mathscr{F}_t$ [see Def. 3.24 (i) (a)]. In this case we also say that *$\mathscr{C}$ is in the filtration $\mathscr{F}_T$*. Furthermore, in condition (b) of this definition, we require that $\mathscr{C}$ is a subset of $\mathscr{F}_t$ if and only if $\mathscr{D}$ is a subset of $\mathscr{F}_t$, for all $t \in T$. Note that the conjunction of (a) and (b) implies that $\mathscr{D}$ is in the filtration $\mathscr{F}_T$ as well, that is, there is a $t \in T$ such that $\mathscr{D} \subset \mathscr{F}_t$. Note again that, in the context of a probability space $(\Omega, \mathscr{A}, P)$, a measurable map is a *random variable* on the probability space $(\Omega, \mathscr{A}, P)$, and measurable set systems $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$ are *sets of events*.

Remark 3.25 [Simultaneous-to Relation of Measurables Sets and of Events] Let $(\Omega, \mathscr{A})$ be a measurable space and $\{C\}$, $\{D\}$ set systems containing the sets $C, D \in \mathscr{A}$ as their only element. We say that *C and D are simultaneous in $\mathscr{F}_T$*, denoted $C \underset{\mathscr{F}_T}{\approx} D$, if $\{C\}$ and $\{D\}$ are simultaneous in $\mathscr{F}_T$. Using this definition, all propositions about the simultaneous-to relation of measurable set systems are easily translated to the simultaneous-to relation of measurable sets. Again remember, in the context of a probability space $(\Omega, \mathscr{A}, P)$, the sets C and D represent events. Again, for brevity, we will not treat any details of the simultaneous-to relation of measurable sets. ◁

Example 3.26 [Joe and Ann With Self-Selection] In Example 3.2, we already specified the probability space $(\Omega, \mathscr{A}, P)$, the random variables U, X, Y, and the filtration $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$. In this example, the person variable U and the event $\{Joe\} \times \Omega_X \times \Omega_Y$ that Joe is sampled are simultaneous in $\mathscr{F}_T$. This is because the set system $\big\{\{Joe\} \times \Omega_X \times \Omega_Y\big\}$ is a subset of the σ-algebra generated by U and because the first σ-algebra $\mathscr{F}_1$ in the filtration $\mathscr{F}_T$ has been defined to be the σ-algebra generated by U. Hence, conditions (a) and (b) of Definition 3.24 (i) hold for $\mathscr{C} = \sigma(U)$ and $\mathscr{D} = \big\{\{Joe\} \times \Omega_X \times \Omega_Y\big\}$. ◁

3.3.1 *Properties of the Simultaneous-to Relation of Measurable Set Systems*

Now we study some elementary properties of the simultaneous-to relation. First of all, we show that this relation is *reflexive*, *symmetric* and *transitive*, that is, we show that it is an *equivalence relation*.

Theorem 3.27 [Reflexivity, Symmetry, and Transitivity]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D}, \mathscr{E} \subset \mathscr{A}$. Then:

(i) $\exists\, t \in T\colon \mathscr{C} \subset \mathscr{F}_t \;\Rightarrow\; \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{C}$ *(reflexivity)*

(ii) $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \;\Rightarrow\; \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{C}$ *(symmetry)*

(iii) $(\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \,\wedge\, \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) \;\Rightarrow\; \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{E}.$ *(transitivity)*

(Proof p. 73)

Now we turn to some properties of the simultaneous-to relation involving the σ-algebras $\sigma(\mathscr{C})$ and $\sigma(\mathscr{C} \cup \mathscr{D})$ generated by the set systems $\mathscr{C}$ and $\mathscr{C} \cup \mathscr{D}$, respectively (see RS-Def. 1.7). The motivation is to obtain some properties of the simultaneous-to relation of measurable maps and random variables (see again Rem. 3.10).

Theorem 3.28 [Properties of the Simultaneous-to Relation Involving σ-Algebras] *Let* $(\Omega, \mathscr{A})$ *be a measurable space,* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$*, and* $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$*. Then:*

(i) $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \quad \Leftrightarrow \quad \sigma(\mathscr{C}) \underset{\mathscr{F}_T}{\approx} \mathscr{D}$

(ii) $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \quad \Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{D}.$

(Proof p. 73)

Hence, according to Theorem 3.28, the set systems $\mathscr{C}$ and $\mathscr{D}$ being simultaneous in $\mathscr{F}_T$ is equivalent to $\sigma(\mathscr{C})$ (i.e., the σ-algebra generated by $\mathscr{C}$) being simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$ [see Prop. (i)]. Furthermore, $\mathscr{C}$ and $\mathscr{D}$ being simultaneous in $\mathscr{F}_T$ implies that $\sigma(\mathscr{C} \cup \mathscr{D})$ (i.e., the σ-algebra generated by the union $\mathscr{C} \cup \mathscr{D}$) is simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$ [see Prop. (ii)].

In the following theorem we treat a property of the simultaneous-to relation involving a third measurable set system $\mathscr{E} \subset \mathscr{A}$.

Theorem 3.29 [A Proposition Involving a Third Measurable Set Systems] *Let* $(\Omega, \mathscr{A})$ *be a measurable space,* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$*, and* $\mathscr{C}, \mathscr{D}, \mathscr{E} \subset \mathscr{A}$*. Then*

$$(\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) \quad \Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{E}. \tag{3.36}$$

(Proof p. 74)

Hence, if the set systems $\mathscr{C}, \mathscr{D}$ and $\mathscr{E}$ are simultaneous in $\mathscr{F}_T$ to each other, then the σ-algebra generated by the union of $\mathscr{C}$ and $\mathscr{D}$ is simultaneous in $\mathscr{F}_T$ to $\mathscr{E}$.

Example 3.30 [Joe and Ann With Self-Selection] In Example 3.2, we already specified the probability space $(\Omega, \mathscr{A}, P)$, the random variables U, X, Y, and the filtration $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ for the random experiment presented in Table 1.2. In this example,

$$\sigma(X) \underset{\mathscr{F}_T}{\approx} \sigma(U, X),$$

and

$$\sigma(Y) \underset{\mathscr{F}_T}{\approx} \sigma(X, Y), \sigma(U, X, Y).$$

That is, $\sigma(X)$ is simultaneous in $\mathscr{F}_T$ to $\sigma(U, X)$, and $\sigma(Y)$ is simultaneous in $\mathscr{F}_T$ to $\sigma(X, Y)$ and to $\sigma(U, X, Y)$. Therefore, according to Definition 3.3 (ii),

$$X \underset{\mathscr{F}_T}{\approx} (U, X)$$

and

$$Y \underset{\mathscr{F}_T}{\approx} (X, Y), (U, X, Y).$$

That is, X is simultaneous in $\mathscr{F}_T$ to (the multivariate random variable) (U, X), and Y is simultaneous in $\mathscr{F}_T$ to (X, Y) and to (U, X, Y). ◁

Now we treat some theorems involving the prior-to and the simultaneous-to relations. In the first one we show that the set system $\mathscr{C}$ being prior in $\mathscr{F}_T$ to the set system $\mathscr{D}$ implies that $\sigma(\mathscr{C} \cup \mathscr{D})$, the σ-algebra generated by the union of $\mathscr{C}$ and $\mathscr{D}$, is simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$ and that $\mathscr{C}$ being simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$ implies that $\mathscr{C}$ is not prior in $\mathscr{F}_T$ to $\mathscr{D}$.

Theorem 3.31 [Combining the Prior-to and Simultaneous-to Relations]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$. Then:

(i) $\mathscr{C} \underset{\mathscr{F}_T}{\lessdot} \mathscr{D} \quad \Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{D}$

(ii) $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \quad \Rightarrow \quad \neg(\mathscr{C} \underset{\mathscr{F}_T}{\lessdot} \mathscr{D}).$

(Proof p. 74)

In the following theorem we study an implication of $\mathscr{C}$ and $\mathscr{D}$ being prior or simultaneous in $\mathscr{F}_T$ for a subset $\mathscr{C}_0$ of the σ-algebra generated by the union of $\mathscr{C}$ and $\mathscr{D}$.

Theorem 3.32 [Combining the Prior-to and Simultaneous-to Relations]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$. If $\mathscr{C}_0 \subset \sigma(\mathscr{C} \cup \mathscr{D})$, then:

(i) $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \quad \Rightarrow \quad (\mathscr{C}_0 \underset{\mathscr{F}_T}{\lessdot} \mathscr{D} \vee \mathscr{C}_0 \underset{\mathscr{F}_T}{\approx} \mathscr{D})$

(ii) $\mathscr{C} \underset{\mathscr{F}_T}{\lessdot} \mathscr{D} \quad \Rightarrow \quad (\mathscr{C}_0 \underset{\mathscr{F}_T}{\lessdot} \mathscr{D} \vee \mathscr{C}_0 \underset{\mathscr{F}_T}{\approx} \mathscr{D}).$

(Proof p. 74)

Hence, according to the two propositions of this theorem, if the set system $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to the set system $\mathscr{D}$ and $\mathscr{C}_0$ is a subset of $\sigma(\mathscr{C} \cup \mathscr{D})$, then $\mathscr{C}_0$ is prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$.

Finally, we show: If $\mathscr{C}$ is simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$ that itself is prior in $\mathscr{F}_T$ to $\mathscr{E}$, then $\mathscr{C}_0$ is prior in $\mathscr{F}_T$ to $\mathscr{E}$, provided that $\mathscr{C}_0$ is a subset of $\sigma(\mathscr{C} \cup \mathscr{D})$. Furthermore, if $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$ that itself is simultaneous in $\mathscr{F}_T$ to $\mathscr{E}$, and $\mathscr{C}_0$ is a subset of $\mathscr{C}$, then we can conclude that $\mathscr{C}_0$ is prior in $\mathscr{F}_T$ to $\mathscr{E}$.

Theorem 3.33 [Combining the Prior-to and Simultaneous-to Relations]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D}, \mathscr{E} \subset \mathscr{A}$. Then:

(i) $(\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\lessdot} \mathscr{E}) \quad \Rightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\lessdot} \mathscr{E}, \quad$ *if* $\mathscr{C}_0 \subset \sigma(\mathscr{C} \cup \mathscr{D})$

(ii) $(\mathscr{C} \underset{\mathscr{F}_T}{\lessdot} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) \quad \Rightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\lessdot} \mathscr{E}, \quad$ *if* $\mathscr{C}_0 \subset \mathscr{C}$.

(Proof p. 75)

In the following remark, we consider the special case of Theorem 3.33 in which $\mathscr{C}_0 = \mathscr{C}$.

Remark 3.34 [A Special Case] Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D}, \mathscr{E} \subset \mathscr{A}$. Then:

(i) $(\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\lessdot} \mathscr{E}) \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\lessdot} \mathscr{E}$

(ii) $(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E}$.

Hence, if the set system $\mathscr{C}$ is simultaneous in $\mathscr{F}_T$ to the set system $\mathscr{D}$ that itself is prior in $\mathscr{F}_T$ to the set system $\mathscr{E}$, then $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{E}$ as well. Furthermore, if $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$ and $\mathscr{D}$ is simultaneous in $\mathscr{F}_T$ to $\mathscr{E}$, then $\mathscr{C}$ is also prior in $\mathscr{F}_T$ to $\mathscr{E}$. ◁

3.3.2 Properties of the Simultaneous-to Relation of Measurable Maps

Now we turn to properties of the simultaneous-to relation of measurable maps, the most important of which are gathered in Box 3.2. Again remember, because random variables on a probability space $(\Omega, \mathscr{A}, P)$ are measurable maps on the measurable space $(\Omega, \mathscr{A})$, all propositions about the simultaneous-to relation of measurable maps in this box and in this section also hold for random variables.

Box 3.2 starts repeating the definition of the simultaneous-to relation of measurable maps. Propositions (i) (*reflexivity*) and (ii) (*symmetry*) of Box 3.2 immediately follow from Propositions (i) and (ii) of Theorem 3.27 because the simultaneous-to relation of measurable maps is defined via their generated σ-algebras [see Def. 3.24 (ii)].

Proposition (iii) of Box 3.2 is an immediate implication of Theorem 3.31 (ii) because the simultaneous-to and prior-to relations of measurable maps are defined via the corresponding relation of their generated σ-algebras [see Defs. 3.3 (ii) and 3.24 (ii)]. According to this proposition, if X is simultaneous in $\mathscr{F}_T$ to Y, then X is not prior in $\mathscr{F}_T$ to Y, nor is Y prior in $\mathscr{F}_T$ to X (which follows from symmetry).

According to Propositions (iv) and (v) of Box 3.2, X being simultaneous or prior in $\mathscr{F}_T$ to Y implies that the bivariate map (X, Y) is also simultaneous in $\mathscr{F}_T$ to Y. The first of these propositions immediately follows from Theorem 3.28 (ii), Definition 3.24 (ii), and Equation (3.20), the second from Theorem 3.31 (i), Definition 3.24 (ii), and Equation (3.20).

Proposition (vi) of Box 3.2 is called *transitivity* of the simultaneous-to relation of measurable maps. According to this proposition, if X is simultaneous in $\mathscr{F}_T$ to Y and Y simultaneous in $\mathscr{F}_T$ to Z, then X is also simultaneous in $\mathscr{F}_T$ to Z. This proposition immediately follows from Theorem 3.27 (iii), Definition 3.24 (ii), and Equation (3.20).

According to Proposition (vii) of Box 3.2, X being simultaneous in $\mathscr{F}_T$ to Y and Y being simultaneous in $\mathscr{F}_T$ to Z implies that the bivariate map (X, Y) is also simultaneous in $\mathscr{F}_T$ to Z. This proposition immediately follows from Theorem 3.29, Definition 3.24 (ii), and Equation (3.20). Symmetry immediately yields

$$\big(X \underset{\mathscr{F}_T}{\approx} Y \wedge Y \underset{\mathscr{F}_T}{\approx} Z\big) \quad \Rightarrow \quad (X, Z) \underset{\mathscr{F}_T}{\approx} Y. \tag{3.37}$$

According to Propositions (viii) and (ix) of Box 3.2, X being prior or simultaneous in $\mathscr{F}_T$ to Y implies that each (X, Y)-measurable map W is prior or simultaneous in $\mathscr{F}_T$ to Y (see again Rem. 3.15 to Example 3.18). These two propositions are proved in the following theorem.

Theorem 3.35 [Two Propositions Involving the Prior-to Relation]
Let W, X, Y be measurable maps on a measurable space $(\Omega, \mathscr{A})$, let $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ be a filtration in $\mathscr{A}$, and assume $\sigma(W) \subset \sigma(X, Y)$. Then

(i) $X \underset{\mathscr{F}_T}{\prec} Y \quad \Rightarrow \quad (W \underset{\mathscr{F}_T}{\prec} Y \vee W \underset{\mathscr{F}_T}{\approx} Y)$

Box 3.2 Properties of the simultaneous-to relation of measurable maps

Let X and Y be measurable maps on a measurable space $(\Omega, \mathscr{A})$, let $\mathscr{F}_T = (\mathscr{F}_t)_{t\in T}$ be a filtration in $\mathscr{A}$, and let $\sigma(X)$ and $\sigma(Y)$ denote the σ-algebras generated X and Y, respectively. Then we say that *X is simultaneous to Y* in $\mathscr{F}_T$, denoted $X \underset{\mathscr{F}_T}{\approx} Y$, if the following two conditions hold:

(a) $\exists\, t \in T\colon\ \sigma(X) \subset \mathscr{F}_t$

(b) $\forall\, t \in T\colon\ \sigma(X) \subset \mathscr{F}_t \ \Leftrightarrow\ \sigma(Y) \subset \mathscr{F}_t$.

Some first properties are

$$\exists\, t \in T\colon \sigma(X) \subset \mathscr{F}_t \quad\Rightarrow\quad X \underset{\mathscr{F}_T}{\approx} X \qquad \text{(reflexivity)} \qquad \text{(i)}$$

$$X \underset{\mathscr{F}_T}{\approx} Y \quad\Rightarrow\quad Y \underset{\mathscr{F}_T}{\approx} X \qquad \text{(symmetry)} \qquad \text{(ii)}$$

$$X \underset{\mathscr{F}_T}{\approx} Y \quad\Rightarrow\quad \neg(X \underset{\mathscr{F}_T}{\lessdot} Y). \qquad \text{(iii)}$$

Additionally, let (X, Y) denote the bivariate measurable map consisting of X and Y. Then:

$$X \underset{\mathscr{F}_T}{\approx} Y \quad\Rightarrow\quad (X, Y) \underset{\mathscr{F}_T}{\approx} Y \qquad \text{(iv)}$$

$$X \underset{\mathscr{F}_T}{\lessdot} Y \quad\Rightarrow\quad (X, Y) \underset{\mathscr{F}_T}{\approx} Y. \qquad \text{(v)}$$

Additionally, let also Z denote a measurable map on $(\Omega, \mathscr{A})$. Then:

$$(X \underset{\mathscr{F}_T}{\approx} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\approx} Z) \quad\Rightarrow\quad X \underset{\mathscr{F}_T}{\approx} Z \qquad \text{(transitivity)} \qquad \text{(vi)}$$

$$(X \underset{\mathscr{F}_T}{\approx} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\approx} Z) \quad\Rightarrow\quad (X, Y) \underset{\mathscr{F}_T}{\approx} Z. \qquad \text{(vii)}$$

Additionally, let also W be a measurable map on $(\Omega, \mathscr{A})$. Then:

$$X \underset{\mathscr{F}_T}{\lessdot} Y \quad\Rightarrow\quad (W \underset{\mathscr{F}_T}{\lessdot} Y \ \vee\ W \underset{\mathscr{F}_T}{\approx} Y), \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(viii)}$$

$$X \underset{\mathscr{F}_T}{\approx} Y \quad\Rightarrow\quad (W \underset{\mathscr{F}_T}{\lessdot} Y \ \vee\ W \underset{\mathscr{F}_T}{\approx} Y), \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(ix)}$$

$$(X \underset{\mathscr{F}_T}{\approx} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\lessdot} Z) \quad\Rightarrow\quad W \underset{\mathscr{F}_T}{\lessdot} Z, \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(x)}$$

$$(X \underset{\mathscr{F}_T}{\lessdot} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\approx} Z) \quad\Rightarrow\quad W \underset{\mathscr{F}_T}{\lessdot} Z, \quad \text{if } \sigma(W) \subset \sigma(X). \qquad \text{(xi)}$$

(ii) $X \underset{\mathscr{F}_T}{\approx} Y \quad\Rightarrow\quad (W \underset{\mathscr{F}_T}{\lessdot} Y \ \vee\ W \underset{\mathscr{F}_T}{\approx} Y)$.

(Proof p. 76)

According to Proposition (x) of Box 3.2, if X is simultaneous in $\mathscr{F}_T$ to Y that itself is prior in $\mathscr{F}_T$ to Z, then each (X, Y)-measurable map W is also prior in $\mathscr{F}_T$ to Z. This proposition immediately follows from Theorem 3.33 (i), Definition 3.3 (ii), and Definition 3.24 (ii) (see Exercise 3-12).

Finally, according to Proposition (xi) of Box 3.2, if X is prior in $\mathscr{F}_T$ to Y that itself is simultaneous in $\mathscr{F}_T$ to Z, then each X-measurable map W is also prior in $\mathscr{F}_T$ to Z. This proposition immediately follows from Theorem 3.33 (ii), Definition 3.3 (ii), and Definition 3.24 (ii). Note that X-measurability of W implies that W is (X, Y)-measurable, but not vice versa.

Remark 3.36 [Two Special Cases] For $W = X$, Proposition (x) of this box yields

$$\left(X \underset{\mathscr{F}_T}{\approx} Y \wedge Y \underset{\mathscr{F}_T}{\prec} Z\right) \quad \Rightarrow \quad X \underset{\mathscr{F}_T}{\prec} Z. \tag{3.38}$$

Hence, if X is simultaneous in $\mathscr{F}_T$ to Y that itself is prior in $\mathscr{F}_T$ to Z, then X is prior in $\mathscr{F}_T$ to Z.

Similarly, for $W = X$, Proposition (xi) of Box 3.2 yields

$$\left(X \underset{\mathscr{F}_T}{\prec} Y \wedge Y \underset{\mathscr{F}_T}{\approx} Z\right) \quad \Rightarrow \quad X \underset{\mathscr{F}_T}{\prec} Z. \tag{3.39}$$

According to this proposition, if X is prior in $\mathscr{F}_T$ to Y and Y is simultaneous in $\mathscr{F}_T$ to Z, then X is also prior in $\mathscr{F}_T$ to Z. ◁

3.4 Prior-or-Simultaneous-to Relations

As noted before, the intuitive idea of a potential confounder W of X is that is it a random variable that is *prior or simultaneous to* X. This concept is now defined for measurable maps — and with it for random variables — and for measurable set systems — and with for sets of events. Reading the following definition, remember that the σ-algebra $\sigma(X)$ generated by a measurable map X is a set system (see RS-Def. 2.12).

Definition 3.37 [Prior-or-Simultaneous-to Relations]
Let $(\Omega, \mathscr{A})$ *be a measurable space,* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$*, and* $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$.

(i) *We say that* $\mathscr{C}$ *is prior or simultaneous in* $\mathscr{F}_T$ *to* $\mathscr{D}$*, denoted* $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$*, if* $\mathscr{C}$ *is prior in* $\mathscr{F}_T$ *to* $\mathscr{D}$ *or simultaneous in* $\mathscr{F}_T$ *to* $\mathscr{D}$.

(ii) *Let* X *and* Y *be measurable maps on* $(\Omega, \mathscr{A})$*. Then we say that* X *is prior or simultaneous in* $\mathscr{F}_T$ *to* Y*, denoted* $X \underset{\mathscr{F}_T}{\precsim} Y$*, if* $\sigma(X)$ *is prior or simultaneous in* $\mathscr{F}_T$ *to* $\sigma(Y)$.

Remark 3.38 [Prior-or-Simultaneous-to Relation of Measurables Sets and of Events]
Let $(\Omega, \mathscr{A})$ be a measurable space and $\{C\}$, $\{D\}$ set systems containing the sets $C, D \in \mathscr{A}$ as their only element. Then we say that *C and D are prior or simultaneous in* $\mathscr{F}_T$, denoted $C \underset{\mathscr{F}_T}{\precsim} D$, if $\{C\}$ and $\{D\}$ are prior or simultaneous in $\mathscr{F}_T$. Remember again that in the context of a probability space $(\Omega, \mathscr{A}, P)$, the sets C and D represent events. Again, for the sake brevity, we will not treat any details of the prior-or-simultaneous-to relation of measurable sets. However, using this definition, all propositions about the prior-or-simultaneous-to relation of measurable set systems treated in this section are easily translated to the prior-or-simultaneous-to relation of measurable sets. ◁

3.4.1 Properties of the Prior-or-Simultaneous-to Relation of Measurable Set Systems

Now we show that the prior-or-simultaneous-to relation is reflexive, pseudo-antisymmetric, linear, and transitive.

Theorem 3.39 [Reflexivity, Pseudo-Antisymmetry, and Transitivity]
Let $(\Omega, \mathscr{A})$ *be a measurable space,* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$*, and* $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$*. Then:*

(i) $\exists\, t \in T\colon \mathscr{C} \subset \mathscr{F}_t \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}$ *(reflexivity)*

(ii) $(\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \,\wedge\, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}) \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ *(pseudo-antisymmetry)*

(iii) $(\exists\, s, t \in T\colon \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \subset \mathscr{F}_t) \quad \Rightarrow \quad (\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \,\vee\, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}).$ *(linearity)*

If, additionally, $\mathscr{E} \subset \mathscr{A}$*, then*

(iv) $(\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \,\wedge\, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}) \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}.$ *(transitivity)*

(Proof p. 76)

Hence, according to Proposition (i) of Theorem 3.39, if the set system $\mathscr{C}$ is in the filtration, then it is prior or simultaneous in $\mathscr{F}_T$ to itself. Furthermore, if $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to the set system $\mathscr{D}$ that itself is prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{C}$, then we can conclude that $\mathscr{C}$ and $\mathscr{D}$ are simultaneous in $\mathscr{F}_T$ to each other [see Prop. (ii)]. Furthermore, if $\mathscr{C}$ and $\mathscr{D}$ are in the filtration $\mathscr{F}_T$, then $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$ or $\mathscr{D}$ is prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{C}$ [see Prop. (iii)]. Finally, according to Proposition (iv), if $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$ that itself is prior or simultaneous in $\mathscr{F}_T$ to the set system $\mathscr{E}$, then $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{E}$.

Now we treat an implication of a set system $\mathscr{C}$ being prior or simultaneous in a filtration $\mathscr{F}_T$ to a set system $\mathscr{D}$ for a subset $\mathscr{C}_0$ of $\sigma(\mathscr{C} \cup \mathscr{D})$, the σ-algebra generated by the union of the sets $\mathscr{C}$ and $\mathscr{D}$.

Theorem 3.40 [An Implication of $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$ for a Subset of $\sigma(\mathscr{C} \cup \mathscr{D})$]
Let $(\Omega, \mathscr{A})$ *be a measurable space, let* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$*, and* $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$*. Then:*

$$\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \quad \Rightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\precsim} \mathscr{D}, \qquad \textit{if } \mathscr{C}_0 \subset \sigma(\mathscr{C} \cup \mathscr{D}). \tag{3.40}$$

(Proof p. 78)

Hence, if the measurable set system $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to a measurable set system $\mathscr{D}$, then each subset of the σ-algebra generated by the union $\mathscr{C} \cup \mathscr{D}$ is also prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$.

The following corollary is a generalization of relexivity of the prior-or-simultaneous-to relation of measurable set systems to subsets of a measurable set system.

Corollary 3.41 [A Generalization of Reflexivity]
Let $(\Omega, \mathscr{A})$ *be a measurable space,* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ *a filtration in* $\mathscr{A}$ *and* $\mathscr{C} \subset \mathscr{A}$*. Then:*

$$\exists\, t \in T\colon \mathscr{C} \subset \mathscr{F}_t \quad \Rightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\precsim} \mathscr{C}, \qquad \textit{if } \mathscr{C}_0 \subset \mathscr{C}. \tag{3.41}$$

(Proof p. 78)

According to Proposition (3.41), if $\mathscr{C}$ is in the filtration $\mathscr{F}_T$, then each subset of $\mathscr{C}$ is prior or simultaneous to $\mathscr{C}$.

In the following theorems we treat some properties of the prior-or-simultaneous-to relation of measurable set systems involving σ-algebra generated by the union of two set systems (for the motivation see again Rem. 3.10).

Theorem 3.42 [An Implication of $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$ for a Subset of $\sigma(\mathscr{C} \cup \mathscr{D})$]
Let $(\Omega, \mathscr{A})$ be a measurable space, let $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$. Then:

$$\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \quad \Rightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{C} \cup \mathscr{D}), \qquad \textit{if } \mathscr{C}_0 \subset \sigma(\mathscr{C} \cup \mathscr{D}). \tag{3.42}$$

(Proof p. 78)

According to this theorem, if the set system $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to the set system $\mathscr{D}$, then each subset $\mathscr{C}_0$ of the σ-algebra $\sigma(\mathscr{C} \cup \mathscr{D})$ generated by the union of $\mathscr{C}$ and $\mathscr{D}$ is prior or simultaneous in $\mathscr{F}_T$ to $\sigma(\mathscr{C} \cup \mathscr{D})$.

Remark 3.43 [A Special Case] In the special case $\mathscr{C}_0 = \mathscr{C}$, Proposition (3.42) yields

$$\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \quad \Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{C} \cup \mathscr{D}). \tag{3.43}$$

Hence, if the measurable set system $\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to a measurable set system $\mathscr{D}$, then $\mathscr{C}$ is also prior or simultaneous in $\mathscr{F}_T$ to $\sigma(\mathscr{C} \cup \mathscr{D})$. ◁

In the next theorem we use the notation

$$\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \quad :\Leftrightarrow \quad (\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \,\wedge\, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}). \tag{3.44}$$

Theorem 3.44 [Propositions Involving the Union of Two Set Systems]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D}, \mathscr{E} \subset \mathscr{A}$. Then:

$$\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \quad \Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\precsim} \mathscr{E}. \tag{3.45}$$

(Proof p. 78)

Hence, if the measurable set systems $\mathscr{C}$ and $\mathscr{D}$ are prior or simultaneous in $\mathscr{F}_T$ to a measurable set system $\mathscr{E}$, then the σ-algebra generated by their union $\mathscr{C} \cup \mathscr{D}$ is also prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{E}$.

In the following theorem we show that the set system $\mathscr{C}$ being prior or simultaneous in $\mathscr{F}_T$ to the set system $\mathscr{D}$ is equivalent to $\mathscr{C}$ being prior or simultaneous in $\mathscr{F}_T$ to the σ-algebra generated by $\mathscr{D}$. Furthermore, it is also equivalent to the σ-algebra generated by $\mathscr{C}$ being prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$.

Theorem 3.45 [More Properties Involving σ-Algebras]
Let $(\Omega, \mathscr{A})$ be a measurable space, $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$, and $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$. Then:

(i) $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \quad \Leftrightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{D})$

(ii) $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \quad \Leftrightarrow \quad \sigma(\mathscr{C}) \underset{\mathscr{F}_T}{\precsim} \mathscr{D}.$

(Proof p. 79)

Box 3.3 Properties of the prior-or-simultaneous-to relation of measurable maps

Let X and Y be measurable maps on a measurable space $(\Omega, \mathscr{A})$, let $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ be a filtration in $\mathscr{A}$, and let $\sigma(X)$ and $\sigma(Y)$ denote the σ-algebras generated by X and Y, respectively. Then we say that *X is prior or simultaneous to Y* in $\mathscr{F}_T$, denoted $X \underset{\mathscr{F}_T}{\precsim} Y$, if $X \underset{\mathscr{F}_T}{\prec} Y$ or $X \underset{\mathscr{F}_T}{\approx} Y$.

Two first properties are:

$$(X \underset{\mathscr{F}_T}{\precsim} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\precsim} X) \ \Rightarrow\ X \underset{\mathscr{F}_T}{\approx} Y \qquad \text{(pseudo-antisymmetry)} \qquad \text{(i)}$$

$$\big(\exists\, s, t \in T\colon \sigma(X) \subset \mathscr{F}_s \wedge \sigma(Y) \subset \mathscr{F}_t\big) \ \Rightarrow\ (X \underset{\mathscr{F}_T}{\precsim} Y \ \vee\ Y \underset{\mathscr{F}_T}{\precsim} X). \qquad \text{(linearity)} \qquad \text{(ii)}$$

Additionally, let also W be a measurable map on $(\Omega, \mathscr{A})$. Then:

$$X \underset{\mathscr{F}_T}{\approx} Y \ \Rightarrow\ W \underset{\mathscr{F}_T}{\precsim} Y, \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(iii)}$$

$$X \underset{\mathscr{F}_T}{\precsim} Y \ \Rightarrow\ W \underset{\mathscr{F}_T}{\precsim} Y, \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(iv)}$$

$$X \underset{\mathscr{F}_T}{\precsim} Y \ \Rightarrow\ W \underset{\mathscr{F}_T}{\precsim} (X, Y), \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(v)}$$

$$\exists\, t \in T\colon \sigma(X) \subset \mathscr{F}_t \ \Rightarrow\ W \underset{\mathscr{F}_T}{\precsim} X, \quad \text{if } \sigma(W) \subset \sigma(X). \qquad \text{(vi)}$$

Additionally, let also Z be a measurable map on $(\Omega, \mathscr{A})$. Then:

$$(X \underset{\mathscr{F}_T}{\approx} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\approx} Z) \ \Rightarrow\ W \underset{\mathscr{F}_T}{\precsim} Z, \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(vii)}$$

$$(X \underset{\mathscr{F}_T}{\precsim} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\precsim} Z) \ \Rightarrow\ W \underset{\mathscr{F}_T}{\precsim} Z, \quad \text{if } \sigma(W) \subset \sigma(X, Y) \qquad \text{(viii)}$$

$$X, Y \underset{\mathscr{F}_T}{\precsim} Z \ \Rightarrow\ W \underset{\mathscr{F}_T}{\precsim} Z, \quad \text{if } \sigma(W) \subset \sigma(X, Y). \qquad \text{(ix)}$$

3.4.2 Properties of the Prior-or-Simultaneous-to Relation of Measurable Maps

Now we turn to the prior-or-simultaneous-to relation of measurable maps, the most important properties of which are gathered in Box 3.3. The first two, *pseudo-antisymmetry* and *linearity*, are immediate implications of Definition 3.37 (ii) and Theorem 3.39 (ii) and (iii), respectively. According to Proposition (i) of Box 3.3, if X is prior or simultaneous in $\mathscr{F}_T$ to Y and Y is prior or simultaneous in $\mathscr{F}_T$ to X, then we can conclude that X and Y are simultaneous in $\mathscr{F}_T$ to each other. And, according to Proposition (ii) of Box 3.3, if X and Y are in the filtration $\mathscr{F}_T$, then X is prior or simultaneous in $\mathscr{F}_T$ to Y, or Y is prior or simultaneous in $\mathscr{F}_T$ to X.

Proposition (iii) of Box 3.3 has already been proved in Theorem 3.35 (ii). According to this proposition, X and Y being simultaneous in $\mathscr{F}_T$ implies that each (X, Y)-measurable map W is prior or simultaneous in $\mathscr{F}_T$ to Y.

Proposition (iv) of Box 3.3 is an immediate implication of Theorem 3.35 (i) and (ii). According to this proposition, if X is prior or simultaneous in $\mathscr{F}_T$ to Y, then W is prior or simultaneous in $\mathscr{F}_T$ to Y as well, provided that W is measurable with respect to (X, Y), that is, provided that $\sigma(W) \subset \sigma(X, Y)$.

According to Proposition (v) of Box 3.3, if X is prior or simultaneous in $\mathscr{F}_T$ to Y, then W is prior or simultaneous in $\mathscr{F}_T$ to the bivariate measurable map (X, Y), provided that W is measurable with respect to (X, Y). This is the case, for example, if $W = X$ or $W = Y$. Proposition (v) is an immediate implication of Theorem 3.40 and Definition 3.37 (ii).

According to Proposition (vi) of Box 3.3, if X is in the filtration $\mathscr{F}_T$ and W is X-measurable, then W is prior or simultaneous in $\mathscr{F}_T$ to X. In the special case $W = X$, this implies

$$\exists\, t \in T : \sigma(X) \subset \mathscr{F}_t \quad \Rightarrow \quad X \underset{\mathscr{F}_T}{\precsim} X. \qquad \text{(reflexivity)} \tag{3.46}$$

Note again that $\sigma(W) \subset \sigma(X)$ implies $\sigma(W) \subset \sigma(X, Y)$, but not vice versa. The analog of Proposition (vi) for measurable set systems has been treated in Theorem 3.39 (i). Together with Definition 3.37 (ii), this proves Proposition (vi) of Box 3.3, and with it, Proposition (3.46).

According to Proposition (vii), if X and Y as well as Y and Z are simultaneous in $\mathscr{F}_T$ to each other, then each (X, Y)-measurable map W is prior or simultaneous in $\mathscr{F}_T$ to Z. For example, if $1_{X=x}$ is the indicator of the event that X takes on the value x, X is simultaneous to Y, and Y simultaneous to Z in $\mathscr{F}_T$, then $1_{X=x}$ is prior or simultaneous to Z.

Furthermore, according to Proposition (viii) of Box 3.3, if X is prior or simultaneous in $\mathscr{F}_T$ to Y and Y is prior or simultaneous in $\mathscr{F}_T$ to Z in $\mathscr{F}_T$, then W is prior or simultaneous in $\mathscr{F}_T$ to Z as well, provided that W is (X, Y)-measurable.

Also note, for $W = X$, Proposition (viii) of Box 3.3 yields

$$(X \underset{\mathscr{F}_T}{\precsim} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\precsim} Z) \quad \Rightarrow \quad X \underset{\mathscr{F}_T}{\precsim} Z, \qquad \text{(transitivity)} \tag{3.47}$$

The proofs of Propositions (vii) and (viii) of Box 3.3 are found in the following theorem.

Theorem 3.46 [Further Implications]
Let W, X, Y, Z be measurable maps on a measurable space $(\Omega, \mathscr{A})$, let $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ be a filtration in $\mathscr{A}$, and assume $\sigma(W) \subset \sigma(X, Y)$. Then:

(i) $(X \underset{\mathscr{F}_T}{\approxeq} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\approxeq} Z) \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\precsim} Z$

(ii) $(X \underset{\mathscr{F}_T}{\precsim} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\precsim} Z) \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\precsim} Z.$

(Proof p. 79)

Finally, we consider Proposition (ix) of Box 3.3. According to this proposition, if X and Y are prior or simultaneous to Z in $\mathscr{F}_T$ and W is (X, Y)-measurable, then W is prior or simultaneous in $\mathscr{F}_T$ to Z as well. The proof of this proposition is found in the following theorem, in which we use the notation

$$X, Y \underset{\mathscr{F}_T}{\precsim} Z \quad :\Leftrightarrow \quad (X \underset{\mathscr{F}_T}{\precsim} Z \ \wedge\ Y \underset{\mathscr{F}_T}{\precsim} Z). \tag{3.48}$$

Theorem 3.47 [An Implication Involving an (X, Y)-Measurable Map]
Let W, X, Y, Z be measurable maps on a measurable space $(\Omega, \mathscr{A})$ and $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ a filtration in $\mathscr{A}$. Then

$$X, Y \underset{\mathscr{F}_T}{\precsim} Z \quad \Rightarrow \quad W \underset{\mathscr{F}_T}{\precsim} Z, \qquad \textit{if } \sigma(W) \subset \sigma(X, Y). \tag{3.49}$$

(Proof p. 80)

Box 3.4 Glossary of new concepts

Let $(\Omega, \mathscr{A})$ be a measurable space, that is, let Ω be a set and $\mathscr{A}$ a σ-algebra on Ω.

$\mathscr{F}_T$ — *Filtration in* $\mathscr{A}$. A family $\mathscr{F}_T := (\mathscr{F}_t)_{t \in T}$ of σ-algebras $\mathscr{F}_t \subset \mathscr{A}$ satisfying $\mathscr{F}_s \subset \mathscr{F}_t$, for all $s, t \in T$ with $s \le t$, $T \subset \mathbb{R}$, and $T \ne \emptyset$.

Additionally, let $\mathscr{C}, \mathscr{D} \subset \mathscr{A}$.

$\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D}$ — *$\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$*. This means:

(a) there is an $s \in T$ such that $\mathscr{C} \subset \mathscr{F}_s$ and $\mathscr{D} \not\subset \mathscr{F}_s$, and

(b) there is a $t \in T$ such that $\mathscr{D} \subset \mathscr{F}_t$.

$\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ — *$\mathscr{C}$ is simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$*. This means:

(c) there is a $t \in T$ such that $\mathscr{C} \subset \mathscr{F}_t$, and

(d) for all $t \in T$: $\mathscr{C} \subset \mathscr{F}_t$ if and only if $\mathscr{D} \subset \mathscr{F}_t$.

$\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$ — *$\mathscr{C}$ is prior or simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$*. This means that $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$ or $\mathscr{C}$ is simultaneous in $\mathscr{F}_T$ to $\mathscr{D}$.

Let X, Y be measurable maps on $(\Omega, \mathscr{A})$ and let $\sigma(X)$, $\sigma(Y)$ denote their generated σ-algebras.

$X \underset{\mathscr{F}_T}{\prec} Y$ — *X is prior in $\mathscr{F}_T$ to Y*. This means that $\sigma(X)$ is prior in $\mathscr{F}_T$ to $\sigma(Y)$.

$X \underset{\mathscr{F}_T}{\approx} Y$ — *X is simultaneous in $\mathscr{F}_T$ to Y*. This means that $\sigma(X)$ is simultaneous in $\mathscr{F}_T$ to $\sigma(Y)$.

$X \underset{\mathscr{F}_T}{\precsim} Y$ — *X is prior or simultaneous in $\mathscr{F}_T$ to Y*. This means that $\sigma(X)$ is prior or simultaneous in $\mathscr{F}_T$ to $\sigma(Y)$.

3.5 Summary and Conclusions

In this chapter we introduced the fundamental concept of a *filtration* $\mathscr{F}_T = (\mathscr{F}_t)_{t \in T}$ in the σ-algebra $\mathscr{A}$ of a measurable space $(\Omega, \mathscr{A})$. Such a filtration induces *time order* among measurable set systems (i.e., subsets of $\mathscr{A}$) and among measurable maps on a measurable space $(\Omega, \mathscr{A})$, provided that we interpret T as a time set. In the context of a probability space $(\Omega, \mathscr{A}, P)$, this means that a filtration induces time order among sets of events, and random variables. Referring to a filtration, we defined the prior-to, simultaneous-to, and prior-or-simultaneous-to relations of measurable set systems and of measurable maps. In the framework of a probability space this yields the prior-to, simultaneous-to, and prior-or-simultaneous-to relations of sets of events and of random variables. Box 3.4 summarizes these concepts. The properties of the prior-to relation, the simultaneous-to relation, and the prior-or-simultaneous-to relation of measurable maps have been gathered in Boxes 3.1, 3.2, and 3.3, respectively.

It should be noted that these concepts are defined in the framework of a measurable space $(\Omega, \mathscr{A})$. No probability measure on $\mathscr{A}$ is involved. Hence, whether or not a random variable X is prior to a random variable Y does not depend on the stochastic dependencies between these random variables, let alone on data that result from conducting a random experiment.

3.6 Proofs

Proof of Theorem 3.6

(i). If $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$, then

(a) $\exists\, s \in T\colon \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$
(b) $\exists\, t \in T\colon \mathscr{D} \subset \mathscr{F}_t$.

Because $\mathscr{F}_T$ is a filtration, $\neg(\exists\, r \in T\colon \mathscr{D} \subset \mathscr{F}_r \wedge \mathscr{C} \not\subset \mathscr{F}_r)$. This implies that $\mathscr{D}$ is not prior to $\mathscr{C}$, which proves asymmetry.

(ii). If $\mathscr{C}$ is prior to $\mathscr{D}$ in $\mathscr{F}_T$, then (a) holds, and if $\mathscr{D}$ is prior to $\mathscr{E}$, then

(c) $\exists\, t \in T\colon \mathscr{D} \subset \mathscr{F}_t \wedge \mathscr{E} \not\subset \mathscr{F}_t$,
(d) $\exists\, r \in T\colon \mathscr{E} \subset \mathscr{F}_r$.

Because $\mathscr{F}_T$ is a filtration, the conjunction of (a) and (c) implies

(e) $\exists\, s \in T\colon \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{E} \not\subset \mathscr{F}_s$.

In conjunction with (d) this is equivalent to $\mathscr{C} \underset{\mathscr{F}_T}{\lessapprox} \mathscr{E}$ [see Def. 3.3 (i)].

Proof of Theorem 3.8

(i). If $\mathscr{C}$ is prior in $\mathscr{F}_T$ to $\mathscr{D}$, then

(a) $\exists\, s \in T\colon \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$
(b) $\exists\, t \in T\colon \mathscr{D} \subset \mathscr{F}_t$.

According to Proposition (3.17), Proposition (a) and $\mathscr{C}_0 \subset \mathscr{C}$ imply

(c) $\exists\, s \in T\colon \mathscr{C}_0 \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$.

However, the conjunction of (c) and (b) is equivalent to $\mathscr{C}_0 \underset{\mathscr{F}_T}{\lessapprox} \mathscr{D}$ [see Def. 3.3 (i)].

Proof of Theorem 3.9

(i). $\mathscr{C} \underset{\mathscr{F}_T}{\lessapprox} \mathscr{D}$ is equivalent to the conjunction of

(a) $\exists s \in T: \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$
(b) $\forall\, t \in T: \mathscr{C} \subset \mathscr{F}_t \Leftrightarrow \mathscr{D} \subset \mathscr{F}_t$.

According to Proposition (3.19), (a) is equivalent to

(c) $\exists s \in T: \mathscr{C} \subset \mathscr{F}_s \wedge \sigma(\mathscr{D}) \not\subset \mathscr{F}_s$,

and for the same reason, (b) is equivalent to

(d) $\forall\, t \in T: \mathscr{C} \subset \mathscr{F}_t \Leftrightarrow \sigma(\mathscr{D}) \subset \mathscr{F}_t$.

However, the conjunction of (c) and (d) is equivalent to $\mathscr{C} \underset{\mathscr{F}_T}{\lessapprox} \sigma(\mathscr{D})$.

(ii). As mentioned above, $\mathscr{C} \underset{\mathscr{F}_T}{\lessapprox} \mathscr{D}$ is equivalent to the conjunction of (a) and (b). According to Proposition (3.19), (a) is equivalent to

(e) $\exists s \in T: \sigma(\mathscr{C}) \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$.

Furthermore, again according to Proposition (3.19), (b) is equivalent to

(f) $\forall\, t \in T\colon \sigma(\mathscr{C}) \subset \mathscr{F}_t \;\Leftrightarrow\; \mathscr{D} \subset \mathscr{F}_t$.

However, the conjunction of (e) and (f) is equivalent to $\sigma(\mathscr{C}) \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$.

Proof of Theorem 3.11

(i).

$$
\begin{array}{lll}
 & (\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \;\wedge\; \exists\, t \in T\colon \mathscr{E} \subset \mathscr{F}_t) & \\
\Leftrightarrow & \big(\exists\, r \in T\colon \mathscr{C} \subset \mathscr{F}_r \wedge \mathscr{D} \not\subset \mathscr{F}_r & [\text{Def. 3.3 (i) (a)}] \\
 & \;\wedge\; \exists\, s \in T\colon \mathscr{D} \subset \mathscr{F}_s & [\text{Def. 3.3 (i) (b)}] \\
 & \;\wedge\; \exists\, t \in T\colon \mathscr{E} \subset \mathscr{F}_t\big) & [\text{part 2 of the premise in (i)}] \\
\Rightarrow & \big(\exists\, r \in T\colon \mathscr{C} \subset \mathscr{F}_r \wedge \sigma(\mathscr{D} \cup \mathscr{E}) \not\subset \mathscr{F}_r & [\mathscr{D} \not\subset \mathscr{F}_r, \mathscr{D} \subset \sigma(\mathscr{D} \cup \mathscr{E})] \\
 & \;\wedge\; \exists\, u \in T\colon \sigma(\mathscr{D} \cup \mathscr{E}) \subset \mathscr{F}_u\big) & [u = \max\{s, t\}] \\
\Leftrightarrow & \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{D} \cup \mathscr{E}). & [\text{Def. 3.3 (i)}]
\end{array}
$$

(ii). First of all, note that $\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$ is equivalent to the conjunction of:

(a) $\exists\, r \in T$: $\mathscr{C} \subset \mathscr{F}_r \wedge \mathscr{E} \not\subset \mathscr{F}_r$
(b) $\exists\, s \in T$: $\mathscr{D} \subset \mathscr{F}_s \wedge \mathscr{E} \not\subset \mathscr{F}_s$
(c) $\exists\, t \in T$: $\mathscr{E} \subset \mathscr{F}_t$

[see Prop. (3.22) and Def. 3.3 (i)]. Without loss of generality, we assume $r \le s$. Because $\mathscr{F}_T$ is a filtration, this implies

(d) $\mathscr{C}, \mathscr{D} \subset \mathscr{F}_s$
(e) $\forall\, u \in T\colon\; \mathscr{C}, \mathscr{D} \subset \mathscr{F}_u \;\Leftrightarrow\; \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_u$

[see RS-Prop. (1.9)].

$\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \;\Rightarrow\; \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$. The conjunction of (b), (d), and (e) implies

(f) $\exists\, s \in T\colon\; \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_s \;\wedge\; \mathscr{E} \not\subset \mathscr{F}_s$.

However, the conjunction of (c) and (f) is equivalent to $\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$ [see Def. 3.3 (i)].

$\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E} \;\Leftarrow\; \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$. As just has been said, the conjunction of (c) and (f) is equivalent to $\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$. Now, the conjunction of (f) and (e) implies (a) and (b). However, the conjunction of (a), (b), and (c) is equivalent to $\mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}$.

Proof of Theorem 3.13

$\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \;\Rightarrow\; \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{C} \cup \mathscr{D})$. This proposition is a special case of Theorem 3.11 (i), in which $\mathscr{C}$ also takes the role of $\mathscr{E}$. The existence of a $t \in T$ with $\mathscr{C} \subset \mathscr{F}_t$ is a part of the premise $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$ [see Def. 3.3 (i)].

$\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \;\Leftarrow\; \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{C} \cup \mathscr{D})$. According to Definition 3.3 (i), $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \sigma(\mathscr{C} \cup \mathscr{D})$ is equivalent to the conjunction of

(a) $\exists\, s \in T: \mathscr{C} \subset \mathscr{F}_s \wedge \sigma(\mathscr{C} \cup \mathscr{D}) \not\subset \mathscr{F}_s$
(b) $\exists\, t \in T: \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t$.

According to RS-Proposition (1.9), Proposition (a) implies

(c) $\exists\, s \in T: \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$.

Furthermore, because $\mathscr{D} \subset \sigma(\mathscr{C} \cup \mathscr{D})$, Proposition (b) implies

(d) $\exists\, t \in T: \mathscr{D} \subset \mathscr{F}_t$.

However, according to Definition 3.3 (i), the conjunction of (c) and (d) is equivalent to $\mathscr{C} \underset{\mathscr{F}_T}{\precneqq} \mathscr{D}$.

Proof of Theorem 3.19

(i).

$$
\begin{aligned}
X, Y \underset{\mathscr{F}_T}{\precneqq} Z \quad &\Leftrightarrow \quad \sigma(X), \sigma(Y) \underset{\mathscr{F}_T}{\precneqq} \sigma(Z) && \text{[Def. 3.3 (ii)]}\\
&\Leftrightarrow \quad \sigma\big(\sigma(X) \cup \sigma(Y)\big) \underset{\mathscr{F}_T}{\precneqq} \sigma(Z) && \text{[Th. 3.11 (ii)]}\\
&\Leftrightarrow \quad \sigma(X, Y) \underset{\mathscr{F}_T}{\precneqq} \sigma(Z) && \text{[(3.20)]}\\
&\Rightarrow \quad \sigma(W) \underset{\mathscr{F}_T}{\precneqq} \sigma(Z) && [\sigma(W) \subset \sigma(X, Y),\ \text{Th. 3.8}]\\
&\Leftrightarrow \quad W \underset{\mathscr{F}_T}{\precneqq} Z. && \text{[Def. 3.3 (ii)]}
\end{aligned}
$$

(ii).

$$
\begin{aligned}
(X \underset{\mathscr{F}_T}{\precneqq} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\precneqq} Z) \quad &\Leftrightarrow \quad \big(\sigma(X) \underset{\mathscr{F}_T}{\precneqq} \sigma(Y) \ \wedge\ \sigma(Y) \underset{\mathscr{F}_T}{\precneqq} \sigma(Z)\big) && \text{[Def. 3.3 (ii)]}\\
&\Rightarrow \quad \big(\sigma(X) \underset{\mathscr{F}_T}{\precneqq} \sigma(Z) \ \wedge\ \sigma(Y) \underset{\mathscr{F}_T}{\precneqq} \sigma(Z)\big) && \text{[Th. 3.6 (ii)]}\\
&\Leftrightarrow \quad \sigma(X), \sigma(Y) \underset{\mathscr{F}_T}{\precneqq} Z && \text{[(3.22)]}\\
&\Leftrightarrow \quad X, Y \underset{\mathscr{F}_T}{\precneqq} Z && \text{[Def. 3.3 (ii)]}\\
&\Rightarrow \quad W \underset{\mathscr{F}_T}{\precneqq} Z. && [\sigma(W) \subset \sigma(X, Y),\ \text{(i)}]
\end{aligned}
$$

Proof of Theorem 3.21

$$
\begin{aligned}
& \big(X \underset{\mathscr{F}_T}{\precneqq} Y \wedge \exists\, t \in T: \sigma(Z) \subset \mathscr{F}_t\big) && \\
\Leftrightarrow\quad & \big(\sigma(X) \underset{\mathscr{F}_T}{\precneqq} \sigma(Y) \wedge \exists\, t \in T: \sigma(Z) \subset \mathscr{F}_t\big) && \text{[Def. 3.3 (ii)]}\\
\Rightarrow\quad & \sigma(X) \underset{\mathscr{F}_T}{\precneqq} \sigma\big(\sigma(Y) \cup \sigma(Z)\big) && \text{[Th. 3.11 (i)]}\\
\Rightarrow\quad & \sigma(W) \underset{\mathscr{F}_T}{\precneqq} \sigma(Y, Z) && \text{[Th. 3.8, (3.20)]}\\
\Leftrightarrow\quad & W \underset{\mathscr{F}_T}{\precneqq} (Y, Z). && \text{[Def. 3.3 (ii)]}
\end{aligned}
$$

Proof of Theorem 3.27

(i). The premise

(a) $\exists\, t \in T : \mathscr{C} \subset \mathscr{F}_t$

implies

(b) $\forall\, t \in T : \mathscr{C} \subset \mathscr{F}_t \Leftrightarrow \mathscr{C} \subset \mathscr{F}_t$.

According to Definition 3.24 (i), the conjunction of (a) and (b) is equivalent to $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{C}$.

(ii). The premise $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ is equivalent to the conjunction of (a) and

(c) $\forall\, t \in T:\;\; \mathscr{C} \subset \mathscr{F}_t \;\;\Leftrightarrow\;\; \mathscr{D} \subset \mathscr{F}_t$.

Now, the conjunction of (a) and (c) implies

(d) $\exists\, t \in T:\;\; \mathscr{D} \subset \mathscr{F}_t$

and

(e) $\forall\, t \in T:\;\; \mathscr{D} \subset \mathscr{F}_t \;\;\Leftrightarrow\;\; \mathscr{C} \subset \mathscr{F}_t$.

However, the conjunction of (d) and (e) is equivalent to $\mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{C}$.

(iii). $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ is equivalent to (a) and (c), and $\mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}$ implies (d) and

(f) $\forall\, t \in T:\;\; \mathscr{D} \subset \mathscr{F}_t \;\;\Leftrightarrow\;\; \mathscr{E} \subset \mathscr{F}_t$.

Now, the conjunction of (c) and (f) implies

(g) $\forall\, t \in T:\;\; \mathscr{C} \subset \mathscr{F}_t \;\;\Leftrightarrow\;\; \mathscr{E} \subset \mathscr{F}_t$.

However, the conjunction of (a) and (g) is equivalent to $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{E}$.

Proof of Theorem 3.28

(i). $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ is equivalent to the conjunction of the following two propositions:

(a) $\exists\, t \in T:\;\; \mathscr{C} \subset \mathscr{F}_t$
(b) $\forall\, t \in T:\;\; \mathscr{C} \subset \mathscr{F}_t \;\;\Leftrightarrow\;\; \mathscr{D} \subset \mathscr{F}_t$.

Furthermore, according to Proposition (3.19), (a) is equivalent to

(c) $\exists\, t \in T : \sigma(\mathscr{C}) \subset \mathscr{F}_t$,

and (b) is equivalent to

(d) $\forall\, t \in T:\;\; \sigma(\mathscr{C}) \subset \mathscr{F}_t \;\;\Leftrightarrow\;\; \mathscr{D} \subset \mathscr{F}_t$.

However, the conjunction of (c) and (d) is equivalent to $\sigma(\mathscr{C}) \underset{\mathscr{F}_T}{\approx} \mathscr{D}$.

(ii). As mentioned above, $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ is equivalent to the conjunction of (a) and (b). $\mathscr{F}_T$ being a filtration implies that each $\mathscr{F}_t$, $t \in T$, is a σ-algebra. Therefore, and because of RS-Proposition (1.9), (b) is equivalent to:

(e) $\forall\, t \in T:\;\; \mathscr{C} \subset \mathscr{F}_t \;\;\Leftrightarrow\;\; \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t$.

However, the conjunction of (a) and (e) is equivalent to $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \sigma(\mathscr{C} \cup \mathscr{D})$ [see Def. 3.24 (i)], to $\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{C}$ [see Th. 3.27 (ii)], and to $\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ [see Th. 3.27 (iii)].

Proof of Theorem 3.29

$$
\begin{aligned}
(\mathscr{C} \underset{\mathscr{F}_T}{\approxeq} \mathscr{D} \ \wedge\ \mathscr{D} \underset{\mathscr{F}_T}{\approxeq} \mathscr{E}) \quad &\Rightarrow \quad \big(\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approxeq} \mathscr{D} \ \wedge\ \mathscr{D} \underset{\mathscr{F}_T}{\approxeq} \mathscr{E}\big) \qquad && [\text{Th. } 3.28 \text{ (ii)}] \\
&\Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approxeq} \mathscr{E}. && [\text{Th. } 3.27 \text{ (iii)}]
\end{aligned}
$$

Proof of Theorem 3.31

(i). $\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D}$ is equivalent to the conjunction of

(a) $\exists\, s \in T: \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$
(b) $\exists\, t \in T: \mathscr{D} \subset \mathscr{F}_t$.

Because $\mathscr{F}_T$ is a filtration, the conjunction of (a) and (b) implies

(c) $\forall\, t \in T: \mathscr{D} \subset \mathscr{F}_t \Rightarrow \mathscr{C} \subset \mathscr{F}_t$.

Because of RS-Proposition (1.9), (c) is equivalent to

(d) $\forall\, t \in T: \mathscr{D} \subset \mathscr{F}_t \Rightarrow \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t$.

Furthermore, because $\mathscr{D} \subset \sigma(\mathscr{C} \cup \mathscr{D})$,

(e) $\forall\, t \in T: \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t \Rightarrow \mathscr{D} \subset \mathscr{F}_t$.

However, the conjunction of (d) and (e) is equivalent to

(f) $\forall\, t \in T: \mathscr{D} \subset \mathscr{F}_t \Leftrightarrow \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t$.

Finally, according to Definition 3.24 (i) and Theorem 3.27 (ii), the conjunction of (b) and (f) is equivalent to $\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{D}$.

(ii). $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ is equivalent to the conjunction of

(g) $\exists\, t \in T: \mathscr{C} \subset \mathscr{F}_t$
(h) $\forall\, t \in T: \mathscr{C} \subset \mathscr{F}_t \Leftrightarrow \mathscr{D} \subset \mathscr{F}_t$.

Now, (h) implies

(i) $\neg \exists\, t \in T: \mathscr{C} \subset \mathscr{F}_t \wedge \mathscr{D} \not\subset \mathscr{F}_t$,

which in turn implies $\neg(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D})$ [see again Def. 3.24 (i)].

Proof of Theorem 3.32

(i). According to Theorem 3.28 (ii), $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ is implies $\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{D}$, which in turn is equivalent to the conjunction of

(a) $\exists\, t \in T: \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t$
(b) $\forall\, t \in T: \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t \Leftrightarrow \mathscr{D} \subset \mathscr{F}_t$

[see Def. 3.24 (i)]. Because of $\mathscr{C}_0 \subset \sigma(\mathscr{C} \cup \mathscr{D})$ and Proposition (3.17), Proposition (a) implies

(c) $\exists\, t \in T\colon \mathscr{C}_0 \subset \mathscr{F}_t$.

Now, we distinguish two cases.

Case 1: $\forall\, t \in T\colon \mathscr{C}_0 \subset \mathscr{F}_t \Leftrightarrow \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t$. Together with (b), this proposition implies

(d) $\forall\, t \in T\colon \mathscr{C}_0 \subset \mathscr{F}_t \Leftrightarrow \mathscr{D} \subset \mathscr{F}_t$.

Now, the conjunction of (c) and (d) is equivalent to $\mathscr{C}_0 \underset{\mathscr{F}_T}{\approx} \mathscr{D}$ [see Def. 3.24 (i)].

Case 2: $\neg\big(\forall\, t \in T\colon \mathscr{C}_0 \subset \mathscr{F}_t \Leftrightarrow \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t\big)$. In conjunction with (c), this proposition implies

(e) $\exists\, s \in T\colon \mathscr{C}_0 \subset \mathscr{F}_s \wedge \sigma(\mathscr{C} \cup \mathscr{D}) \not\subset \mathscr{F}_s$.

Together with (b), this proposition implies

(f) $\exists\, s \in T\colon \mathscr{C}_0 \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s$.

Furthermore, the conjunction of (a) and (b) implies

(g) $\exists\, t \in T\colon \mathscr{D} \subset \mathscr{F}_t$.

However, the conjunction of (f) and (g) is equivalent to $\mathscr{C}_0 \underset{\mathscr{F}_T}{\lessapprox} \mathscr{D}$ [see Def. 3.3 (i)].

(ii).

$$
\begin{aligned}
\mathscr{C} \underset{\mathscr{F}_T}{\lessapprox} \mathscr{D} \quad &\Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{D} && \text{[Th. 3.31 (i)]} \\
&\Rightarrow \quad (\mathscr{C}_0 \underset{\mathscr{F}_T}{\lessapprox} \mathscr{D} \vee \mathscr{C}_0 \underset{\mathscr{F}_T}{\approx} \mathscr{D}). && \big[\sigma\big(\sigma(\mathscr{C} \cup \mathscr{D}) \cup \mathscr{D}\big),\ \text{(i)}\big]
\end{aligned}
$$

Proof of Theorem 3.33

(i). According to Theorem 3.28 (ii),

$$
(\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\lessapprox} \mathscr{E}) \quad \Rightarrow \quad \big(\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\approx} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\lessapprox} \mathscr{E}\big).
$$

The right-hand side of this proposition is equivalent to the conjunction of

(a) $\exists\, t \in T\colon \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t$
(b) $\forall\, t \in T\colon \sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_t \Leftrightarrow \mathscr{D} \subset \mathscr{F}_t$
(c) $\exists\, s \in T\colon (\mathscr{D} \subset \mathscr{F}_s \wedge \mathscr{E} \not\subset \mathscr{F}_s)$
(d) $\exists\, r \in T\colon \mathscr{E} \subset \mathscr{F}_r$.

Now the conjunction of (b) and (c) implies

(e) $\exists\, s \in T\colon \big(\sigma(\mathscr{C} \cup \mathscr{D}) \subset \mathscr{F}_s \wedge \mathscr{E} \not\subset \mathscr{F}_s\big)$.

However, according to Definition 3.3 (i), the conjunction of (e) and (d) is equivalent to $\sigma(\mathscr{C} \cup \mathscr{D}) \underset{\mathscr{F}_T}{\lessapprox} \mathscr{E}$. Now, Proposition (i) follows from Proposition (3.18).

(ii). $\mathscr{C} \underset{\mathscr{F}_T}{\lessapprox} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}$ is equivalent to the conjunction of

(f) $\exists\, s \in T\colon (\mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \not\subset \mathscr{F}_s)$
(g) $\exists\, t \in T\colon \mathscr{D} \subset \mathscr{F}_t$
(h) $\forall\, t \in T\colon \mathscr{D} \subset \mathscr{F}_t \Leftrightarrow \mathscr{E} \subset \mathscr{F}_t$.

Now, the conjunction of (f) and (h) implies

(i) $\exists\, s \in T : (\mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{E} \not\subset \mathscr{F}_s)$,

and the conjunction of (g) and (h) implies

(j) $\exists\, t \in T : \mathscr{E} \subset \mathscr{F}_t$.

However, according to Definition 3.3 (i), the conjunction of (i) and (j) is equivalent to $\mathscr{C} \underset{\mathscr{F}_T}{\lesssim} \mathscr{E}$. Now, Proposition (ii) follows from Proposition (3.18).

Proof of Theorem 3.35

(i).

$$
\begin{array}{lll}
 & X \underset{\mathscr{F}_T}{\lesssim} Y & \\
\Leftrightarrow & \sigma(X) \underset{\mathscr{F}_T}{\lesssim} \sigma(Y) & \text{[Def. 3.3 (ii)]} \\
\Rightarrow & \sigma\big(\sigma(X) \cup \sigma(Y)\big) \underset{\mathscr{F}_T}{\approx} \sigma(Y) & \text{[Th. 3.31 (i)]} \\
\Leftrightarrow & \sigma(X,Y) \underset{\mathscr{F}_T}{\approx} \sigma(Y) & \text{[(3.20)]} \\
\Rightarrow & \big(\sigma(W) \underset{\mathscr{F}_T}{\lesssim} \sigma(Y) \ \vee\ \sigma(W) \underset{\mathscr{F}_T}{\approx} \sigma(Y)\big) & [\sigma(W) \subset \sigma(X,Y),\ \text{Th. 3.32 (i)}] \\
\Leftrightarrow & (W \underset{\mathscr{F}_T}{\lesssim} Y \ \vee\ W \underset{\mathscr{F}_T}{\approx} Y). & \text{[Defs. 3.3 (ii), 3.24 (ii)]}
\end{array}
$$

(ii). Similarly,

$$
\begin{array}{lll}
 & X \underset{\mathscr{F}_T}{\approx} Y & \\
\Leftrightarrow & \sigma(X) \underset{\mathscr{F}_T}{\approx} \sigma(Y) & \text{[Def. 3.24 (ii)]} \\
\Rightarrow & \sigma\big(\sigma(X) \cup \sigma(Y)\big) \underset{\mathscr{F}_T}{\approx} \sigma(Y) & \text{[Th. 3.28 (ii)]} \\
\Leftrightarrow & \sigma(X,Y) \underset{\mathscr{F}_T}{\approx} \sigma(Y) & \text{[(3.20)]} \\
\Rightarrow & \big(\sigma(W) \underset{\mathscr{F}_T}{\lesssim} \sigma(Y) \ \vee\ \sigma(W) \underset{\mathscr{F}_T}{\approx} \sigma(Y)\big) & [\sigma(W) \subset \sigma(X,Y),\ \text{Th. 3.32 (i)}] \\
\Leftrightarrow & (W \underset{\mathscr{F}_T}{\lesssim} Y \ \vee\ W \underset{\mathscr{F}_T}{\approx} Y). & \text{[Defs. 3.3 (ii), 3.24 (ii)]}
\end{array}
$$

Proof of Theorem 3.39

(i).

$$
\begin{array}{llll}
\exists\, t \in T : \mathscr{C} \subset \mathscr{F}_t & \Rightarrow & \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{C} & \text{[Th. 3.27 (i)]} \\
 & \Rightarrow & \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}. & \text{[Def. 3.37 (i)]}
\end{array}
$$

(ii). According to Theorem 3.6 (i),

(a) $\mathscr{C} \underset{\mathscr{F}_T}{\lesssim} \mathscr{D} \quad\Rightarrow\quad \neg(\mathscr{D} \underset{\mathscr{F}_T}{\lesssim} \mathscr{C})$

and, according to Theorems 3.31 (ii) and 3.27 (ii),

(b) $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \quad \Rightarrow \quad \neg(\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C})$.

Hence, in both cases in which $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$ holds, we can conclude $\neg(\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C})$, and in both cases in which $\mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}$ holds, we can conclude $\neg(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D})$. Therefore,

$$
\begin{aligned}
(\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \;\wedge\; \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}) \quad &\Rightarrow \quad \neg(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D}) \;\wedge\; \neg(\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C}) && \text{[(a), (b), Def. 3.37 (i)]}\\
&\Rightarrow \quad \neg(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \;\wedge\; \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C})\\
&\Leftrightarrow \quad \neg\,(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D}) \;\vee\; \neg\,(\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C}). && \text{[de Morgan]}
\end{aligned}
$$

Now $\neg\,(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D})$ implies $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$, because we assume $\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}$ [see Def. 3.37 (i)], and $\neg\,(\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C})$ implies $\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}$, because we assume $\mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}$ [see Def. 3.37 (i) and Th. 3.27 (ii)].

(iii). We presume $\exists\, s, t \in T : \mathscr{C} \subset \mathscr{F}_s \wedge \mathscr{D} \subset \mathscr{F}_t$.

$$
\begin{aligned}
&\text{Case 1:} && (\forall\, t \in T : \mathscr{C} \subset \mathscr{F}_t \;\Leftrightarrow\; \mathscr{D} \subset \mathscr{F}_t)\\
&\Rightarrow && \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} && [\exists\, s \in T : \mathscr{C} \subset \mathscr{F}_s, \text{ Def. 3.24 (i)}]\\
&\Rightarrow && \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D}. && \text{[Def. 3.37 (i)]}\\
&\Rightarrow && (\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \;\vee\; \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}).\\
&\text{Case 2:} && \neg\,(\forall\, t \in T : \mathscr{C} \subset \mathscr{F}_t \;\Leftrightarrow\; \mathscr{D} \subset \mathscr{F}_t)\\
&\Rightarrow && \big(\exists\, t \in T : (\mathscr{C} \subset \mathscr{F}_t \;\wedge \mathscr{D} \not\subset \mathscr{F}_t) \;\vee\; (\mathscr{C} \not\subset \mathscr{F}_t \;\wedge \mathscr{D} \subset \mathscr{F}_t)\big)\\
&\Rightarrow && (\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \;\vee\; \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{C}) && \text{[Def. 3.3 (i)]}\\
&\Rightarrow && (\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \;\vee\; \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{C}). && \text{[Def. 3.37 (i)]}
\end{aligned}
$$

(iv). Note that

$$
\begin{aligned}
&(\mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\precsim} \mathscr{E})\\
\Leftrightarrow\;& \big((\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \vee \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}) \wedge (\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E} \vee \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E})\big). && \text{[Def. 3.37 (i)]}\\
\Leftrightarrow\;& \big((\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E}) \vee (\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) \vee (\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E}) \vee (\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E})\big).
\end{aligned}
$$

The latter proposition follows from repeatedly applying the first distributive law. According to the last proposition, we consider four cases.

$$
\begin{aligned}
&\text{Case 1:} && (\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \;\wedge\; \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E})\\
&\Rightarrow && \mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E} && \text{[Th. 3.6 (ii)]}\\
&\Rightarrow && \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}. && \text{[Def. 3.37 (i)]}\\
&\text{Case 2:} && (\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \;\wedge\; \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E})\\
&\Rightarrow && \mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E} && \text{[Th. 3.33 (ii)]}\\
&\Rightarrow && \mathscr{C} \underset{\mathscr{F}_T}{\precsim} \mathscr{E}. && \text{[Def. 3.37 (i)]}\\
&\text{Case 3:} && (\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \;\wedge\; \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E})
\end{aligned}
$$

$$
\begin{aligned}
&\Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E} && \text{[Th. 3.33 (i)]}\\
&\Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precapprox} \mathscr{E}. && \text{[Def. 3.37 (i)]}\\
\text{Case 4:} \quad & \quad\quad\ (\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D} \ \wedge\ \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) && \\
&\Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{E} && \text{[Th. 3.27 (iii)]}\\
&\Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precapprox} \mathscr{E}. && \text{[Def. 3.37 (i)]}
\end{aligned}
$$

Proof of Theorem 3.40

$$
\begin{aligned}
\mathscr{C} \underset{\mathscr{F}_T}{\precapprox} \mathscr{D} \quad &\Leftrightarrow \quad (\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \ \vee\ \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}) && \text{[Def. 3.37 (i)]}\\
&\Rightarrow \quad (\mathscr{C}_0 \underset{\mathscr{F}_T}{\prec} \mathscr{D} \ \vee\ \mathscr{C}_0 \underset{\mathscr{F}_T}{\approx} \mathscr{D}) && \text{[Th. 3.32]}\\
&\Leftrightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\precapprox} \mathscr{D}. && \text{[Def. 3.37 (i)]}
\end{aligned}
$$

Proof of Corollary 3.41

$$
\begin{aligned}
\exists\, t \in T : \mathscr{C} \subset \mathscr{F}_t \quad &\Rightarrow \quad \mathscr{C} \underset{\mathscr{F}_T}{\precapprox} \mathscr{C} && \text{[Th. 3.39 (i)]}\\
&\Rightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\precapprox} \mathscr{C}. && \text{[Th. 3.40]}
\end{aligned}
$$

Proof of Theorem 3.42

$$
\begin{aligned}
\mathscr{C} \underset{\mathscr{F}_T}{\precapprox} \mathscr{D} \quad &\Leftrightarrow \quad (\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{D} \ \vee\ \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{D}) && \text{[Def. 3.37 (i)]}\\
&\Rightarrow \quad \big(\mathscr{C} \underset{\mathscr{F}_T}{\prec} \sigma(\mathscr{C} \cup \mathscr{D}) \ \vee\ \mathscr{C} \underset{\mathscr{F}_T}{\approx} \sigma(\mathscr{C} \cup \mathscr{D})\big) && \text{[Ths. 3.13, 3.27 (ii), 3.28 (ii)]}\\
&\Rightarrow \quad \Big(\mathscr{C}_0 \underset{\mathscr{F}_T}{\prec} \sigma(\mathscr{C} \cup \mathscr{D}) \ \vee\ \big(\mathscr{C}_0 \underset{\mathscr{F}_T}{\prec} \sigma(\mathscr{C} \cup \mathscr{D}) \ \vee\ \mathscr{C}_0 \underset{\mathscr{F}_T}{\approx} \sigma(\mathscr{C} \cup \mathscr{D})\big)\Big) && \text{[Th. 3.32]}\\
&\Leftrightarrow \quad \big(\mathscr{C}_0 \underset{\mathscr{F}_T}{\prec} \sigma(\mathscr{C} \cup \mathscr{D}) \ \vee\ \mathscr{C}_0 \underset{\mathscr{F}_T}{\approx} \sigma(\mathscr{C} \cup \mathscr{D})\big) && \\
&\Leftrightarrow \quad \mathscr{C}_0 \underset{\mathscr{F}_T}{\precapprox} \sigma(\mathscr{C} \cup \mathscr{D}). && \text{[Def. 3.37 (i)]}
\end{aligned}
$$

Proof of Theorem 3.44

First of all, note that

$$
\begin{aligned}
& \mathscr{C}, \mathscr{D} \underset{\mathscr{F}_T}{\precapprox} \mathscr{E} && \\
\Leftrightarrow\ & (\mathscr{C} \underset{\mathscr{F}_T}{\precapprox} \mathscr{E} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\precapprox} \mathscr{E}) && \text{[(3.44)]}\\
\Leftrightarrow\ & \big((\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E} \vee \mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) \wedge (\mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E} \vee \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E})\big) && \text{[Def. 3.37 (i)]}\\
\Leftrightarrow\ & \big((\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E}) \vee (\mathscr{C} \underset{\mathscr{F}_T}{\prec} \mathscr{E} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E}) \vee (\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{E} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\prec} \mathscr{E}) \vee (\mathscr{C} \underset{\mathscr{F}_T}{\approx} \mathscr{E} \wedge \mathscr{D} \underset{\mathscr{F}_T}{\approx} \mathscr{E})\big). &&
\end{aligned}
$$

The last proposition follows from repeatedly applying the first distributive law. Accordingly, we consider four cases.

$$\begin{aligned} \text{Case 1:}\quad (\mathscr{C} \prec_{\mathscr{F}_T} \mathscr{E} \wedge \mathscr{D} \prec_{\mathscr{F}_T} \mathscr{E}) \quad &\Leftrightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \prec_{\mathscr{F}_T} \mathscr{E} && \text{[Th. 3.11 (ii)]} \\ &\Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \precapprox_{\mathscr{F}_T} \mathscr{E}. && \text{[Def. 3.37 (i)]} \end{aligned}$$

$$\begin{aligned} \text{Case 2:}\quad (\mathscr{C} \prec_{\mathscr{F}_T} \mathscr{E} \wedge \mathscr{D} \approxeq_{\mathscr{F}_T} \mathscr{E}) \quad &\Rightarrow \quad \big(\sigma(\mathscr{C} \cup \mathscr{E}) \approxeq_{\mathscr{F}_T} \mathscr{E} \wedge \mathscr{E} \approxeq_{\mathscr{F}_T} \mathscr{D}\big) && \text{[Ths. 3.31 (i), 3.27 (ii)]} \\ &\Rightarrow \quad \sigma\big(\sigma(\mathscr{C} \cup \mathscr{D}) \cup \mathscr{E}\big) \approxeq_{\mathscr{F}_T} \mathscr{E} && \text{[Ths. 3.29, 3.27 (iii)]} \\ &\Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \precapprox_{\mathscr{F}_T} \mathscr{E}. && \text{[Th. 3.32 (i), Def. 3.37 (i)]} \end{aligned}$$

$$\begin{aligned} \text{Case 3:}\quad (\mathscr{C} \approxeq_{\mathscr{F}_T} \mathscr{E} \wedge \mathscr{D} \prec_{\mathscr{F}_T} \mathscr{E}) \quad &\Leftrightarrow \quad (\mathscr{D} \prec_{\mathscr{F}_T} \mathscr{E} \wedge \mathscr{C} \approxeq_{\mathscr{F}_T} \mathscr{E}) \\ &\Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \precapprox_{\mathscr{F}_T} \mathscr{E}. \qquad \text{[Case 2, } \sigma(\mathscr{D} \cup \mathscr{C}) = \sigma(\mathscr{C} \cup \mathscr{D})\text{]} \end{aligned}$$

$$\begin{aligned} \text{Case 4:}\quad (\mathscr{C} \approxeq_{\mathscr{F}_T} \mathscr{E} \wedge \mathscr{D} \approxeq_{\mathscr{F}_T} \mathscr{E}) \quad &\Rightarrow \quad (\mathscr{C} \approxeq_{\mathscr{F}_T} \mathscr{D} \wedge \mathscr{D} \approxeq_{\mathscr{F}_T} \mathscr{E}) && \text{[Th. 3.27 (ii), (iii)]} \\ &\Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \approxeq_{\mathscr{F}_T} \mathscr{E} && \text{[Th. 3.29]} \\ &\Rightarrow \quad \sigma(\mathscr{C} \cup \mathscr{D}) \precapprox_{\mathscr{F}_T} \mathscr{E}. && \text{[Def. 3.37 (i)]} \end{aligned}$$

Proof of Theorem 3.45

(i).

$$\begin{aligned} \mathscr{C} \precapprox_{\mathscr{F}_T} \mathscr{D} \quad &\Leftrightarrow \quad (\mathscr{C} \prec_{\mathscr{F}_T} \mathscr{D} \vee \mathscr{C} \approxeq_{\mathscr{F}_T} \mathscr{D}) && \text{[Def. 3.37 (i)]} \\ &\Leftrightarrow \quad \big(\mathscr{C} \prec_{\mathscr{F}_T} \sigma(\mathscr{D}) \vee \mathscr{C} \approxeq_{\mathscr{F}_T} \sigma(\mathscr{D})\big) && \text{[Ths. 3.9 (i), 3.28 (i), 3.27 (ii)]} \\ &\Leftrightarrow \quad \mathscr{C} \precapprox_{\mathscr{F}_T} \sigma(\mathscr{D}). && \text{[Def. 3.37 (i)]} \end{aligned}$$

(ii). Similarly,

$$\begin{aligned} \mathscr{C} \precapprox_{\mathscr{F}_T} \mathscr{D} \quad &\Leftrightarrow \quad (\mathscr{C} \prec_{\mathscr{F}_T} \mathscr{D} \vee \mathscr{C} \approxeq_{\mathscr{F}_T} \mathscr{D}) && \text{[Def. 3.37 (i)]} \\ &\Leftrightarrow \quad \big(\sigma(\mathscr{C}) \prec_{\mathscr{F}_T} \mathscr{D} \vee \sigma(\mathscr{C}) \approxeq_{\mathscr{F}_T} \mathscr{D}\big) && \text{[Ths. 3.9 (ii), 3.28 (i)]} \\ &\Leftrightarrow \quad \sigma(\mathscr{C}) \precapprox_{\mathscr{F}_T} \mathscr{D}. && \text{[Def. 3.37 (i)]} \end{aligned}$$

Proof of Theorem 3.46

(i).

$$(X \approxeq_{\mathscr{F}_T} Y \wedge Y \approxeq_{\mathscr{F}_T} Z) \quad \Rightarrow \quad X \approxeq_{\mathscr{F}_T} Z \qquad \text{[Box 3.2 (vi)]}$$

$$\Rightarrow \quad W \underset{\mathscr{F}_T}{\precsim} Z. \qquad \text{[Th. 3.35 (ii), Def. 3.37 (ii)]}$$

(ii).

$$\begin{aligned}
(X \underset{\mathscr{F}_T}{\precsim} Y \ \wedge\ Y \underset{\mathscr{F}_T}{\precsim} Z) \quad &\Leftrightarrow \quad \big(\sigma(X) \underset{\mathscr{F}_T}{\precsim} \sigma(Y) \ \wedge\ \sigma(Y) \underset{\mathscr{F}_T}{\precsim} \sigma(Z)\big) && \text{[Def. 3.37 (ii)]}\\
&\Rightarrow \quad \sigma(X) \underset{\mathscr{F}_T}{\precsim} \sigma(Z) && \text{[Th. 3.39 (iv)]}\\
&\Leftrightarrow \quad X \underset{\mathscr{F}_T}{\precsim} Z && \text{[Def. 3.37 (ii)]}\\
&\Rightarrow \quad W \underset{\mathscr{F}_T}{\precsim} Z. && \text{[Th. 3.35 (ii), (i)]}
\end{aligned}$$

Proof of Theorem 3.47

$$\begin{aligned}
X, Y \underset{\mathscr{F}_T}{\precsim} Z \quad &\Leftrightarrow \quad (X \underset{\mathscr{F}_T}{\precsim} Z \ \wedge\ Y \underset{\mathscr{F}_T}{\precsim} Z) && \text{[Prop. (3.48)]}\\
&\Leftrightarrow \quad \big(\sigma(X) \underset{\mathscr{F}_T}{\precsim} \sigma(Z) \ \wedge\ \sigma(Y) \underset{\mathscr{F}_T}{\precsim} \sigma(Z)\big) && \text{[Def. 3.37 (ii)]}\\
&\Rightarrow \quad \sigma\big(\sigma(X) \cup \sigma(Y)\big) \underset{\mathscr{F}_T}{\precsim} \sigma(Z) && \text{[Th. 3.44]}\\
&\Leftrightarrow \quad \sigma(X, Y) \underset{\mathscr{F}_T}{\precsim} \sigma(Z) && \text{[(3.20)]}\\
&\Rightarrow \quad (X, Y) \underset{\mathscr{F}_T}{\precsim} Z && \text{[Def. 3.37 (ii)]}\\
&\Rightarrow \quad W \underset{\mathscr{F}_T}{\precsim} Z. && [\sigma(W) \subset \sigma(X, Y),\ \text{Box 3.3 (iv)}]
\end{aligned}$$

3.7 Exercises

▷ **Exercise 3-1** Which are the elements of the σ-algebra $\sigma(U)$ in Example 3.2?

▷ **Exercise 3-2** Consider Table 1.2 and write down the inverse image $X^{-1}(A')$ of the set $A' = \{1\}$ in terms of a subset of $\Omega = \{\omega_1, \omega_2, \dots, \omega_8\}$.

▷ **Exercise 3-3** Consider the random experiment presented in Example 3.2 and enumerate all elements of the σ-algebra generated by X. Instead of specifying the value space $\big(\Omega'_X, \mathscr{P}(\Omega'_X)\big)$ of X with $\Omega'_X = \{0, 1\}$ as in Example 3.2, choose $(\mathbb{R}, \mathscr{B})$, where $\mathscr{B}$ denotes the Borel σ-algebra on the set $\mathbb{R}$ of real numbers (see RS-Rem. 1.14).

▷ **Exercise 3-4** Consider Table 1.2 and write down the inverse images

$$(U, X)^{-1}(A'_i) \ = \ \big\{\omega \in \Omega \colon (U, X)(\omega) \in A'_i\big\}, \quad i = 1, 2,$$

in terms of subsets of $\Omega = \{\omega_1, \omega_2, \dots, \omega_8\}$ for $A'_1 = \{(Joe, 0)\}$ and $A'_2 = \{(Joe, 1), (Ann, 0), (Ann, 1)\}$.

▷ **Exercise 3-5** Consider Example 3.5 and name three more pairs of sets systems such that one is prior in $\mathscr{F}_T$ to the other.

▷ **Exercise 3-6** Prove $(A \subset B \ \wedge\ A \subset C) \quad \Leftrightarrow \quad A \subset (B \cap C)$.

▷ **Exercise 3-7** Prove$(A \subset C \ \wedge\ B \subset C) \quad \Leftrightarrow \quad (A \cup B) \subset C$.

▷ **Exercise 3-8** Prove: $\big(A \subset B \ \wedge\ B \subset C\big) \ \Rightarrow \ A \subset C$.

▷ **Exercise 3-9** Prove: If X is a real-valued measurable map on $(\Omega, \mathscr{A})$ and $W = \alpha \cdot X$, $\alpha \in \mathbb{R}$, then W is X-measurable.

▷ **Exercise 3-10** Show: If $W = \beta_1 X_1 + \beta_2 X_2$ with $\beta_1, \beta_2 \in \mathbb{R}$, then W is (X_1, X_2)-measurable.

▷ **Exercise 3-11** Prove Proposition (iv) of Box 3.1.

▷ **Exercise 3-12** Prove Propositions (x) and (xi) of Box 3.2.

Solutions

▷ **Solution 3-1** The probability space $(\Omega, \mathscr{A}, P)$ representing the random experiment described in Table 1.2 is specified in Example 3.2. Furthermore, the random variable U is specified in Table 1.2. The σ-algebra $\sigma(U)$ has four elements. Aside from Ω and $\emptyset$, these are the events

$$C = \{Joe\} \times \Omega_X \times \Omega_Y = \big\{(Joe, no, -), (Joe, yes, -), (Joe, no, +), (Joe, yes, +)\big\}$$

that *Joe is drawn* and

$$C^c = \{Ann\} \times \Omega_X \times \Omega_Y = \big\{(Ann, no, -), (Ann, yes, -), (Ann, no, +), (Ann, yes, +)\big\}$$

that *Ann is drawn*.

▷ **Solution 3-2** This inverse image is $X^{-1}(A') = X^{-1}(\{1\}) = \{\omega_3, \omega_4, \omega_7, \omega_8\}$. This is the event that X takes on the value 1.

▷ **Solution 3-3** There are four different inverse images of sets $B \in \mathscr{B}$ under X:

$$\forall B \in \mathscr{B}: \quad X^{-1}(B) = \begin{cases} \{\omega_1, \omega_2, \omega_5, \omega_6\}, & \text{if } 0 \in B \text{ and } 1 \notin B \\ \{\omega_3, \omega_4, \omega_7, \omega_8\}, & \text{if } 0 \notin B \text{ and } 1 \in B \\ \Omega, & \text{if } 0 \in B \text{ and } 1 \in B \\ \emptyset, & \text{if } 0 \notin B \text{ and } 1 \notin B. \end{cases}$$

These four inverse images are the elements of $\sigma(X)$. The same four elements are obtained if, instead of $(\mathbb{R}, \mathscr{B})$, we choose $\big(\Omega'_X, \mathscr{P}(\Omega'_X)\big) = \big(\{0,1\}, \big\{\{0\}, \{1\}, \{0,1\}, \emptyset\big\}\big)$ as the value space of X (cf. Exercise 3-2).

▷ **Solution 3-4** The first inverse image is

$$(U, X)^{-1}(A'_1) = (U, X)^{-1}\big(\{(Joe, 0)\}\big) = \{\omega_1, \omega_2\}.$$

This is the event that *Joe is drawn and not treated*. The second inverse image is

$$(U, X)^{-1}(A'_2) = (U, X)^{-1}\big(\{(Joe, 1), (Ann, 0), (Ann, 1)\}\big) = \{\omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\}$$

[consider again Eqs. (3.11) to (3.13)]. This is the event that *Joe is drawn and treated or Ann is drawn.*

▷ **Solution 3-5** $\sigma(U)$ is prior in $\mathscr{F}_T$ to $\sigma(U, X)$, $\sigma(U)$ is prior to $\sigma(U, X, Y)$, and $\sigma(X)$ is prior to $\sigma(U, X, Y)$.

▷ **Solution 3-6**

$$\begin{aligned} (A \subset B \ \wedge \ A \subset C) \quad &\Leftrightarrow \quad \forall a \in A\colon a \in B \ \wedge \ a \in C \\ &\Leftrightarrow \quad \forall a \in A\colon a \in (B \cap C) \\ &\Leftrightarrow \quad A \subset (B \cap C). \end{aligned}$$

▷ **Solution 3-7**

$$\begin{aligned}(A \subset C \,\wedge\, B \subset C) &\Leftrightarrow \forall a \in A\colon a \in C \,\wedge\, \forall b \in B\colon b \in C\\ &\Leftrightarrow \forall c \in A \cup B\colon c \in C\\ &\Leftrightarrow A \cup B \subset C.\end{aligned}$$

▷ **Solution 3-8**

$$\begin{aligned}(A \subset B \,\wedge\, B \subset C) &\Leftrightarrow \forall a \in A\colon a \in B \,\wedge\, \forall b \in B\colon b \in C\\ &\Rightarrow \forall a \in A\colon a \in C\\ &\Leftrightarrow A \subset C.\end{aligned}$$

▷ **Solution 3-9** If $\alpha \in \mathbb{R}$, then $W = \alpha \cdot X$ is the composition $g(X)$ of the measurable maps $X\colon (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ and $g\colon (\overline{\mathbb{R}}, \overline{\mathscr{B}}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ defined by

$$g(x) = \alpha \cdot x, \quad \forall x \in \overline{\mathbb{R}}.$$

According to Remark 3.17, W is X-measurable.

▷ **Solution 3-10** If $\beta_1, \beta_2 \in \mathbb{R}$, then $W = \beta_1 X_1 + \beta_2 X_2$ is the composition $g(X_1, X_2)$ of the bivariate measurable map $(X_1, X_2)\colon (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}^{\,2}, \overline{\mathscr{B}}_2)$ and the measurable map $g\colon (\overline{\mathbb{R}}^{\,2}, \overline{\mathscr{B}}_2) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ defined by

$$g(x_1, x_2) = \beta_1 x_1 + \beta_2 x_2, \quad \forall (x_1, x_2) \in \overline{\mathbb{R}}^{\,2}.$$

According to Remark 3.17, W is (X_1, X_2)-measurable.

▷ **Solution 3-11**

$$\begin{aligned}X \underset{\mathscr{F}_T}{\lesssim} Y &\Leftrightarrow \sigma(X) \underset{\mathscr{F}_T}{\lesssim} \sigma(Y) && \text{[Def. 3.3 (ii)]}\\ &\Leftrightarrow \sigma(X) \underset{\mathscr{F}_T}{\lesssim} \sigma\big(\sigma(X) \cup \sigma(Y)\big) && \text{[Th. 3.13]}\\ &\Leftrightarrow \sigma(X) \underset{\mathscr{F}_T}{\lesssim} \sigma(X, Y) && \text{[(3.20)]}\\ &\Rightarrow \sigma(W) \underset{\mathscr{F}_T}{\lesssim} \sigma(X, Y) && \text{[Th. 3.8]}\\ &\Leftrightarrow W \underset{\mathscr{F}_T}{\lesssim} (X, Y). && \text{[Def. 3.3 (ii)]}\end{aligned}$$

▷ **Solution 3-12** Box 3.2 (x).

$$\begin{aligned}&(X \underset{\mathscr{F}_T}{\approx} Y \,\wedge\, Y \underset{\mathscr{F}_T}{\lesssim} Z)\\ \Leftrightarrow\ &\big(\sigma(X) \underset{\mathscr{F}_T}{\approx} \sigma(Y) \,\wedge\, \sigma(Y) \underset{\mathscr{F}_T}{\lesssim} \sigma(Z)\big) && \text{[Defs. 3.3 (ii), 3.24 (ii)]}\\ \Leftrightarrow\ &\Big(\sigma\big(\sigma(X) \cup \sigma(Y)\big) \underset{\mathscr{F}_T}{\approx} \sigma(Y) \,\wedge\, \sigma(Y) \underset{\mathscr{F}_T}{\lesssim} \sigma(Z)\Big) && \text{[Th. 3.28 (i)]}\\ \Rightarrow\ &\sigma(W) \underset{\mathscr{F}_T}{\lesssim} \sigma(Z) && [\sigma(W) \subset \sigma(X, Y),\ \text{Th. 3.33 (i)}]\\ \Leftrightarrow\ &W \underset{\mathscr{F}_T}{\lesssim} Z. && \text{[Def. 3.3 (ii)]}\end{aligned}$$

Box 3.2 (xi).

$$\begin{aligned}&(X \underset{\mathscr{F}_T}{\lesssim} Y \,\wedge\, Y \underset{\mathscr{F}_T}{\approx} Z)\\ \Leftrightarrow\ &\big(\sigma(X) \underset{\mathscr{F}_T}{\lesssim} \sigma(Y) \,\wedge\, \sigma(Y) \underset{\mathscr{F}_T}{\approx} \sigma(Z)\big) && \text{[Defs. 3.3 (ii), 3.24 (ii)]}\\ \Rightarrow\ &\sigma(W) \underset{\mathscr{F}_T}{\lesssim} \sigma(Z) && [\sigma(W) \subset \sigma(X),\ \text{Th. 3.33 (ii)}]\\ \Leftrightarrow\ &W \underset{\mathscr{F}_T}{\lesssim} Z. && \text{[Def. 3.3 (ii)]}\end{aligned}$$

Chapter 4
Regular Causality Space and Potential Confounder

In chapter 1 we studied some examples showing that the conditional expectation values $E(Y|X{=}x)$ of an outcome variable Y and their differences $E(Y|X{=}x) - E(Y|X{=}x')$, the *prima facie effects*, can be seriously misleading in evaluating the causal effect of a treatment variable X on an outcome variable Y. These examples demonstrate that the standard probabilistic concepts — such as conditional expectation values or conditional probabilities — cannot be used offhandedly to define the causal effects in which we are interested when we have to evaluate a treatment, an intervention, or an exposition. In chapter 2, we described random experiments of various research designs in which a causal total effect is of interest. In chapter 3 we started the mathematical theory of causal effects introducing the prior-to, simultaneous-to, and prior-or-simultaneous-to relations among measurable set systems and among measurable maps. In the context of a probability space $(\Omega, \mathscr{A}, P)$, these relations also apply to sets of events and random variables, respectively.

Overview

In the present chapter, we introduce the concepts of a *regular causality space* and a *regular causality setup*. A *regular causality setup* provides the mathematical structure that allows us to define a *putative cause variable* X, an *outcome variable* Y *of* X, a *potential confounder of* X, and a *potential mediator between* X *and* Y. In many cases, conditional expectations describing a causal dependence can be distinguished from conditional expectations that have no such causal interpretation by their relationship to the potential confounders of the putative cause variable X. This will be substantiated in the chapters on causality conditions.

Just as in chapter 3, none of the concepts treated in this chapter involves a probability measure. Instead, a measurable space $(\Omega, \mathscr{A})$ (i. e., a set Ω and a σ-algebra $\mathscr{A}$ on Ω) suffices. Note, however, that all properties of a regular causality space still hold if we add a probability measure P on $(\Omega, \mathscr{A})$, considering a probability space $(\Omega, \mathscr{A}, P)$. This will be necessary as soon as we turn to random variables and their stochastic dependencies on each other (see the chapters to come). In empirical applications, $(\Omega, \mathscr{A}, P)$ represents a concrete random experiment, Ω the set of possible outcomes, and $\mathscr{A}$ the set of possible events in this random experiment.

Prerequisites

Reading this chapter requires that the reader is familiar with the concepts of a σ-*algebra*, a *measurable space*, and a *measurable map* as treated, for example, in the first two chapters of Steyer (2024). These chapters will be referred to as RS-chapter 1 and RS-chapter 2, and the same kind of shortcut is used when referring to other parts of that book.

4.1 Regular Causality Space and Setup

In this section, we introduce the notions of a *regular causality space* and a *regular probabilistic causality space*. A regular causality space is the minimal formal framework in which we can introduce the concepts of a putative cause variable, a potential confounder, a potential mediator, and an outcome variable. In a regular probabilistic causality space we can define true outcome variables, (causal) unbiasedness, and causality conditions, which imply unbiasedness of various conditional expectations and their differences.

4.1.1 Regular Causality Space

Defining a regular causality space (see Def. 4.6), we refer to a measurable space $(\Omega, \mathscr{A})$ that is assumed to be the product of four measurable spaces $(\Omega_t, \mathscr{A}_t)$, $t \in T = \{1, 2, 3\}$ (see RS-Def. 1.15). This implies that Ω is the set product of the four sets Ω_1, Ω_2, and Ω_3, each of which can itself be a set product of other sets.

Remark 4.1 [Intuitive Background of the Product Space $(\Omega, \mathscr{A})$**]** When we consider the effect of a putative cause variable X (treatment, intervention, or exposition variable) on an outcome variable Y that is assessed for an observational unit to be sampled (oftentimes a person), then there are variables that are *prior* to X. They represent attributes of the observational unit *before* treatment. It is obvious that these pretreatment attributes or their fallible observations cannot be caused by a subsequent treatment. Simple examples are *age*, *sex*, *race*, *socioeconomic status*, or *educational status before* treatment. The set Ω_1 occurring in Definition 4.6 should be chosen such that these pretreatment variables solely depend on the elements of Ω_1.

Next, there might be variables that are *simultaneous* to X, for example, a second treatment variable that varies at the same time as X. For example, you can drink coffee (or not) and alcohol (or not) at the same time. The set Ω_2 in Definition 4.6 should be chosen such that variables that are simultaneous to a focused putative cause variable X (including X itself) only depend on the elements of Ω_2.

Finally, the set Ω_3 should be chosen such that the outcome variable Y depends, at least in part, on the elements of this set. For details and a mathematical formulation of these ideas, see Definitions 4.6 and 4.11. $\triangleleft$

Remark 4.2 [Intuitive Background of the Projections] Definition 4.6 refers to the projections (or coordinate maps) π_1 to π_3 (see, e. g., RS-Def. 2.27). Intuitively speaking, the projection π_1 is a map that contains the information about all events that are *prior to the putative cause variable*. In any case, all nonconstant maps that only depend — in the sense of measurability, not in the sense of a probabilistic dependence — on π_1 are potential confounders.

In contrast, π_2 contains the information about all events that are *simultaneous to the putative cause variable*, say X, including the events represented by X itself. The projection π_2 can be multidimensional, that is, it may consist of several projections π_{2j}. Hence, $\pi_2 = (\pi_{2j}, j \in J)$ is a family of projections for a nonempty index set J.

Finally, the projection π_3 contains the information about all events that are *simultaneous or posterior to the outcome variable*, say Y. We will require that Y depends at least in part on π_3. This will allow us to choose $V_2 - V_1$ as an outcome variable, where V_2 is a function of π_3, and V_1 a function of π_1. Examples for V_1 and V_2 are *achievement*, *well-being*, or *health scores*, before and after treatment, respectively. $\lhd$

Remark 4.3 [Filtration $(\mathscr{F}_t)_{t\in T}$] In Definition 4.6 we specify a specific filtration $(\mathscr{F}_t)_{t\in T}$, $T = \{1,2,3\}$. From the perspective of the focused putative cause variable, the σ-algebras $\mathscr{F}_1$ and $\mathscr{F}_2$ represent the sets of past and present events, respectively. In contrast, $\mathscr{F}_3$ represents all events that are posterior to X. $\lhd$

Remark 4.4 [Cause σ-Algebra] In Remark 4.2, we already discussed the relation between the projection π_2 and a putative cause variable X. In Definition 4.6 (i), this relationship is made explicit for the cause σ-algebra $\mathscr{C}$ with respect to which we will require a putative cause variable X to be measurable (see Def. 4.11). $\lhd$

Remark 4.5 [Potential Confounder σ-Algebra] Intuitively speaking, we consider all variables as potential confounders of a putative cause variable X that are prior or simultaneous to X, except for X itself. In Definition 4.6 we introduce the concept of a *confounder σ-algebra of* $\mathscr{C}$, denoted $\mathscr{D}_{\mathscr{C}}$. This σ-algebra $\mathscr{D}_{\mathscr{C}}$ can be interpreted to represent the set of past and present possible events from the perspective of the cause σ-algebra $\mathscr{C}$, except for those that are represented by $\mathscr{C}$ itself. The σ-algebra $\mathscr{D}_{\mathscr{C}}$ also plays a crucial role in the definition of a *true outcome variable* and all other concepts based thereon (see ch. 5). $\lhd$

Definition 4.6 [Regular Causality Space]
Let $(\Omega,\mathscr{A})$ denote the product of the measurable spaces $(\Omega_t,\mathscr{A}_t)$, $t\in T=\{1,2,3\}$, where for all $\omega_t\in\Omega_t$, $t\in T$, we assume $\{\omega_t\}\in\mathscr{A}_t$. Furthermore, let $\pi_t\colon\Omega\to\Omega_t$, $t\in T$, denote the projection with $\pi_t(\omega)=\omega_t$, for all $\omega\in\Omega$, and define the filtration $(\mathscr{F}_t)_{t\in T}$ by

$$\mathscr{F}_1 := \sigma(\pi_1), \qquad \mathscr{F}_2 := \sigma(\pi_1,\pi_2), \qquad \mathscr{F}_3 := \sigma(\pi_1,\pi_2,\pi_3). \tag{4.1}$$

Finally, let $\pi_2 = (\pi_{2j},\, j\in J)$, where J is a nonempty finite subset of $\mathbb{N}$ and assume that $\{\Omega,\emptyset\}\neq\sigma(\pi_{2j},\, j\in K)$, $\emptyset\neq K\subset J$. Then we call

(i) $\mathscr{C} := \sigma(\pi_{2j},\, j\in K)$, $\emptyset\neq K\subset J$, *the cause σ-algebra*
(ii) $\mathscr{D}_{\mathscr{C}} := \sigma(\pi_1,\pi_{2j},\, j\in J\setminus K)$ *the confounder σ-algebra of* $\mathscr{C}$
(iii) $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$ *a regular causality space.*

Remark 4.7 [Regular Probabilistic Causality Space] Beginning with chapter 5, we will additionally consider a probability measure P on the σ-algebra $\mathscr{A}$, and then

$$\big((\Omega,\mathscr{A},P),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$$

will be called a *regular probabilistic causality space*. $\lhd$

Remark 4.8 [Cause σ-Algebra $\mathscr{C}$ and Confounder σ-Algebra $\mathscr{D}_{\mathscr{C}}$] According to Definition 4.6 (i), the cause σ-algebra $\mathscr{C}$ can be chosen to be generated by π_2 itself or by any proper subset of the projections π_{2j}, $j\in J$, that constitute π_2. This choice of the cause σ-algebra $\mathscr{C}$ also determines the confounder σ-algebra $\mathscr{D}_{\mathscr{C}}$. In particular, the choice of the index set $K\subset J$ and with it, the choice of $\mathscr{C}$, determines which projections π_{2j} generate, together with π_1, the confounder σ-algebra $\mathscr{D}_{\mathscr{C}}$ [see Def. 4.6 (ii)]. $\lhd$

Table 4.1. Joe and Ann with a single treatment variable

Possible outcomes	Unit	Treatment	Success	π_1	π_2	π_3	Treatment variable X	Outcome variable Y
$\omega_1 = (Joe,\ no,\ -)$				*Joe*	*no*	$-$	0	0
$\omega_2 = (Joe,\ no,\ +)$				*Joe*	*no*	$+$	0	1
$\omega_3 = (Joe,\ yes,\ -)$				*Joe*	*yes*	$-$	1	0
$\omega_4 = (Joe,\ yes,\ +)$				*Joe*	*yes*	$+$	1	1
$\omega_5 = (Ann,\ no,\ -)$				*Ann*	*no*	$-$	0	0
$\omega_6 = (Ann,\ no,\ +)$				*Ann*	*no*	$+$	0	1
$\omega_7 = (Ann,\ yes,\ -)$				*Ann*	*yes*	$-$	1	0
$\omega_8 = (Ann,\ yes,\ +)$				*Ann*	*yes*	$+$	1	1

(Column groups: Possible outcomes — Unit, Treatment, Success; Projections — π_1, π_2, π_3; Observables — Treatment variable X, Outcome variable Y.)

Remark 4.9 [$(\mathscr{F}_t)_{t\in T}$ **Is a Filtration in** $\mathscr{A}$] The family of σ-algebras $(\mathscr{F}_t)_{t\in T}$, $T = \{1,2,3\}$, specified in the definition of a regular causality space [see Eq. (4.1)] is a filtration in $\mathscr{A}$. This immediately follows from RS-Equation (1.3) and RS-Equation (2.15). ◁

Example 4.10 [Joe and Ann With a Single Treatment Variable] We illustrate the concepts treated in Definition 4.6 by the kind of experiment presented in Table 4.1. It is the same kind of experiment as already presented in Table 1.2. The first part of the experiment consists of sampling a person u from the set $\Omega_1 = \{Joe, Ann\}$. Then the sampled person receives (*yes*) or does not receive (*no*) a treatment. The two treatment conditions are the elements of the set $\Omega_2 = \{no, yes\}$. Finally, it is observed whether or not a success criterion is reached at some appropriate time *after* treatment. These two possible outcomes are the elements of the set $\Omega_3 = \{-,+\}$. Hence, the sets Ω_t, $t \in T = \{1,2,3\}$, occurring in Definition 4.6 are

$$\Omega_1 = \{Joe, Ann\}, \qquad \Omega_2 = \{no, yes\}, \qquad \Omega_3 = \{-,+\}.$$

Furthermore, we choose the σ-algebras on these sets to be $\mathscr{A}_t = \mathscr{P}(\Omega_t)$, $t \in T$, the corresponding power sets. Because, in Definition 4.6, we require that $(\Omega, \mathscr{A})$ is the product of the measurable spaces $(\Omega_t, \mathscr{A}_t)$ (see again RS-Def. 1.15), the measurable space $(\Omega, \mathscr{A})$ is completely specified, and with it the projections π_1, π_2, and π_3. The assignment of values of these projections to all elements ω_i of Ω is explicitly shown in Table 4.1. Note that the projection π_1 is identical to the person variable U specified in Table 1.2.

The σ-algebras of the filtration $(\mathscr{F}_t)_{t\in T}$, $T = \{1,2,3\}$, are

$$\mathscr{F}_1 = \sigma(\pi_1), \qquad \mathscr{F}_2 = \sigma(\pi_1, \pi_2), \qquad \mathscr{F}_3 = \sigma(\pi_1, \pi_2, \pi_3) \tag{4.2}$$

(see Exercises 4-1 to 4-4). Furthermore, according to Definition 4.6, in this kind of experiment,

$$\mathscr{C} \;=\; \sigma(\pi_2) \;=\; \big\{\{\omega_1,\omega_2,\omega_5,\omega_6\},\{\omega_3,\omega_4,\omega_7,\omega_8\},\Omega,\emptyset\big\} \tag{4.3}$$

is necessarily the cause σ-algebra because $J = K = \{1\}$ implying $\pi_2 = \pi_{21} = (\pi_{2j}, j \in K)$ (see Exercises 4-5 and 4-6). Finally,

$$\mathscr{D}_{\mathscr{C}} \;=\; \sigma(\pi_1) \;=\; \big\{\{\omega_1,\dots,\omega_4\},\{\omega_5,\dots,\omega_8\},\Omega,\emptyset\big\} \tag{4.4}$$

is the confounder σ-algebra of $\mathscr{C}$ because $J = K$. Hence, we completely specified the regular causality space $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$.

Note that the structure of this kind of experiment is essentially the same for every simple experiment of the type described in section 2.1. Only the numbers of values of the projections π_t increase if we consider more than two observational units, more than two treatment conditions, or more than two values of the outcome variable. ◁

4.1.2 Regular Causality Setup and Potential Confounder

Now we introduce a *regular causality setup*, which, additionally to a regular causality space, comprises a *putative cause variable* and an *outcome variable*. Furthermore, we define the concepts of a *potential confounder of a putative cause variable* and a *global potential confounder of a putative cause variable*. In particular it is explicated how these maps are related to a regular causality space.

Definition 4.11 [Regular Causality Setup and Potential Confounder]
Let $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$ be a regular causality space and X, Y, D_X, W measurable maps on $(\Omega,\mathscr{A})$. Then we call

(i) *X a putative cause variable of Y if*

  (a) $\{\Omega,\emptyset\} \neq \sigma(X) \subset \mathscr{C} = \sigma(\pi_{2j}, j \in K)$

  (b) $\neg\exists\, K_0 \subset K,\ K_0 \neq K,$ *such that* $\sigma(X) \subset \sigma(\pi_{2j}, j \in K_0)$

(ii) *Y an outcome or response variable if* $\sigma(Y) \not\subset \mathscr{F}_2$.

(iii) *D_X a global potential confounder of X if* $\sigma(D_X) = \mathscr{D}_{\mathscr{C}}$

(iv) *W a potential confounder of X if* $\sigma(W) \subset \sigma(D_X)$

(v) $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$ *a regular causality setup.*

A potential confounder W of X is called trivial if $\sigma(W) = \{\Omega,\emptyset\}$.

In Definition 4.11, we require that all variables introduced in this definition are measurable maps on the same measurable space $(\Omega,\mathscr{A})$. Specifying $(\Omega,\mathscr{A})$ and $(\mathscr{F}_t)_{t\in T}$, we fix which kind of experiment we are talking about, and with X and Y, we choose the putative cause variable and the outcome variable, respectively. Finally, with D_X we specify the global potential confounder of X. Note that the *set of all potential confounders of X* is the set of all measurable maps on $(\Omega,\mathscr{A})$ that are D_X-measurable.

Remark 4.12 [Regular Probabilistic Causality Setup] As mentioned before, in chapter 5, we will additionally consider a probability measure P on the σ-algebra $\mathscr{A}$, and then

$$\big((\Omega,\mathscr{A},P),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$$

will be called a *regular probabilistic causality setup*. ◁

Remark 4.13 [Putative Cause Variable] In Definition 4.11 (i), we postulate $\sigma(X) \neq \{\Omega, \emptyset\}$. Hence, a putative cause variable X is not a constant. Requiring $\sigma(X) \subset \mathscr{C}$ — and not $\sigma(X) = \mathscr{C}$ — means that X is not necessarily the only putative cause variable we might consider in the framework of a given regular causality space. Instead we can also coarsen an original putative cause variable and consider a new one with fewer values (see the example in sect. 4.2.1). According to condition (b) of Definition 4.11 (i), there is no proper subset K_0 of K such that a putative cause variable X is measurable with respect to $(\pi_{2j}, j \in K_0)$. This requirement secures, for example, that we cannot choose a putative cause variable that is measurable with respect to a single projection, say π_{21}, if we choose $\mathscr{C}$ to be generated by (π_{21}, π_{22}). If, in a setting with $\pi_2 = (\pi_{21}, \pi_{22})$, a π_{21}-measurable putative cause variable X is desired, then we have to choose $\mathscr{C} = \sigma(\pi_{21})$, implying $\sigma(\pi_{22}) \subset \mathscr{D}_{\mathscr{C}} = \sigma(\pi_1, \pi_{22})$. Hence, if $\pi_2 = (\pi_{21}, \pi_{22})$ and $\sigma(X) \subset \sigma(\pi_{21})$, then the projection π_{22} is a potential confounder of X. For more details see the first regular causality setup presented in the example of section 4.2.2. ◁

Remark 4.14 [Outcome Variable] Requiring $\sigma(Y) \not\subset \mathscr{F}_2$, we secure that an outcome variable is neither a constant nor a potential confounder of X. We allow for outcome variables that are functions of a potential confounder and a map that is measurable with respect to $\sigma(\pi_3)$. A typical example is a change variable $Y = V_2 - V_1$ between a $\sigma(\pi_3)$-measurable variable V_2 and a $\sigma(\pi_1)$-measurable V_1 — for example, a pretest — assessing the same attribute as V_2, but before treatment. In this case, the change variable still depends on the elements of Ω_3, but not exclusively (see Rem. 4.2).

Of course, a $\sigma(\pi_3)$-measurable map is an outcome variable as well. The definition of an outcome variable also allows us to rescale, coarsen, or aggregate a new outcome variable from other outcome variables. For example, instead of considering a ten-dimensional outcome variable $(Y_1, \ldots, Y_{10})$ representing the binary answers to ten items in a questionnaire, one might be interested in a uni-dimensional outcome variable Y defined as the sum $Y_1 + \ldots + Y_{10}$. ◁

Remark 4.15 [Potential Confounder] According to Definition 4.11 (iv), each non-constant measurable map W satisfying $\sigma(W) \subset \mathscr{D}_{\mathscr{C}}$ is a potential confounder of X. Hence, each pretreatment variable W — which is measurable with respect to the projection π_1 — is a potential confounder of a putative cause variable X [see Def. 4.11 (iv)]. However, a treatment variable can also be a potential confounder of another treatment variable if this other treatment variable is focused as the putative cause variable [see Defs. 4.6 (ii) and 4.11 (i)]. ◁

Remark 4.16 [Covariate of X] If W is a potential confounder of X, then we also call it a *covariate of X*, in particular in a context in which we condition on W and X in a conditional expectation such as $E(Y \mid X, W)$ or in a conditional distribution $P_{Y \mid X, W}$. From a mathematical point of view, the terms *potential confounder of X* and *covariate of X* are synonyms. Note, however, that, in the statistical literature, the term 'covariate' is not used unanimously. Sometimes it even refers to a putative cause variable of Y. ◁

Remark 4.17 [Global Potential Confounder] According to Definition 4.11 (iii), each non-constant measurable map D_X satisfying $\sigma(D_X) = \mathscr{D}_{\mathscr{C}}$ is a *global* or *comprehensive potential confounder of X*. Hence, instead of specifying the σ-algebra $\mathscr{D}_{\mathscr{C}}$ we may also specify a global potential confounder D_X, which may often be more convenient. The concept of a global potential confounder also plays a crucial role in the definition of a *true outcome*

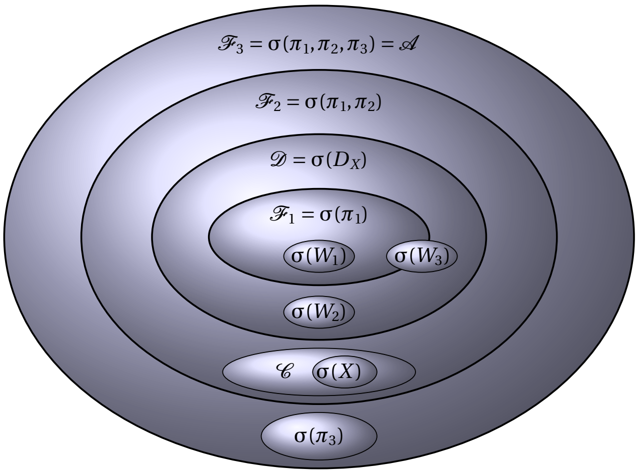

Figure 4.1. The filtration $(\mathscr{F}_t)_{t \in T}$ and various σ-algebras in a regular causality space.

variable and all other concepts based thereon (see ch. 5). The σ-algebra $\sigma(D_X)$ generated by D_X can be interpreted to represent, from the perspective of the cause σ-algebra $\mathscr{C}$, the set of past and present possible events, except for those that are represented by $\mathscr{C}$ itself. ◁

Remark 4.18 [There Are Several Global Potential Confounders] There are always several global potential confounders of X. For example, in Table 4.1, the projection π_1 and the indicator variable $1_{U=Joe}$ of the event that Joe is sampled generate the same σ-algebra, which is also the confounder σ-algebra if we consider the treatment variable X as the putative cause variable. Therefore, we talk about 'a' — and not about 'the' — global potential confounder of X unless a specific global potential confounder is already specified. In contrast, the confounder σ-algebra $\mathscr{D}_{\mathscr{C}} = \sigma(D_X)$ is uniquely determined, once the measurable space $(\Omega, \mathscr{A})$, the filtration $(\mathscr{F}_t)_{t \in T}$, and the cause σ-algebra $\mathscr{C}$ are specified (see Def. 4.6). ◁

Remark 4.19 [Reference to X] According to Definition 4.11, a potential confounder of X is every non-constant map that is measurable with respect to the σ-algebra generated by a global potential confounder D_X of X. Note that the reference to X is important. Choosing another putative cause variable may lead to another confounder σ-algebra [see the examples in sect. 4.2] and with it, to other potential confounders. ◁

Figure 4.1 shows the subset relationships between the various σ-algebras in a regular causality setup. These relationships follow from the conditions specified in Definitions 4.6 and 4.11. The figure also shows which of the σ-algebras occur for the first time in one of the four σ-algebras $\mathscr{F}_1$, $\mathscr{F}_2$, and $\mathscr{F}_3$. Note that W_1, W_2, and W_3 denote potential confounders of X. The σ-algebra $\sigma(Y)$ generated by the outcome variable Y is not shown in the figure. It can be a subset of $\sigma(\pi_3)$ or a any other subset of $\mathscr{A}$ as long as it is not a subset of $\mathscr{F}_2$.

4.1.3 Restriction of a Regular Causality Space and Setup

In experiments in which we have more than two treatment conditions, for example, treatment a, treatment b, and control, we might be interested in comparing treatment a or treatment b against control, or treatment a against treatment b. In each of these cases, we use putative cause variables that are not maps on the original measurable space $(\Omega, \mathscr{A})$ but on $(\Omega_0, \mathscr{A}|_{\Omega_0})$, where Ω_0 is a subset of Ω and $\mathscr{A}|_{\Omega_0}$ denotes the *restriction of the σ-algebra* $\mathscr{A}$ *to* Ω_0. For detailed examples see sections 4.2.1 and 4.2.2.

Remark 4.20 [Restriction of a Set System and a σ-Algebra] Let $\mathscr{E}$ be a set of subsets of a nonempty set Ω and $\Omega_0 \subset \Omega$. Then

$$\mathscr{E}|_{\Omega_0} := \{\Omega_0 \cap A \colon A \in \mathscr{E}\} \tag{4.5}$$

is called the *restriction of* $\mathscr{E}$ *to* Ω_0 or the *trace of* $\mathscr{E}$ *in* Ω_0. If $\mathscr{A}$ is a σ-algebra on Ω, then $\mathscr{A}|_{\Omega_0}$ is a σ-algebra on Ω_0. (For a proof see SN-Exercise 1-5). ◁

Remark 4.21 [Restriction of a Map] In the sequel we will also use the *restriction of a map* $f\colon \Omega \to \Omega'$ *to the subset* Ω_0 of Ω, defined by

$$f|_{\Omega_0}(\omega) = f(\omega), \quad \text{for all } \omega \in \Omega_0. \tag{4.6}$$

This concept also applies to multivariate maps $f = (f_1, \ldots, f_m)$ consisting of several maps $f_1, \ldots, f_m$. In this case

$$f|_{\Omega_0} = (f_1, \ldots, f_m)|_{\Omega_0} = \big(f_1|_{\Omega_0}, \ldots, f_m|_{\Omega_0}\big). \tag{4.7}$$

Note that we can also consider the restrictions

$$\pi_t|_{\Omega_0}(\omega) = \pi_t(\omega), \quad \text{for all } \omega \in \Omega_0, \tag{4.8}$$

of the projections π_t, $t \in T$, introduced in Definition 4.6. That is, each map $\pi_t|_{\Omega_0}$ is the *restriction of* the corresponding map π_t *to the subset* Ω_0 of Ω. ◁

A regular causality space $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$ is a structure that entirely consists of σ-algebras on the set Ω. According to the following theorem, the restrictions of these σ-algebras to a subset Ω_0 of Ω constitute a new regular causality space, which is the framework for putative cause variables that are measurable maps on Ω_0 (see sect. 4.2.1 for a detailed example).

Theorem 4.22 [Restriction of a Regular Causality Space]
Let $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$ *be a regular causality space and*

$$\Omega_0 = (\Omega_{01} \times \Omega_{02} \times \Omega_{03}) \subset \Omega.$$

Then

$$\begin{aligned} \mathscr{F}_1|_{\Omega_0} &= \sigma(\pi_1|_{\Omega_0}) \\ \mathscr{F}_2|_{\Omega_0} &= \sigma\big((\pi_1, \pi_2)|_{\Omega_0}\big) \\ \mathscr{F}_3|_{\Omega_0} &= \sigma\big((\pi_1, \pi_2, \pi_3)|_{\Omega_0}\big) \end{aligned} \tag{4.9}$$

and

$$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)|_{\Omega_0} := \Big((\Omega_0,\mathscr{A}|_{\Omega_0}),(\mathscr{F}_t|_{\Omega_0})_{t\in T},\mathscr{C}|_{\Omega_0},\mathscr{D}_{\mathscr{C}}|_{\Omega_0}\Big) \tag{4.10}$$

is a regular causality space. It is called the restriction of $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$ *to* Ω_0.

(Proof p. 106)

Remark 4.23 [Restriction of a Putative Cause Variable and a Potential Confounder] Let us also consider the *restriction of the putative cause variable X to* Ω_0,

$$X|_{\Omega_0}(\omega) \;=\; X(\omega), \quad \text{for all } \omega \in \Omega_0, \tag{4.11}$$

and the *restriction of the outcome variable Y to* Ω_0,

$$Y|_{\Omega_0}(\omega) \;=\; Y(\omega), \quad \text{for all } \omega \in \Omega_0. \tag{4.12}$$

Adding these maps to the restriction of a regular causality space yields a new regular causality setup. ◁

Definition 4.24 [Restriction of a Regular Causality Setup]
Let the assumptions of Theorem 4.22 hold. Then

$$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)|_{\Omega_0} := \Big((\Omega_0,\mathscr{A}|_{\Omega_0}),(\mathscr{F}_t|_{\Omega_0})_{t\in T},\mathscr{C}|_{\Omega_0},\mathscr{D}_{\mathscr{C}}|_{\Omega_0},X|_{\Omega_0},Y|_{\Omega_0}\Big)$$

is called the restriction of $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$ *to* Ω_0.

4.2 Examples

4.2.1 Joe and Ann With Three Treatment Conditions

We illustrate the concepts treated in section 4.1 by the kind of experiment presented in Table 4.2. In this kind of experiment, we first sample a person u from the set $\Omega_1 = \{Joe, Ann\}$. Then the sampled person receives treatment *a*, *b*, or *c*. These treatment conditions are the elements of the set $\Omega_2 = \{a,b,c\}$. In this experiment, there is no other treatment variable. Finally, it is observed whether or not a success criterion is reached at some appropriate time after treatment. These two possible outcomes are the elements of the set $\Omega_3 = \{-,+\}$. Hence, the sets Ω_t, $t \in T = \{1,2,3\}$, occurring in Definition 4.6 are

$$\Omega_1 = \{Joe, Ann\}, \qquad \Omega_2 = \{a,b,c\}, \qquad \Omega_3 = \{-,+\},$$

and $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$ contains the twelve elements listed in first column of Table 4.2.

Furthermore, we choose the σ-algebras on these sets to be $\mathscr{A}_t = \mathscr{P}(\Omega_t)$, $t \in T$, the corresponding power sets. Because, in Definition 4.6, we require that $(\Omega,\mathscr{A})$ is the product of the measurable spaces $(\Omega_t,\mathscr{A}_t)$ (see RS-Def. 1.15), the measurable space $(\Omega,\mathscr{A})$ is completely specified, and with it the projections π_1,π_2, and π_3. The assignment of values of these projections to all elements of Ω is explicitly shown in Table 4.2.

Table 4.2. Joe and Ann with three treatment conditions

Possible outcomes			Projections			Observables					Restrictions	
Unit	Treatment condition	Success	π_1	π_2	π_3	Treatment variable X	Treatment variable 1_A	Treatment variable 1_B	Treatment variable 1_C	Outcome variable Y	Treatment variable $X\|_{\Omega_{ab}}$	Outcome variable $Y\|_{\Omega_{ab}}$
$\omega_1 = (Joe,\ a,\ -)$			*Joe*	*a*	−	0	1	0	0	0	1	0
$\omega_2 = (Joe,\ a,\ +)$			*Joe*	*a*	+	0	1	0	0	1	1	1
$\omega_3 = (Joe,\ b,\ -)$			*Joe*	*b*	−	1	0	1	0	0	0	0
$\omega_4 = (Joe,\ b,\ +)$			*Joe*	*b*	+	1	0	1	0	1	0	1
$\omega_5 = (Joe,\ c,\ -)$			*Joe*	*c*	−	2	0	0	1	0		
$\omega_6 = (Joe,\ c,\ +)$			*Joe*	*c*	+	2	0	0	1	1		
$\omega_7 = (Ann,\ a,\ -)$			*Ann*	*a*	−	0	1	0	0	0	1	0
$\omega_8 = (Ann,\ a,\ +)$			*Ann*	*a*	+	0	1	0	0	1	1	1
$\omega_9 = (Ann,\ b,\ -)$			*Ann*	*b*	−	1	0	1	0	0	0	0
$\omega_{10} = (Ann,\ b,\ +)$			*Ann*	*b*	+	1	0	1	0	1	0	1
$\omega_{11} = (Ann,\ c,\ -)$			*Ann*	*c*	−	2	0	0	1	0		
$\omega_{12} = (Ann,\ c,\ +)$			*Ann*	*c*	+	2	0	0	1	1		

Note. For the definition of Ω_{ab} see the text.

The filtration $(\mathscr{F}_t)_{t \in T}$, $T = \{1,2,3\}$, is specified by

$$\mathscr{F}_1 = \sigma(\pi_1), \quad \mathscr{F}_2 = \sigma(\pi_1, \pi_2), \quad \mathscr{F}_3 = \sigma(\pi_1, \pi_2, \pi_3). \tag{4.13}$$

Furthermore, according to Definition 4.6, in this kind of experiment,

$$\begin{aligned} \mathscr{C} = \sigma(\pi_2) = \big\{ & \{\omega_1, \omega_2, \omega_7, \omega_8\}, \{\omega_3, \omega_4, \omega_9, \omega_{10}\}, \{\omega_5, \omega_6, \omega_{11}, \omega_{12}\}, \\ & \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_7, \omega_8, \omega_9, \omega_{10}\}, \\ & \{\omega_1, \omega_2, \omega_5, \omega_6, \omega_7, \omega_8, \omega_{11}, \omega_{12}\}, \\ & \{\omega_3, \omega_4, \omega_5, \omega_6, \omega_9, \omega_{10}, \omega_{11}, \omega_{12}\}, \Omega, \emptyset \big\} \end{aligned} \tag{4.14}$$

is necessarily the cause σ-algebra because $J = K = \{1\}$, implying $\pi_2 = \pi_{21} = (\pi_{2j}, j \in K)$ (see Def. 4.6 and Exercise 4-7). Finally,

$$\mathscr{D}_{\mathscr{C}} = \sigma(\pi_1, \pi_{2j}, j \in J \setminus K) = \sigma(\pi_1) = \big\{\{\omega_1, \ldots, \omega_6\}, \{\omega_7, \ldots, \omega_{12}\}, \Omega, \emptyset\big\}$$

is the confounder σ-algebra of $\mathscr{C}$ because $J = K$. Hence, we completely specified the regular causality space $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$.

In the framework of this space, we can consider several putative cause variables. The first one is X as specified in Table 4.2. The σ-algebra generated by X is identical to $\mathscr{C} = \sigma(\pi_2)$. (Comparing X to π_2 in the table shows why.) Other putative cause variables are the *indicator variable of treatment a* defined by

$$1_A(\omega) \;=\; \begin{cases} 1, \text{ if } \omega \in \{\omega_1, \omega_2, \omega_7, \omega_8\} \\ 0, \text{ if } \omega \in \Omega \setminus \{\omega_1, \omega_2, \omega_7, \omega_8\}, \end{cases} \tag{4.15}$$

the *indicator variable of treatment b* defined by

$$1_B(\omega) \;=\; \begin{cases} 1, \text{ if } \omega \in \{\omega_3, \omega_4, \omega_9, \omega_{10}\} \\ 0, \text{ if } \omega \in \Omega \setminus \{\omega_3, \omega_4, \omega_9, \omega_{10}\}, \end{cases} \tag{4.16}$$

and the *indicator variable of treatment c* defined by

$$1_C(\omega) \;=\; \begin{cases} 1, \text{ if } \omega \in \{\omega_5, \omega_6, \omega_{11}, \omega_{12}\} \\ 0, \text{ if } \omega \in \Omega \setminus \{\omega_5, \omega_6, \omega_{11}, \omega_{12}\}. \end{cases} \tag{4.17}$$

The σ-algebras generated by these three putative cause variables are subsets of $\mathscr{C}$. They can be used for comparing one treatment condition *against the other two.* All four putative cause variables, X and the three indicator variables listed above, are maps on Ω, sharing the same cause and confounder σ-algebras, and the same set of potential confounders. Hence, we specified the four regular causality setups

$$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big), \qquad \big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},1_A,Y\big),$$
$$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},1_B,Y\big), \qquad \big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},1_C,Y\big),$$

which only differ in their putative cause variables (see Exercise 4-8).

Comparing Treatment *a* to *b*

However, if we intend to compare treatment conditions a and b to each other, then we have to use the restriction of the regular causality space $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$ to the subset

$$\Omega_{ab} := \{\omega_1,\omega_2,\omega_3,\omega_4,\omega_7,\omega_8,\omega_9,\omega_{10}\}$$

of Ω. That is, we use the regular causality space $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)|_{\Omega_{ab}}$. All σ-algebras in this space are the restrictions of the corresponding σ-algebras in $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$ to the set $\Omega_{ab} \subset \Omega$ (see Def. 4.24). For example, for the cause σ-algebra $\mathscr{C}$ specified in Equation (4.14), this means

$$\begin{aligned} \mathscr{C}|_{\Omega_{ab}} &= \{\Omega_{ab}\cap C\colon A\in\mathscr{C}\} \\ &= \big\{\Omega_{ab}\cap\{\omega_1,\omega_2,\omega_7,\omega_8\},\Omega_{ab}\cap\{\omega_3,\omega_4,\omega_9,\omega_{10}\},\Omega_{ab}\cap\{\omega_5,\omega_6,\omega_{11},\omega_{12}\}, \\ &\quad\ \Omega_{ab}\cap\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_7,\omega_8,\omega_9,\omega_{10}\}, \\ &\quad\ \Omega_{ab}\cap\{\omega_1,\omega_2,\omega_5,\omega_6,\omega_7,\omega_8,\omega_{11},\omega_{12}\}, \\ &\quad\ \Omega_{ab}\cap\{\omega_3,\omega_4,\omega_5,\omega_6,\omega_9,\omega_{10},\omega_{11},\omega_{12}\},\ \Omega_{ab}\cap\Omega,\Omega_{ab}\cap\emptyset\big\} \\ &= \big\{\{\omega_1,\omega_2,\omega_7,\omega_8\},\{\omega_3,\omega_4,\omega_9,\omega_{10}\},\ \Omega_{ab},\emptyset\big\}. \end{aligned} \tag{4.18}$$

Note that $\mathscr{C}|_{\Omega_{ab}}$ contains only four elements (see the last equation), whereas the original σ-algebra $\mathscr{C}$ contains eight elements. Checking conditions (a) to (c) of RS-Definition 1.4 for the set in the last displayed line shows that $\mathscr{C}|_{\Omega_{ab}}$ is in fact a σ-algebra on Ω_{ab} (see Exercise 4-9).

Finally, we may consider the restriction $X|_{\Omega_{ab}}$ of the original putative cause variable X to Ω_{ab} specified by

$$X|_{\Omega_{ab}}(\omega) \;=\; X(\omega) \;=\; \begin{cases} 0, \text{ if } \omega \in \{\omega_1, \omega_2, \omega_7, \omega_8\} \\ 1, \text{ if } \omega \in \{\omega_3, \omega_4, \omega_9, \omega_{10}\}, \end{cases}$$

and the restriction $Y|_{\Omega_{ab}}$ of the original outcome variable Y to Ω_{ab} specified by

$$Y|_{\Omega_{ab}}(\omega) \;=\; Y(\omega) \;=\; \begin{cases} 0, \text{ if } \omega \in \{\omega_1, \omega_3, \omega_7, \omega_9\} \\ 1, \text{ if } \omega \in \{\omega_2, \omega_4, \omega_8, \omega_{10}\} \end{cases}$$

(see the last two colums of Table 4.2). Hence, $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)|_{\Omega_{ab}}$ is a regular causality setup in which we can compare treatments a and b to each other with respect to the outcome variable $Y|_{\Omega_{ab}}$.

Comparing Two Other Treatment Conditions To Each Other

If we intend to compare two other treatment conditions to each other, for example, a to c or b to c, then each of these comparisons would be based on the restriction of the regular causality space $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$ to another subset of Ω. For comparing a to c, this subset of Ω is

$$\Omega_{ac} := \{\omega_1, \omega_2, \omega_5, \omega_6, \omega_7, \omega_8, \omega_{11}, \omega_{12}\},$$

and for comparing b to c, it is

$$\Omega_{bc} := \{\omega_3, \omega_4, \omega_5, \omega_6, \omega_9, \omega_{10}, \omega_{11}, \omega_{12}\}.$$

The corresponding restrictions of the regular causality space $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$ are

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)|_{\Omega_{ac}} \qquad \text{and} \qquad \big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)|_{\Omega_{bc}}.$$

4.2.2 Joe and Ann With Two Simultaneous Treatment Variables

Now we illustrate the concepts introduced in section 4.1 by the kind of experiment presented in Table 4.3. Again, we sample a person u from the set $\Omega_1 = \{Joe, Ann\}$. Then the sampled person receives or does not receive treatment a (e.g., drug a) and, simultaneously, he or she receives or does not receive treatment b (e.g., drug b). The pairs of these treatment conditions are the elements of the set

$$\begin{aligned} \Omega_2 \;&=\; \Omega_{21} \times \Omega_{22} \;=\; \{no, yes\} \times \{no, yes\} \\ &=\; \{(no, no),\ (no, yes),\ (yes, no),\ (yes, yes)\}. \end{aligned} \tag{4.19}$$

Finally, it is observed whether or not a success criterion is reached at some appropriate time *after* treatment. The set $\Omega_3 = \{-, +\}$ contains these two possible outcomes as elements.

Furthermore, we choose the σ-algebras on these sets to be the power sets $\mathscr{A}_t = \mathscr{P}(\Omega_t)$, $t \in T = \{1, 2, 3\}$, and specify $(\Omega, \mathscr{A})$ to be the product of the measurable spaces $(\Omega_t, \mathscr{A}_t)$ (see Def. 4.6 and RS-Def. 1.15). Hence, the measurable space

Table 4.3. Joe and Ann with two simultaneous treatment variables

Possible outcomes					Projections			Observables				Restrictions	
	Unit	Treatment a	Treatment b	Success	π_1	$\pi_2 = (\pi_{21}, \pi_{22})$	π_3	Treatment variable Z	Treatment variable X	Treatment variable 1_A	Outcome variable Y	Treatment variable $X\|_{\Omega_0}$	Outcome variable $Y\|_{\Omega_0}$
$\omega_1 =$	(*Joe*,	*no*,	*no*,	−)	*Joe*	(*no*, *no*)	−	0	0	0	0	0	0
$\omega_2 =$	(*Joe*,	*no*,	*no*,	+)	*Joe*	(*no*, *no*)	+	0	0	0	1	0	1
$\omega_3 =$	(*Joe*,	*no*,	*yes*,	−)	*Joe*	(*no*, *yes*)	−	0	1	0	0		
$\omega_4 =$	(*Joe*,	*no*,	*yes*,	+)	*Joe*	(*no*, *yes*)	+	0	1	0	1		
$\omega_5 =$	(*Joe*,	*yes*,	*no*,	−)	*Joe*	(*yes*, *no*)	−	1	0	0	0		
$\omega_6 =$	(*Joe*,	*yes*,	*no*,	+)	*Joe*	(*yes*, *no*)	+	1	0	0	1		
$\omega_7 =$	(*Joe*,	*yes*,	*yes*,	−)	*Joe*	(*yes*, *yes*)	−	1	1	1	0	1	0
$\omega_8 =$	(*Joe*,	*yes*,	*yes*,	+)	*Joe*	(*yes*, *yes*)	+	1	1	1	1	1	1
$\omega_9 =$	(*Ann*,	*no*,	*no*,	−)	*Ann*	(*no*, *no*)	−	0	0	0	0	0	0
$\omega_{10} =$	(*Ann*,	*no*,	*no*,	+)	*Ann*	(*no*, *no*)	+	0	0	0	1	0	1
$\omega_{11} =$	(*Ann*,	*no*,	*yes*,	−)	*Ann*	(*no*, *yes*)	−	0	1	0	0		
$\omega_{12} =$	(*Ann*,	*no*,	*yes*,	+)	*Ann*	(*no*, *yes*)	+	0	1	0	1		
$\omega_{13} =$	(*Ann*,	*yes*,	*no*,	−)	*Ann*	(*yes*, *no*)	−	1	0	0	0		
$\omega_{14} =$	(*Ann*,	*yes*,	*no*,	+)	*Ann*	(*yes*, *no*)	+	1	0	0	1		
$\omega_{15} =$	(*Ann*,	*yes*,	*yes*,	−)	*Ann*	(*yes*, *yes*)	−	1	1	1	0	1	0
$\omega_{16} =$	(*Ann*,	*yes*,	*yes*,	+)	*Ann*	(*yes*, *yes*)	+	1	1	1	1	1	1

Note. For the definition of Ω_0 see the text.

$$(\Omega, \mathscr{A}) = (\Omega_1 \times \Omega_2 \times \Omega_3, \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3)$$

is completely specified, and with it the projections π_1, π_2, π_3, and the σ-algebras of the filtration $(\mathscr{F}_t)_{t \in T}$ [see Eqs. (4.1)].

Note that the first projection π_1 is identical to the *person variable* U that has been used in the examples of chapter 1, for instance. In this example, the second projection $\pi_2 = (\pi_{21}, \pi_{22})$ consists of two projections $\pi_{21}: \Omega \to \Omega_{21}$ and $\pi_{22}: \Omega \to \Omega_{22}$ [see Eq. (4.19)]. Note that π_2 generates the same σ-algebra as (Z, X).

Table 4.3 shows the sixteen elements ω_i of the set $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$ of all possible outcomes of this kind of experiment. It also shows the values assigned to these elements by the projections π_1 to π_3, the treatment variable Z (that represents treatment a vs. $\neg a$), the treatment variable X (representing treatment b vs. $\neg b$), and by the outcome variable Y.

First Regular Causality Setup. Again, in this kind of experiment, we can choose among several maps as the focused putative cause variable. However, in contrast to the example presented in section 4.2.1, now the cause and confounder σ-algebras change, even if we stick to the unrestricted measurable space $(\Omega, \mathscr{A})$.

Referring to Definition 4.6, the index set J is $\{1,2\}$ because $\pi_2 = (\pi_{21}, \pi_{22})$. If we want to choose X to take the role of the putative cause variable, then we have to specify $K = \{2\}$, which yields

$$\mathscr{C}_1 \;=\; \sigma(\pi_{2j}, j \in K) \;=\; \sigma(\pi_{22}). \tag{4.20}$$

Another choice of the index set K would not meet conditions (a) and (b) of Definition 4.11 (i) for X. For $K = \{2\}$, condition (a) is satisfied because $\emptyset \neq \sigma(X) = \sigma(\pi_{22}) = \mathscr{C}_1$. Condition (b) is met as well because there is no proper subset K_0 of $K = \{2\}$ such that X is measurable with respect to $(\pi_{2j}, j \in K_0)$. In contrast, for $K = \{1\}$, condition (a) is not satisfied because $\sigma(X) \not\subset \sigma(\pi_{21})$. Finally, for $K = \{1,2\}$, condition (a) is satisfied, however, condition (b) does not hold because $K_0 = \{2\}$ is a proper subset of K and the σ-algebra generated by X is identical to (and therefore a subset of) $\sigma(\pi_{2j}, j \in K_0) = \sigma(\pi_{22})$.

Choosing $K = \{2\}$ does not only imply $\mathscr{C}_1 = \sigma(\pi_{22})$ but also that the corresponding confounder σ-algebra is

$$\mathscr{D}_{\mathscr{C}_1} \;=\; \sigma(\pi_1, \pi_{2j}, j \in J \setminus K) \;=\; \sigma(\pi_1, \pi_{21}). \tag{4.21}$$

In Table 4.3 we already specified the putative cause variable, the treatment variable X, and the outcome variable Y. In this example, $\sigma(X) = \sigma(\pi_{22}) = \mathscr{C}_1$, $\sigma(Z) = \sigma(\pi_{21}) \subset \mathscr{D}_{\mathscr{C}_1}$, and $\sigma(Y) = \sigma(\pi_3)$. Hence, the treatment variable Z, which is simultaneous to X, as well as π_1 and π_{21} are potential confounders of X because they all are $\mathscr{D}_{\mathscr{C}_1}$-measurable.

Finally, the bivariate projection (π_1, π_{21}) as well as (π_1, Z) are global potential confounders of X and can take the role of D_X [see Def. 4.11 (iii)] because $\sigma(\pi_1, \pi_{21}) = \sigma(\pi_1, Z) = \mathscr{D}_{\mathscr{C}_1}$. Hence, we completely specified the regular causality setup

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}_1, \mathscr{D}_{\mathscr{C}_1}, X, Y\big).$$

Second Regular Causality Setup. Because X and Z are simultaneous treatment variables, their roles can be exchanged. Hence, now we consider Z as the focused putative cause variable and X as a potential confounder of Z (see again Table 4.3). For this purpose, we specify $K = \{1\}$, which yields $\mathscr{C}_2 = \sigma(\pi_{21})$ and that $\mathscr{D}_{\mathscr{C}_2} = \sigma(\pi_1, \pi_{22})$ is the confounder σ-algebra of Z (see again Def. 4.6). Now, the bivariate projection (π_1, π_{22}) as well as (π_1, X) are global potential confounders of Z and can take the role of D_Z [see Def. 4.11 (iii)]. Hence, the regular causality setup is now

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}_2, \mathscr{D}_{\mathscr{C}_2}, Z, Y\big),$$

and (π_1, π_{22}), π_1, as well as X are potential confounders of Z. The outcome variable is still Y.

Third Regular Causality Setup. In this example, we can also choose the bivariate map (X, Z) to take the role of a putative cause variable. This allows us to study the joint effects of treatment a and b. In this case, $K = \{1,2\}$, $\mathscr{C}_3 = \sigma(\pi_{21}, \pi_{22})$ is the cause σ-algebra, and $\mathscr{D}_{\mathscr{C}_3} = \sigma(\pi_1)$ the confounder σ-algebra (see again Def. 4.6). Now, π_1 is a global potential confounder of the putative cause variable (X, Z). Hence, a third regular causality setup in the example presented in Table 4.3 is

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}_3, \mathscr{D}_{\mathscr{C}_3}, (X, Z), Y\big),$$

with the outcome variable Y specified in Table 4.3.

Fourth Regular Causality Setup. In the regular causality space $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C}_3,\mathscr{D}_{\mathscr{C}_3}\big)$ we can choose among several putative cause variables. The first one, the bivariate map (X,Z), has been selected above. However, we can also choose the indicator variable

$$1_A(\omega) = \begin{cases} 1, \text{ if } \omega \in \{\omega_7,\omega_8,\omega_{15},\omega_{16}\} \\ 0, \text{ otherwise.} \end{cases}$$

This new putative cause variable allows us to compare receiving both treatments to receiving at most one of the two treatments. The cause and the confounder σ-algebras are still $\mathscr{C}_3 = \sigma(\pi_2) = \sigma(\pi_{21},\pi_{22})$ and $\mathscr{D}_{\mathscr{C}_3} = \sigma(\pi_1)$, respectively, because 1_A is measurable with respect to π_2 but neither with respect to π_{21} nor with respect to π_{22}. Hence, we specified the fourth regular causality setup

$$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C}_3,\mathscr{D}_{\mathscr{C}_3},1_A,Y\big).$$

Fifth Regular Causality Setup. If, for example, we intend to compare receiving *both treatments* to receiving *no treatment*, then we have to use the restriction of the regular causality space $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C}_3,\mathscr{D}_{\mathscr{C}_3}\big)$ to the subset

$$\Omega_0 = \{\omega_1,\omega_2,\omega_7,\omega_8,\omega_9,\omega_{10},\omega_{15},\omega_{16}\}$$

of Ω. That is, we have to use $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C}_3,\mathscr{D}_{\mathscr{C}_3}\big)|_{\Omega_0}$ because the comparison of receiving *both treatments* to receiving *no treatment* only involves the elements of Ω_0 and not the other elements of Ω (see Table 4.3). All σ-algebras involved in this space are the restrictions of the corresponding σ-algebras to Ω_0. For example, for the cause σ-algebra

$$\begin{aligned}\mathscr{C}_3 &= \sigma(\pi_2) = \sigma(\pi_2)^{-1}(\mathscr{A}_2) = \sigma(\pi_2)^{-1}\big(\mathscr{P}(\Omega_2)\big)\\ &= \big\{\{\omega_1,\omega_2,\omega_9,\omega_{10}\},\{\omega_3,\omega_4,\omega_{11},\omega_{12}\},\{\omega_5,\omega_6,\omega_{13},\omega_{14}\},\{\omega_7,\omega_8,\omega_{15},\omega_{16}\},\\ &\quad \{\omega_1,\omega_2,\omega_3,\omega_4,\omega_9,\omega_{10},\omega_{11},\omega_{12}\},\{\omega_1,\omega_2,\omega_5,\omega_6,\omega_9,\omega_{10},\omega_{13},\omega_{14}\},\\ &\quad \{\omega_1,\omega_2,\omega_7,\omega_8,\omega_9,\omega_{10},\omega_{15},\omega_{16}\},\{\omega_3,\omega_4,\omega_5,\omega_6,\omega_{11},\omega_{12},\omega_{13},\omega_{14}\},\\ &\quad \{\omega_3,\omega_4,\omega_7,\omega_8,\omega_{11},\omega_{12},\omega_{15},\omega_{16}\},\{\omega_5,\omega_6,\omega_7,\omega_8,\omega_{13},\omega_{14},\omega_{15},\omega_{16}\},\\ &\quad \{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6,\omega_9,\omega_{10},\omega_{11},\omega_{12},\omega_{13},\omega_{14}\},\\ &\quad \{\omega_1,\omega_2,\omega_3,\omega_4,\omega_7,\omega_8,\omega_9,\omega_{10},\omega_{11},\omega_{12},\omega_{15},\omega_{16}\},\\ &\quad \{\omega_3,\omega_4,\omega_5,\omega_6,\omega_7,\omega_8,\omega_{11},\omega_{12},\omega_{13},\omega_{14},\omega_{15},\omega_{16}\},\\ &\quad \{\omega_1,\omega_2,\omega_5,\omega_6,\omega_7,\omega_8,\omega_9,\omega_{10},\omega_{13},\omega_{14},\omega_{15},\omega_{16}\},\Omega,\emptyset\big\}\end{aligned}$$

this means

$$\mathscr{C}_3|_{\Omega_0} = \big\{\{\omega_1,\omega_2,\omega_9,\omega_{10}\},\{\omega_7,\omega_8,\omega_{15},\omega_{16}\},\Omega_0,\emptyset\big\}.$$

For, the intended comparison we can use the restriction $X|_{\Omega_0}$ of the original putative cause variable X (see Table 4.3) to Ω_0 specified by,

$$X|_{\Omega_0}(\omega) = X(\omega) = \begin{cases} 0, \text{ if } \omega \in \{\omega_1,\omega_2,\omega_9,\omega_{10}\} \\ 1, \text{ if } \omega \in \{\omega_7,\omega_8,\omega_{15},\omega_{16}\}, \end{cases}$$

and the restriction $Y|_{\Omega_0}$ of the original outcome variable Y to Ω_0 specified by,

$$Y|_{\Omega_0}(\omega) = Y(\omega) = \begin{cases} 0, \text{ if } \omega \in \{\omega_1, \omega_7, \omega_9, \omega_{15}\} \\ 1, \text{ if } \omega \in \{\omega_2, \omega_8, \omega_{10}, \omega_{16}\}. \end{cases}$$

Hence, the fifth regular causality setup is

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}_3, \mathscr{D}_{\mathscr{C}_3}, X, Y\big)|_{\Omega_0} = \Big(\big(\Omega_0, \mathscr{A}|_{\Omega_0}\big), \big(\mathscr{F}_t|_{\Omega_0}\big)_{t\in T}, \mathscr{C}_3|_{\Omega_0}, \mathscr{D}_{\mathscr{C}_3}|_{\Omega_0}, X|_{\Omega_0}, Y|_{\Omega_0}\Big).$$

Of course, we might also want to compare receiving only treatment a to only receiving treatment b, or receiving both treatments to receiving only one single treatment, and so on. All such comparisons would need their own restriction of the regular causality space $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}_3, \mathscr{D}_{\mathscr{C}_3}\big)$ to the appropriate subset of Ω (see Exercise 4-10).

4.2.3 Nonorthogonal Factors

Consider again the example presented in Table 1.5. In contrast to the previous examples, now we consider an outcome variable Y whose possible values are not only 0 and 1 any more. Therefore, this example cannot be presented in the same format as before.

Nevertheless, we can specify the sets Ω_t, $t \in T = \{1,2,3\}$, occurring in Definition 4.6. In this example, they are $\Omega_1 = \{Tom, Tim, \ldots, Mia\}$, $\Omega_2 = \{control, treatment\,1, treatment\,2\}$, and $\Omega_3 = \mathbb{R}$. Furthermore, we choose the σ-algebras on these sets to be $\mathscr{A}_t = \mathscr{P}(\Omega_t)$, $t = 1,2$, and $\mathscr{A}_3 = \mathscr{B}$, the Borel σ-algebra on $\mathbb{R}$ (see RS-Rem. 1.14).

Requiring $(\Omega, \mathscr{A})$ to be the product of the measurable spaces $(\Omega_t, \mathscr{A}_t)$, $t \in T$ (see Def. 4.6), the measurable space $(\Omega, \mathscr{A})$ is completely specified and with it the projections π_1 to π_3. The projection $\pi_1 \colon \Omega \to \Omega_1$ has the possible values *Tom*, *Tim*, ..., *Mia*. Hence, π_1 is identical to the person variable U used in Table 1.5. The projection $\pi_2 \colon \Omega \to \Omega_2$ has the possible values *control*, *treatment* 1, *treatment* 2, and the possible values of $\pi_3 \colon \Omega \to \Omega_3$ are the real numbers.

The filtration $(\mathscr{F}_t)_{t\in T}$ is specified by

$$\mathscr{F}_1 = \sigma(\pi_1), \qquad \mathscr{F}_2 = \sigma(\pi_1, \pi_2), \qquad \mathscr{F}_3 = \sigma(\pi_1, \pi_2, \pi_3).$$

Furthermore, the index sets occurring in Definition 4.6 are $J = K = \{1\}$, implying that the cause σ-algebra is

$$\mathscr{C} = \sigma(\pi_{2j}, j \in K) = \sigma(\pi_2)$$

and that the confounder σ-algebra of $\mathscr{C}$ is

$$\mathscr{D}_{\mathscr{C}} = \sigma(\pi_1, \pi_{2j}, j \in J \setminus K) = \sigma(\pi_1).$$

Choosing X as the putative cause variable and Y as the outcome variable (see Table 1.5), we specify the regular causality setup

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big).$$

Note that the putative cause variable X is the composition of the projection π_2 and a measurable map $g \colon \big(\Omega_2, \mathscr{P}(\Omega_2)\big) \to (\mathbb{R}, \mathscr{B})$ (see RS-sect. 2.1.7), that is, $X = g(\pi_2)$. This implies $\sigma(X) \subset \sigma(\pi_2)$ (see RS-Lemma 2.35). Although X is a measurable map on $(\Omega, \mathscr{A})$, intuitively speaking, $X = g(\pi_2)$ means that the values of X only depend on the elements of the set $\Omega_2 = \{control, treatment\,1, treatment\,2\}$, the second factor set in the set product $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$.

Similarly, the outcome variable Y is the composition of the projection π_4 and a measurable map $f: (\mathbb{R}, \mathscr{B}) \to (\mathbb{R}, \mathscr{B})$, that is, $Y = f(\pi_4)$. This implies $\sigma(Y) \subset \sigma(\pi_4)$ (see again RS-Lemma 2.35). Hence, the values of Y only depend on the elements of $\Omega_4 = \mathbb{R}$, the fourth set in the set product $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$.

According to Definition 4.11 (iv), each non-constant measurable map W that satisfies $\sigma(W) \subset \mathscr{D}_{\mathscr{C}}$ is a potential confounder of X. In this example, this applies, for instance, to *educational status* specified in Table 1.5, but also to *sex*, which is also an attribute of the persons in the set $\Omega_1 = \{Tom, Tim, \dots, Mia\}$. Again, each of these variables can be written as the composition of some map and the projection π_1, implying that they are measurable with respect to $\sigma(\pi_1) = \mathscr{D}_{\mathscr{C}}$ (see Exercise 4-11).

4.3 Properties

In this section, we study some properties of a regular causality space, a putative cause variable, an outcome variable, a potential confounder, and other maps that are measurable with respect to these maps.

4.3.1 Cause σ-Algebra and Potential Confounder σ-Algebra

According to the following theorem, the intersection of the cause σ-algebra $\mathscr{C}$ and the confounder σ-algebra $\mathscr{D}_{\mathscr{C}}$ is the trivial σ-algebra, and the cause and confounder σ-algebras, $\mathscr{C}$ and $\mathscr{D}_{\mathscr{C}}$, are measurable with respect to $\mathscr{F}_2$. In contrast to $\mathscr{C}$, the confounder σ-algebra $\mathscr{D}_{\mathscr{C}}$ can also be measurable with respect to $\mathscr{F}_1$. According to Definition 4.6 (ii), this is the case if $J = K$, implying $\mathscr{D}_{\mathscr{C}} = \sigma(\pi_1)$. This is the case if there are no events $A \in \mathscr{A}$ that are simultaneous to the cause σ-algebra $\mathscr{C}$ and that are not elements of $\mathscr{C}$.

Theorem 4.25 [Properties of $\mathscr{C}$ and $\mathscr{D}_{\mathscr{C}}$]
If $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$ is a regular causality space, then

$$\mathscr{C} \cap \mathscr{D}_{\mathscr{C}} = \{\Omega, \emptyset\} \tag{4.22}$$
$$\mathscr{C} \not\subset \mathscr{F}_1 \tag{4.23}$$
$$\mathscr{C} \subset \mathscr{F}_2 \tag{4.24}$$
$$\mathscr{D}_{\mathscr{C}} \subset \mathscr{F}_2. \tag{4.25}$$

(Proof p. 109)

4.3.2 Putative Cause Variable

Proposition (4.22) has two immediate implications for the σ-algebra generated by a putative cause variable and the σ-algebra generated by a potential confounder.

Remark 4.26 [Two Immediate Implications] If $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ is a regular causality setup, then

$$\sigma(X) \cap \mathscr{D}_{\mathscr{C}} = \{\Omega, \emptyset\} \tag{4.26}$$

because $\sigma(X) \subset \mathscr{C}$. Furthermore, if W is a potential confounder of X, then

$$\sigma(X) \cap \sigma(W) = \{\Omega, \emptyset\}, \tag{4.27}$$

because $\sigma(W) \subset \mathscr{D}_{\mathscr{C}}$. ◁

Remark 4.27 [Implications Involving a Global Potential Confounder] Remember, by definition, a global potential confounder D_X of X satisfies $\sigma(D_X) = \mathscr{D}_{\mathscr{C}}$. Hence, Equation (4.26) is equivalent to

$$\sigma(X) \cap \sigma(D_X) = \{\Omega, \emptyset\}. \tag{4.28}$$

In the framework of a regular probabilistic causality setup $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$, this equation neither implies stochastic independence nor dependence of X and D_X. That is, X and D_X can be stochastically dependent or independent of each other, and this solely is determined by the probability measure P on $(\Omega, \mathscr{A})$. ◁

Example 4.28 [Joe and Ann With a Single Treatment Variable] The issue addressed in Remark 4.27 is exemplified by RS-Tables 1.2 and 1.4. These tables present random experiments that differ from each other only in the probability measure P, whereas the structure is the same as described in Example 4.10. Hence, in both examples, $\sigma(X) \cap \sigma(U) = \{\Omega, \emptyset\}$ holds for the putative cause variable X and the global potential confounder U. Because, in this example, $\sigma(X) = \mathscr{C}$ and $\sigma(U) = \mathscr{D}_{\mathscr{C}}$, this can be seen from Equations (4.3) and (4.4). Obviously, the only elements shared by these σ-algebras are Ω and $\emptyset$.

Furthermore, while X and U are stochastically *independent* in the example of RS-Table 1.2, they are stochastically *dependent* in the example of RS-Table 1.4. This can be seen from comparing the columns $P(X{=}1 \mid U{=}u)$ in the two tables to each other. In RS-Table 1.2, the conditional probability $P(X{=}1 \mid U{=}u)$ is the same for both persons u, which implies stochastic independence of X and U. In contrast, in RS-Table 1.4, the conditional probabilities $P(X{=}1 \mid U{=}\mathit{Joe})$ and $P(X{=}1 \mid U{=}\mathit{Ann})$ differ from each other, and this implies stochastic dependence of X and U. ◁

According to the following theorem, a putative cause variable is not a potential confounder of itself, it is measurable with respect to $\mathscr{F}_2$, and it is not measurable with respect to $\mathscr{F}_1$.

Theorem 4.29 [Properties of a Putative Cause Variable]
If $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *is a regular causality setup, then*

$$\sigma(X) \not\subset \mathscr{D}_{\mathscr{C}} \tag{4.29}$$

$$\sigma(X) \subset \mathscr{F}_2 \tag{4.30}$$

$$\sigma(X) \not\subset \mathscr{F}_1. \tag{4.31}$$

(Proof p. 110)

4.3.3 Potential Confounder

Now we study some properties of a potential confounder. In the following theorem, we show that every potential confounder is measurable with respect to the σ-algebra $\mathscr{F}_2$, but not measurable with respect to the cause σ-algebra $\sigma(X)$.

Theorem 4.30 [Two Properties of a Potential Confounder]
If $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$ *is a regular causality setup and* W *a potential confounder of* X*, then*

$$\sigma(W) \subset \mathscr{F}_2 \tag{4.32}$$

$$\sigma(W) \not\subset \sigma(X)\,. \tag{4.33}$$

(Proof p. 110)

Now we turn to time order concerning a putative cause variable and a potential confounder. According to the first proposition of the following theorem, a measurable map V that is prior in $(\mathscr{F}_t)_{t\in T}$ to a putative cause variable X is measurable with respect to $\mathscr{D}_{\mathscr{C}}$. And, according to the second proposition, if V is $\mathscr{D}_{\mathscr{C}}$-measurable, then it is prior or simultaneous in $(\mathscr{F}_t)_{t\in T}$ to X. Note that V can also be a global potential confounder of X.

Theorem 4.31 [A Measurable Map That Is Prior to X Is $\mathscr{D}_{\mathscr{C}}$-Measurable]
Let $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$ *be a regular causality setup,* V *a measurable map on* $(\Omega,\mathscr{A})$*, and let* $\mathscr{F}_T$ *denote the filtration* $(\mathscr{F}_t)_{t\in T}$*. Then*

$$V \underset{\mathscr{F}_T}{\lessdot} X \quad\Rightarrow\quad \sigma(V)\subset\mathscr{D}_{\mathscr{C}} \tag{4.34}$$

$$\sigma(V)\subset\mathscr{D}_{\mathscr{C}} \quad\Rightarrow\quad V \underset{\mathscr{F}_T}{\lesssim} X\,. \tag{4.35}$$

(Proof p. 110)

Proposition (4.34) and Definition 4.11 (iv) immediately imply the following corollary, in which we consider a *non-constant* measurable map on $(\Omega,\mathscr{A})$ that is prior in $(\mathscr{F}_t)_{t\in T}$ to X.

Corollary 4.32 [A Non-Constant Measurable Map That Is Prior to X]
Under the assumptions of Theorem 4.31, if V *is prior in* $\mathscr{F}_T$ *to* X *and* $\sigma(V) \neq \{\Omega,\emptyset\}$*, then* V *is a potential confounder of* X*.*

An immediate implication of Proposition (4.35) and Definition 4.11 applies to potential confounders, which, by definition are non-constant.

Corollary 4.33 [A Potential Confounder Is Prior or Simultaneous to X]
Under the assumptions of Theorem 4.31, if W *is a potential confounder of* X*, then* W *is prior or simultaneous in* $\mathscr{F}_T$ *to* X*.*

Hence, according to this corollary, each potential confounder of X is prior or simultaneous in $\mathscr{F}_T$ to X. Remember, however, a potential confounder of X is not measurable with respect to X [see Prop. (4.33)].

According to the following corollary, each non-constant W-measurable map is a potential confounder of X, if W is a potential confounder of X.

Corollary 4.34 [Measurable Maps of a Potential Confounder]
Let $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ be a regular causality setup and let V, W be measurable maps on $(\Omega, \mathscr{A})$. If W is a potential confounder of X and $\{\Omega, \emptyset\} \neq \sigma(V) \subset \sigma(W)$, then V is a potential confounder of X as well.

(Proof p. 111)

4.3.4 Outcome Variable

Now we study some measurability properties of an outcome variable Y. According to the following theorem, an outcome variable Y is $\mathscr{F}_3$-measurable but not measurable with respect to $\mathscr{F}_2$.

Theorem 4.35 [An Outcome Variable Is $\mathscr{F}_3$-Measurable]
If $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ is a regular causality setup, then $\sigma(Y) \not\subset \mathscr{F}_2$ and

$$\sigma(Y) \subset \mathscr{F}_3 . \tag{4.36}$$

(Proof p. 111)

Remark 4.36 [Immediate Implications] Note that $\sigma(Y) \not\subset \mathscr{F}_2$ implies $\sigma(Y) \not\subset \mathscr{F}_1$ because $(\mathscr{F}_t)_{t\in T}$ is a filtration. It also implies $\sigma(Y) \not\subset \mathscr{D}_{\mathscr{C}}$ because $\mathscr{D}_{\mathscr{C}} \subset \mathscr{F}_2$ [see Prop. (4.25)]. $\lhd$

According to the following corollary, a potential confounder of X is prior in $(\mathscr{F}_t)_{t\in T}$ to Y. According to Remark 4.16, this also applies to a covariate of X. This corollay immediately follows from Theorems 4.30, 4.35, and Definition 3.3.

Corollary 4.37 [A Potential Confounder of X Is Prior in $(\mathscr{F}_t)_{t\in T}$ to Y]
If $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ is a regular causality setup and W a potential confounder of X, then

$$W \underset{\mathscr{F}_T}{\prec} Y. \tag{4.37}$$

Now we consider a map V on $(\Omega, \mathscr{A})$ that is not $\mathscr{F}_2$-measurable. If it is measurable with respect to (W, Y), where W is $\mathscr{F}_2$-measurable and Y is $\mathscr{F}_3$-measurable, then, according to the following theorem, V is measurable with respect to $\mathscr{F}_3$.

Theorem 4.38 [Measurable Map of Y and W]
Let $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_\mathscr{C}, X, Y\big)$ be a regular causality setup, let W be $\mathscr{F}_2$-measurable, and V a measurable map on $(\Omega, \mathscr{A})$ that is not $\mathscr{F}_2$-measurable. Then

$$\sigma(V) \subset \sigma(W, Y) \quad \Rightarrow \quad \sigma(V) \subset \mathscr{F}_3. \tag{4.38}$$

(Proof p. 111)

Under the assumptions of Theorem 4.38, a (W, Y)-measurable map V on $(\Omega, \mathscr{A})$ that is not $\mathscr{F}_2$-measurable is posterior in $(\mathscr{F}_t)_{t \in T}$ to X.

Corollary 4.39 [X **Is Prior to** V]
Let the assumptions of Theorem 4.38 hold and let $\mathscr{F}_T$ denote the filtration $(\mathscr{F}_t)_{t \in T}$. Then

$$\sigma(V) \subset \sigma(W, Y) \quad \Rightarrow \quad X \underset{\mathscr{F}_T}{\lessdot} V. \tag{4.39}$$

(Proof p. 111)

According to RS-Lemma 2.34, the premise $\sigma(V) \subset \sigma(W, Y)$ holds for the composition of (W, Y) and a measurable map g. Hence, the propositions of Theorem 4.38 and Corollary 4.39 apply to such a composition $g(W, Y)$.

Corollary 4.40 [X **Is Prior to** $g(W, Y)$]
Let $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_\mathscr{C}, X, Y\big)$ be a regular causality setup, let $\mathscr{F}_T$ denote the filtration $(\mathscr{F}_t)_{t \in T}$, and let $W : (\Omega, \mathscr{A}) \to (\Omega'_W, \mathscr{A}'_W)$ be measurable with respect to $\mathscr{F}_2$. Furthermore, assume that $g : (\Omega'_Y \times \Omega'_W, \mathscr{A}'_Y \otimes \mathscr{A}'_W) \to (\Omega'_g, \mathscr{A}'_g)$ is a measurable map and $\sigma\big(g(W, Y)\big) \not\subset \mathscr{F}_2$. Then

$$\sigma\big(g(W, Y)\big) \subset \mathscr{F}_3 \tag{4.40}$$

and

$$X \underset{\mathscr{F}_T}{\lessdot} g(W, Y). \tag{4.41}$$

Hence, if the composition $g(W, Y)$ is not measurable with respect to the σ-algebra $\mathscr{F}_2$, then $g(W, Y)$ is $\mathscr{F}_3$-measurable and X is prior in $\mathscr{F}_T$ to $g(W, Y)$.

Example 4.41 [Linear Combination of W and Y] A special case of a map $g(W, Y)$ mentioned in Corollary 4.40 that is not $\mathscr{F}_2$-measurable is a linear combination $g(W, Y) = \alpha W + \beta Y$, $\alpha, \beta \in \mathbb{R}$, provided that $\beta \neq 0$. Note that in Definition 4.11 (ii) we assume $\sigma(Y) \not\subset \mathscr{F}_2$, which implies $\sigma(Y) \neq \{\Omega, \emptyset\}$. ◁

Example 4.42 [Change Score Variable] A special case of a linear combination of W and Y mentioned in Example 4.41 is the difference variable $g(W, Y) = Y - W$, where W assesses the same attribute as Y, only before treatment (see sect. 2.2). Again, remember the requirement $\sigma(Y) \not\subset \mathscr{F}_2$ [see Def. 4.11 (ii)], which implies $\sigma(Y) \neq \{\Omega, \emptyset\}$. ◁

Example 4.43 [Residual] Another special case of a function $g(W,Y)$ mentioned in Corollary 4.40 that is not $\mathscr{F}_2$-measurable is the residual $g(W,Y) = Y - E(Y|W)$ of Y with respect to its W-conditional expectation. [Here, we additionally assume that P is a probability measure on $(\Omega, \mathscr{A})$ (see RS-sect. 4.5).] Of course, the requirements for Y mentioned in the last examples still stand. ◁

The conjunction of Definition 4.11, Remark 4.9, and Corollary 4.40 implies the following corollary.

Corollary 4.44 [A New Regular Causality Setup]
Let the assumptions of Corollary 4.40 hold. Then

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, g(W,Y)\big) \tag{4.42}$$

is a regular causality setup.

According to this corollary, X is a putative cause variable of the outcome variable $g(W,Y)$, provided that the assumptions of Corollary 4.40 including $\sigma\big(g(W,Y)\big) \not\subset \mathscr{F}_2$ hold. Hence, according to Example 4.41,

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, \alpha W + \beta Y\big) \tag{4.43}$$

is a regular causality setup, provided that $\beta \neq 0$. Furthermore, according to Example 4.42,

$$\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y - W\big) \tag{4.44}$$

is a regular causality setup. Finally, according to Example 4.43,

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y - E(Y|W)\big) \tag{4.45}$$

is a regular probabilistic causality setup. Note that in all of these setups we assume that Y is not measurable with respect to $\mathscr{F}_2$, which implies $\sigma(Y) \neq \{\Omega, \emptyset\}$.

4.4 Summary and Conclusions

In this chapter, we specified the mathematical structure, and formulated the assumptions under which we can define causal effects. The most important concepts are gathered in Box 4.1 and their properties are summarized in Box 4.2.

The fundamental concept is a *regular causality space* $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$, which consists of a *measurable space* $(\Omega, \mathscr{A})$, a *filtration* $(\mathscr{F}_t)_{t\in T}$, a *cause* σ*-algebra* $\mathscr{C}$ and a *confounder* σ*-algebra* $\mathscr{D}_{\mathscr{C}}$. None of these concepts involves a probability measure. However, in the context of a probability measure P on $(\Omega, \mathscr{A})$, a *measurable space* $(\Omega, \mathscr{A})$ is the mathematical framework for any statement about *events*, their probabilities, and their (causal or noncausal) dependencies. It is also the framework for the definition of *random variables*, their distributions, and their (causal or noncausal) dependencies (see, e. g., Steyer & Nagel, 2017).

A *filtration* $(\mathscr{F}_t)_{t\in T}$ in $\mathscr{A}$ allows to define time order among elements of $\mathscr{A}$, among subsets of $\mathscr{A}$, and among measurable maps on $(\Omega, \mathscr{A})$ (see ch. 3). In the context of a probability

Box 4.1 Glossary of new concepts

$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$	*Regular causality space.* It consists of a measurable space, a filtration on $\mathscr{A}$, and two σ-algebras $\mathscr{C}$ and $\mathscr{D}_{\mathscr{C}}$ on $\mathscr{A}$.
$(\Omega,\mathscr{A})$	*Measurable space.* $(\Omega,\mathscr{A})$ is assumed to be the product of the measurable spaces $(\Omega_t,\mathscr{A}_t)$, $T = \{1,2,3\}$. For all $t \in T$ and all $\omega_t \in \Omega_t$, we assume $\{\omega_t\} \in \mathscr{A}_t$.
$(\mathscr{F}_t)_{t\in T}$	*Filtration* $(\mathscr{F}_t)_{t\in T}$ *in* $\mathscr{A}$. It consists of the σ-algebras $\mathscr{F}_1 = \sigma(\pi_1)$, $\mathscr{F}_2 = \sigma(\pi_1,\pi_2)$, and $\mathscr{F}_3 = \sigma(\pi_1,\pi_2,\pi_3)$, where π_t, $t \in T$, are the projections $\pi_t\colon \Omega \to \Omega_t$ defined by $\pi_t(\omega) = \omega_t$, for all $\omega \in \Omega$. The projection π_2 may itself consist of several projections, that is, $\pi_2 = (\pi_{2j}, j \in J)$, $J \subset \mathbb{N}$.
$\mathscr{C}$	*Cause* σ*-algebra*. For $\emptyset \neq K \subset J$ it is defined by $$\mathscr{C} = \sigma(\pi_{2j}, j \in K).$$
$\mathscr{D}_{\mathscr{C}}$	*Confounder* σ*-algebra of* $\mathscr{C}$. It is defined by $$\mathscr{D}_{\mathscr{C}} = \sigma(\pi_1,\pi_{2j}, j \in J\setminus K).$$
$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$	*Regular causality setup.* It consists of a regular causality space, a putative cause variable X, and an outcome variable Y, which are measurable maps on $(\Omega,\mathscr{A})$.
X	*Putative cause variable.* By definition, it satisfies (a) $\{\Omega,\emptyset\} \neq \sigma(X) \subset \mathscr{C}$ (b) $\neg\exists\, K_0 \subset K,\ K_0 \neq K,$ such that $\sigma(X) \subset \sigma(\pi_{2j}, j \in K_0)$.
Y	*Outcome variable* or *response variable.* It is defined by $$\sigma(Y) \not\subset \mathscr{F}_2.$$ In the framework of a regular causality space this implies that Y is measurable with respect to $\mathscr{F}_3$.
W	*Potential confounder of* X or *covariate of* X. It is a measurable map on $(\Omega,\mathscr{A})$ satisfying $\sigma(W) \subset \mathscr{D}_{\mathscr{C}}$.
D_X	*Global potential confounder of* X. A potential confounder of X satisfying $\sigma(D_X) = \mathscr{D}_{\mathscr{C}}$.
$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)\big|_{\Omega_0}$	*Restriction of a regular causality space* to $\Omega_0 \subset \Omega$. It is a new regular causality space in which all σ-algebras of $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)$ are restricted to Ω_0.
$\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)\big|_{\Omega_0}$	*Restriction of a regular causality setup* to $\Omega_0 \subset \Omega$. It is a regular causality setup in which we consider the restrictions of X and Y to Ω_0, and $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}}\big)\big|_{\Omega_0}$.

space $(\Omega, \mathscr{A}, P)$, this is tantamount to time order among events, among sets of events, and among random variables, respectively. Such a time order is indispensable for a formalization of the intuitive idea that a cause precedes its outcome. The filtration $(\mathscr{F}_t)_{t\in T}$ consists of three σ-algebras. Their substantive interpretation is that, from the perspective of the cause, $\mathscr{F}_1$ represents the *set of all past events*, $\mathscr{F}_2$ the *set of all past and present events*, and $\mathscr{F}_3$ the *set of all past, present, and future events*.

The third and fourth components of a regular causality space are the *cause* σ*-algebra* $\mathscr{C}$ and its *confounder* σ*-algebra* $\mathscr{D}_{\mathscr{C}}$. In the context of a probability space and from the perspective of a focused putative cause variable X, the σ-algebra $\mathscr{C}$ represents the set of all *present events* with respect to which a putative cause variable X is measurable. In contrast, $\mathscr{D}_{\mathscr{C}}$ consists of all *past and all other present events*. They are the events that might confound or disturb the stochastic dependency of the outcome variable Y on the putative cause variable X.

Adding a putative cause variable X and an outcome variable Y to a regular causality space constitutes a *regular causality setup* $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$. By definition, the *putative cause variable* X satisfies two conditions. According to condition (a), it is $\mathscr{C}$-measurable and its generated σ-algebra is not trivial. The latter would be the case if X would be a constant map. According to the condition (b), there is no family of projections $(\pi_{2j}, j \in K_0)$, $K_0 \subset K$, such that X is measurable with respect to this family of projections. This condition secures that the confounder σ-algebra $\mathscr{D}_{\mathscr{C}}$ is chosen such that all other putative cause variables that are simultaneous to X are measurable with respect to $\mathscr{D}_{\mathscr{C}}$.

By definition, the *outcome variable* Y is not measurable with respect to $\mathscr{F}_2$. Hence, an outcome variable Y can either solely depend on π_3 or depend on π_3 *and* one or more of the projections π_1 and π_2. This allows, for instance, to consider pre-post difference variables ('pre' and 'post' with respect to the focused putative cause variable) as outcome variables (see Examples 4.41 to 4.43).

Furthermore, we introduced the concepts of a potential confounder and a global potential confounder of a putative cause variable X. By definition, a *potential confounder* W of a putative cause variable X is a nonconstant $\mathscr{D}_{\mathscr{C}}$-measurable map, and a *global potential confounder* D_X of X is a potential confounder that generates the confounder σ-algebra $\mathscr{D}_{\mathscr{C}}$. Hence, all potential confounders are measurable with respect to D_X.

Last but not least, in Theorem 4.22 we showed that the *restriction of a regular causality space* $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)|_{\Omega_0}$ to $\Omega_0 \subset \Omega$ is a new regular causality space in which all σ-algebras involved in $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}\big)$ are restricted to Ω_0. Adding the restrictions of the maps X and Y to Ω_0 then yields $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)|_{\Omega_0}$, the *restriction of a regular causality setup* to Ω_0, which is the formal framework in which we can consider putative cause variables and outcome variables that are measurable maps on Ω_0 (see the Examples in sect. 4.2.1 and 4.2.2).

4.5 Proofs

Proof of Theorem 4.22

Remember, $(\Omega, \mathscr{A})$ is the product of the measurable spaces $(\Omega_t, \mathscr{A}_t)$, $t \in T$. Hence, first we show that $\big(\Omega_0, \mathscr{A}|_{\Omega_0}\big)$ is the product of the measurable spaces $\big(\Omega_{0t}, \mathscr{A}_t|_{\Omega_{0t}}\big)$, $t \in T$. That is, we show $\mathscr{A}|_{\Omega_0} = \mathscr{A}_1|_{\Omega_{01}} \otimes \mathscr{A}_2|_{\Omega_{02}} \otimes \mathscr{A}_3|_{\Omega_{03}}$. Now consider the set system

Box 4.2 Properties of a regular causality setup and potential confounders

All propositions gathered in this box refer to the same regular causality setup $\big((\Omega,\mathscr{A}),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$ and throughout this box, W denotes a potential confounder of X, which by definition, satisfies $\sigma(W)\subset\mathscr{D}_{\mathscr{C}}$.

$$\mathscr{C}\cap\mathscr{D}_{\mathscr{C}} = \{\Omega,\emptyset\} \tag{i}$$
$$\mathscr{C} \not\subset \mathscr{F}_1 \tag{ii}$$
$$\mathscr{C} \subset \mathscr{F}_2 \tag{iii}$$
$$\mathscr{D}_{\mathscr{C}} \subset \mathscr{F}_2 \tag{iv}$$

$$\sigma(X)\cap\sigma(W) = \{\Omega,\emptyset\} \tag{v}$$
$$\sigma(X) \subset \mathscr{F}_2 \tag{vi}$$
$$\sigma(X) \not\subset \mathscr{F}_1 \tag{vii}$$
$$\sigma(X) \not\subset \mathscr{D}_{\mathscr{C}} \tag{viii}$$

$$\sigma(Y) \subset \mathscr{F}_3 \tag{ix}$$
$$\forall\, t\in\{1,2\}\colon\ \sigma(Y) \not\subset \mathscr{F}_t \tag{x}$$
$$\sigma(Y) \not\subset \mathscr{D}_{\mathscr{C}} \tag{xi}$$

$$\sigma(W) \subset \mathscr{F}_2 \tag{xii}$$
$$\sigma(W) \not\subset \sigma(X) \tag{xiii}$$
$$W \underset{\mathscr{F}_T}{\precneqq} X. \tag{xiv}$$

If V is a measurable map on $(\Omega,\mathscr{A})$ that is not $\mathscr{F}_2$-measurable, then

$$\sigma(V)\subset\sigma(W,Y) \quad\Rightarrow\quad (a)\ \ \sigma(V)\subset\mathscr{F}_3 \tag{xv}$$
$$(b)\ \ X \underset{\mathscr{F}_T}{\prec} V \tag{xvi}$$
$$(c)\ \ V \text{ is an outcome variable.} \tag{xvii}$$

If Z is a measurable map on $(\Omega,\mathscr{A})$, then

$$\sigma(Z)\subset\sigma(W) \quad\Rightarrow\quad V \text{ is a potential confounder of } X. \tag{xviii}$$

$$\begin{aligned}
\mathscr{E}_0 &:= \Big\{\bigtimes_{t=1}^{3} A_{0t}\colon A_{0t}\in\mathscr{A}_t|_{\Omega_{0t}}\Big\} \\
&= \Big\{\bigtimes_{t=1}^{3} A_{0t}\colon A_{0t}\in\{\Omega_{0t}\cap A_t\colon A_t\in\mathscr{A}_t\}\Big\} && [(4.5)] \\
&= \Big\{\bigtimes_{t=1}^{3} (\Omega_{0t}\cap A_t)\colon A_t\in\mathscr{A}_t\Big\} && [A_{0t}=\Omega_{0t}\cap A_t] \\
&= \Big\{\bigtimes_{t=1}^{3} \Omega_{0t}\cap\bigtimes_{t=1}^{3} A_t\colon A_t\in\mathscr{A}_t\Big\} \\
&= \Big\{\Omega_0\cap\bigtimes_{t=1}^{3} A_t\colon A_t\in\mathscr{A}_t\Big\}. && \Big[\Omega_0=\bigtimes_{t=1}^{3}\Omega_{0t}\Big]
\end{aligned}$$

According to RS-Equation (1.10), the set system $\mathscr{E}_0$ is a generating system of the product σ-algebra $\mathscr{A}_1|_{\Omega_{01}} \otimes \mathscr{A}_2|_{\Omega_{02}} \otimes \mathscr{A}_3|_{\Omega_{03}}$. Hence, using this fact and the last equation yields

$$
\begin{aligned}
\bigotimes_{t=1}^{3} \mathscr{A}_t|_{\Omega_{0t}} &= \sigma(\mathscr{E}_0) && \text{[RS-Eq. (1.10)]}\\
&= \sigma\Big(\Big\{\Omega_0 \cap \mathop{\times}_{t=1}^{3} A_t \colon A_t \in \mathscr{A}_t\Big\}\Big) && \Big[\mathscr{E}_0 = \Big\{\Omega_0 \cap \mathop{\times}_{t=1}^{3} A_t \colon A_t \in \mathscr{A}_t\Big\}\Big]\\
&= \sigma\Big(\Big\{\Omega_0 \cap A \colon A \in \bigotimes_{t=1}^{3} \mathscr{A}_t\Big\}\Big) && \Big[\text{RS-(1.3)}, \Big\{\mathop{\times}_{t=1}^{3} A_t \colon A_t \in \mathscr{A}_t\Big\} \subset \bigotimes_{t=1}^{3} \mathscr{A}_t\Big]\\
&= \Big\{\Omega_0 \cap A \colon A \in \bigotimes_{t=1}^{3} \mathscr{A}_t\Big\} && \text{[RS-(1.6)]}\\
&= \big\{\Omega_0 \cap A \colon A \in \mathscr{A}\big\} && \text{[Def. 4.6, RS-Eq. (1.10)]}\\
&= \mathscr{A}|_{\Omega_0}. && \text{[(4.5)]}
\end{aligned}
$$

Hence, $\mathscr{A}|_{\Omega_0}$ is in fact the product of the σ-algebras $\mathscr{A}_t|_{\Omega_{0t}}$, $t \in T$, as required in the definition of a regular causality space.

Next we show that $(\mathscr{F}_t|_{\Omega_0})_{t \in T}$ is a filtration in $\mathscr{A}|_{\Omega_0}$.

$$
\begin{aligned}
& \mathscr{F}_1 \subset \mathscr{F}_2 \subset \mathscr{F}_3 \\
\Rightarrow\ & \{\Omega_0 \cap A \colon A \in \mathscr{F}_1\} \subset \{\Omega_0 \cap A \colon A \in \mathscr{F}_2\} \subset \{\Omega_0 \cap A \colon A \in \mathscr{F}_3\} \\
\Leftrightarrow\ & \mathscr{F}_1|_{\Omega_0} \subset \mathscr{F}_2|_{\Omega_0} \subset \mathscr{F}_3|_{\Omega_0}. && \text{[(4.5)]}
\end{aligned}
$$

The filtration $(\mathscr{F}_t|_{\Omega_0})_{t \in T}$ satisfies Equations (4.9), which can be seen as follows:

$$
\begin{aligned}
\mathscr{F}_1|_{\Omega_0} &= \{\Omega_0 \cap A \colon A \in \mathscr{F}_1\} && \text{[(4.5)]}\\
&= \big\{A_0 \colon A_0 \in \mathscr{F}_1|_{\Omega_0}\big\} && [A_0 := \Omega_0 \cap A]\\
&= \Big\{\big\{\omega \in \Omega_0 \colon \pi_1|_{\Omega_0}(\omega) \in A_0\big\} \colon A_0 \in \mathscr{F}_1|_{\Omega_0}\Big\} && \text{[(4.8)]}\\
&= \Big\{\big(\pi_1|_{\Omega_0}\big)^{-1}(A_1) \colon A_1 \in \mathscr{A}_1|_{\Omega_{01}}\Big\} && \text{[RS-(2.1)]}\\
&= \sigma\big(\pi_1|_{\Omega_0}\big). && \text{[RS-(2.11), (2.12)]}
\end{aligned}
$$

$$
\begin{aligned}
\mathscr{F}_2|_{\Omega_0} &= \{\Omega_0 \cap A \colon A \in \mathscr{F}_2\} && \text{[(4.5)]}\\
&= \big\{A_0 \colon A_0 \in \mathscr{F}_2|_{\Omega_0}\big\} && [A_0 := \Omega_0 \cap A]\\
&= \Big\{\big\{\omega \in \Omega_0 \colon \big(\pi_1|_{\Omega_0}, \pi_2|_{\Omega_0}\big)(\omega) \in A_0\big\} \colon A_0 \in \mathscr{F}_2|_{\Omega_0}\Big\} && \text{[(4.8)]}\\
&= \big\{\big(\pi_1|_{\Omega_0}, \pi_2|_{\Omega_0}\big)^{-1}(A_{12}) \colon A_{12} \in \mathscr{A}_1|_{\Omega_{01}} \otimes \mathscr{A}_2|_{\Omega_{02}}\big\} && \text{[RS-(2.1)]}\\
&= \sigma\big(\pi_1|_{\Omega_0}, \pi_2|_{\Omega_0}\big). && \text{[RS-(2.11), (2.12)]}
\end{aligned}
$$

$$
\begin{aligned}
\mathscr{F}_3|_{\Omega_0} &= \{\Omega_0 \cap A \colon A \in \mathscr{F}_3\} && \text{[(4.5)]}\\
&= \big\{A_0 \colon A_0 \in \mathscr{F}_3|_{\Omega_0}\big\} && [A_0 := \Omega_0 \cap A]\\
&= \big\{\big\{\omega \in \Omega_0 \colon \big(\pi_1|_{\Omega_0}, \pi_2|_{\Omega_0}, \pi_3|_{\Omega_0}\big)(\omega) \in A_0\big\} \colon A_0 \in \mathscr{F}_3|_{\Omega_0}\big\} && \text{[(4.8)]}\\
&= \Big\{\big(\pi_1|_{\Omega_0}, \pi_2|_{\Omega_0}, \pi_3|_{\Omega_0}\big)^{-1}(A_{123}) \colon A_{123} \in \bigotimes_{t=1}^{3} \mathscr{A}_t|_{\Omega_{01}}\Big\} && \text{[RS-(2.1)]}\\
&= \sigma\big(\pi_1|_{\Omega_0}, \pi_2|_{\Omega_0}, \pi_3|_{\Omega_0}\big). && \text{[RS-(2.11), (2.12)]}
\end{aligned}
$$

Now we show that $\mathscr{C}|_{\Omega_0}$ is a cause σ-algebra. Let $\emptyset \neq K \subset J$,

$$\begin{aligned}
\mathscr{C}|_{\Omega_0} &= \{\Omega_0 \cap A\colon A \in \mathscr{C}\} && [(4.5)]\\
&= \big\{\Omega_0 \cap A\colon A \in \sigma(\pi_{2j}, j \in K)\big\} && [\text{Def. } 4.6\ (i)]\\
&= \big\{A_0\colon A_0 \in \sigma\big(\pi_{2j}|_{\Omega_0}, j \in K\big)\big\} && [A_0 = \Omega_0 \cap A]\\
&= \sigma\big(\pi_{2j}|_{\Omega_0}, j \in K\big).
\end{aligned}$$

Finally, we show that $\mathscr{D}_{\mathscr{C}}|_{\Omega_0}$ is the confounder σ-algebra of $\mathscr{C}|_{\Omega_0}$.

$$\begin{aligned}
\mathscr{D}_{\mathscr{C}}|_{\Omega_0} &= \{\Omega_0 \cap A\colon A \in \mathscr{D}_{\mathscr{C}}\} && [(4.5)]\\
&= \big\{\Omega_0 \cap A\colon A \in \sigma(\pi_1, \pi_{2j}, j \in J \setminus K)\big\} && [\text{Def. } 4.6\ (ii)]\\
&= \big\{A_0\colon A_0 \in \sigma\big(\pi_1|_{\Omega_0}, \pi_{2j}|_{\Omega_0}, j \in J \setminus K\big)\big\} && [A_0 = \Omega_0 \cap A]\\
&= \sigma\big(\pi_1|_{\Omega_0}, \pi_{2j}|_{\Omega_0}, j \in J \setminus K\big).
\end{aligned}$$

Proof of Theorem 4.25

Equation (4.22). Note that $K \cap (\{1\} \cup J \setminus K) = \emptyset$ holds for the sets used in Definition 4.6. Hence,

$$\begin{aligned}
\mathscr{C} \cap \mathscr{D}_{\mathscr{C}} &= \sigma(\pi_{2j}, j \in K) \ \cap\ \sigma(\pi_1, \pi_{2j}, j \in J \setminus K) && [\text{Def. } 4.6\ (i), (ii)]\\
&= \{\Omega, \emptyset\}. && [\text{RS-Eq. } (2.39)]
\end{aligned}$$

Proposition (4.23). The proof is by contradiction.

$$\begin{aligned}
\mathscr{C} \subset \mathscr{F}_1 \quad &\Leftrightarrow \quad \mathscr{C} \subset \sigma(\pi_1) && [(4.1)]\\
&\Rightarrow \quad \mathscr{C} \subset \mathscr{D}_{\mathscr{C}}, && [(3.16), \text{Def. } 4.6\ (ii)]
\end{aligned}$$

which is a contradiction to Proposition (4.22) because $\mathscr{C} \neq \{\Omega, \emptyset\}$ [see Def. 4.6 (i)].

Proposition (4.24). We assume $\emptyset \neq K \subset J$ (see Def. 4.6). Hence,

$$\begin{aligned}
& \mathscr{C} = \sigma(\pi_{2j}, j \in K) && [\text{Def. } 4.6\ (i)]\\
\Rightarrow\ & \mathscr{C} \subset \sigma(\pi_{2j}, j \in J) && [(3.16)]\\
\Leftrightarrow\ & \mathscr{C} \subset \sigma(\pi_2) && [\text{Def. } 4.6]\\
\Rightarrow\ & \mathscr{C} \subset \mathscr{F}_2. && [(3.16), (4.1)]
\end{aligned}$$

Proposition (4.25).

$$\begin{aligned}
& \mathscr{D}_{\mathscr{C}} = \sigma(\pi_1, \pi_{2j}, j \in J \setminus K) && [\text{Def. } 4.6\ (ii)]\\
\Rightarrow\ & \mathscr{D}_{\mathscr{C}} \subset \sigma(\pi_1, \pi_{2j}, j \in J) && [(3.16)]\\
\Leftrightarrow\ & \mathscr{D}_{\mathscr{C}} \subset \sigma(\pi_1, \pi_2) && [\text{Def. } 4.6]\\
\Leftrightarrow\ & \mathscr{D}_{\mathscr{C}} \subset \mathscr{F}_2. && [(4.1)]
\end{aligned}$$

Proof of Theorem 4.29

Proposition (4.29). In Definition 4.11 (i), we assume $\sigma(X) \neq \{\Omega, \emptyset\}$. Hence, $\sigma(X) \subset \mathscr{D}_{\mathscr{C}}$ would be a contradiction to Equation (4.26).

Proposition (4.30).

$$
\begin{aligned}
& \sigma(X) \subset \mathscr{C} && \text{[Def. 4.11 (i)]} \\
\Rightarrow \quad & \sigma(X) \subset \sigma(\pi_1) \cup \sigma(\pi_2) && \text{[Def. 4.6 (i), (3.16)]} \\
\Rightarrow \quad & \sigma(X) \subset \sigma\big(\sigma(\pi_1) \cup \sigma(\pi_2)\big) && \text{[RS-(1.6), RS-(1.7)]} \\
\Leftrightarrow \quad & \sigma(X) \subset \sigma(\pi_1, \pi_2) && \text{[(3.20)]} \\
\Leftrightarrow \quad & \sigma(X) \subset \mathscr{F}_2\,. && \text{[(4.1)]}
\end{aligned}
$$

Proposition (4.31). The proof is by contradiction.

$$
\begin{aligned}
\sigma(X) \subset \mathscr{F}_1 \quad & \Leftrightarrow \quad \sigma(X) \subset \sigma(\pi_1) && \text{[(4.1)]} \\
& \Rightarrow \quad \sigma(X) \subset \mathscr{D}_{\mathscr{C}}, && \text{[(3.16), Def. 4.6 (ii)]}
\end{aligned}
$$

which is a contradiction to Proposition (4.29) because we assume $\sigma(X) \neq \{\Omega, \emptyset\}$ [see Def. 4.11 (i)].

Proof of Theorem 4.30

Proposition (4.32).

$$
\begin{aligned}
& \sigma(W) \subset \mathscr{D}_{\mathscr{C}} && \text{[Def. 4.11 (iv)]} \\
\Rightarrow \quad & \sigma(W) \subset \sigma(\pi_1, \pi_2) && \text{[Def. 4.6 (ii), (3.16)]} \\
\Leftrightarrow \quad & \sigma(W) \subset \mathscr{F}_2\,. && \text{[(4.1)]}
\end{aligned}
$$

Proposition (4.33). The proof is by contradiction. Hence, additionally to $\sigma(W) \subset \mathscr{D}_{\mathscr{C}}$ and $\sigma(W) \neq \{\Omega, \emptyset\}$ [see Def. 4.11 (iv)], we assume $\sigma(W) \subset \sigma(X)$. Then

$$
\begin{aligned}
\big(\sigma(W) \subset \mathscr{D}_{\mathscr{C}} \wedge \sigma(W) \subset \sigma(X)\big) \quad & \Leftrightarrow \quad \sigma(W) \subset \big(\mathscr{D}_{\mathscr{C}} \cap \sigma(X)\big) && \text{[(3.14)]} \\
& \Rightarrow \quad \sigma(W) \subset \{\Omega, \emptyset\}, && \text{[(4.26)]}
\end{aligned}
$$

which is a contradiction to the assumption $\sigma(W) \neq \{\Omega, \emptyset\}$.

Proof of Theorem 4.31

Proposition (4.34). According to Propositions (4.30) and (4.31), $\sigma(X) \subset \mathscr{F}_2$ and $\sigma(X) \not\subset \mathscr{F}_1$. Hence,

$$
\begin{aligned}
V \underset{\mathscr{F}_T}{\precsim} X \quad & \Rightarrow \quad \sigma(V) \subset \mathscr{F}_1 && \text{[(4.30), (4.31), Def. 3.3]} \\
& \Leftrightarrow \quad \sigma(V) \subset \sigma(\pi_1) && \text{[(4.1)]} \\
& \Rightarrow \quad \sigma(V) \subset \mathscr{D}_{\mathscr{C}}\,. && \text{[Def. 4.6 (ii), (3.16)]}
\end{aligned}
$$

Proposition (4.35).

$$
\begin{array}{lll}
& \sigma(V) \subset \mathscr{D}_{\mathscr{C}} & \\
\Rightarrow & \sigma(V) \subset \sigma(\pi_1, \pi_2) & \text{[Def. 4.6 (ii), (3.16)]} \\
\Leftrightarrow & \sigma(V) \subset \mathscr{F}_2 & \text{[(4.1)]} \\
\Rightarrow & \sigma(V) \subset \mathscr{F}_1 \vee \big(\sigma(V) \subset \mathscr{F}_2 \wedge \sigma(V) \not\subset \mathscr{F}_1\big) & [\mathscr{F}_1 \subset \mathscr{F}_2] \\
\Rightarrow & \sigma(V) \underset{\mathscr{F}_T}{\lesssim} \sigma(X) \vee \sigma(V) \underset{\mathscr{F}_T}{\approx} \sigma(X) & [\text{Def. } 3.3, \mathscr{F}_1 \not\supset \sigma(X) \subset \mathscr{F}_2, \text{Def. } 3.24] \\
\Leftrightarrow & V \underset{\mathscr{F}_T}{\precsim} X. & \text{[Def. 3.37 (ii)]}
\end{array}
$$

Proof of Corollary 4.34

$$
\begin{array}{lll}
& \{\Omega, \emptyset\} \neq \sigma(V) \subset \sigma(W) & \\
\Rightarrow & \{\Omega, \emptyset\} \neq \sigma(V) \subset \mathscr{D}_{\mathscr{C}} & [\sigma(W) \subset \mathscr{D}_{\mathscr{C}}, (3.16)] \\
\Leftrightarrow & V \text{ is a potential confounder of } X. & \text{[Def. 4.11 (iv)]}
\end{array}
$$

Proof of Theorem 4.35

The property $\sigma(Y) \not\subset \mathscr{F}_2$ is required in Definition 4.11 (ii). Furthermore,

$$
\begin{array}{lll}
& \sigma(Y) \subset \mathscr{A} & \text{[Def. 4.11, SN-Cor. 2.28]} \\
\Leftrightarrow & \sigma(Y) \subset \mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3 & \text{[Def. 4.6, RS-Def. 1.15]} \\
\Leftrightarrow & \sigma(Y) \subset \sigma(\pi_1, \pi_2, \pi_3) & \text{[RS-Th. 2.30]} \\
\Leftrightarrow & \sigma(Y) \subset \mathscr{F}_3. & \text{[(4.1)]}
\end{array}
$$

Proof of Theorem 4.38

$$
\begin{array}{lll}
& \sigma(W) \subset \mathscr{F}_2 \wedge \sigma(Y) \subset \mathscr{F}_3 & \text{[(4.36)]} \\
\Rightarrow & \sigma(W) \subset \mathscr{F}_3 \wedge \sigma(Y) \subset \mathscr{F}_3 & [\mathscr{F}_2 \subset \mathscr{F}_3, (3.16)] \\
\Rightarrow & \big(\sigma(Y) \cup \sigma(W)\big) \subset \mathscr{F}_3 & \text{[(3.15)]} \\
\Leftrightarrow & \sigma(W, Y) \subset \mathscr{F}_3 & \text{[(3.20), (3.21)]} \\
\Rightarrow & \sigma(V) \subset \mathscr{F}_3. & [\sigma(V) \subset \sigma(W, Y), (3.16)]
\end{array}
$$

Proof of Corollary 4.39

According to Proposition (4.30), $\sigma(X) \subset \mathscr{F}_2$ and according to Theorem 4.38, $\sigma(V) \subset \mathscr{F}_3$. However, this implies $X \underset{\mathscr{F}_T}{\lesssim} V$ (see Def. 3.3).

4.6 Exercises

▷ **Exercise 4-1** Write down the set of all values of the bivariate projection (π_1, π_2) occurring in Equation (4.2).

▷ **Exercise 4-2** Consider Table 4.1 and gather in a set all elements of the σ-algebra $\sigma(\pi_1)$, using RS-Equations (2.1), (2.11), and (2.12).

▷ **Exercise 4-3** Consider Table 4.1 and write down the product σ-algebra $\mathscr{A}_1 \otimes \mathscr{A}_2$ with all its elements. The σ-algebras $\mathscr{A}_1$ and $\mathscr{A}_2$ are specified in Example 4.10 to be the power sets of Ω_1 and Ω_2, respectively. Use RS-Equation (1.10) as well as RS-Definitions 1.4 and 1.7.

▷ **Exercise 4-4** Consider Table 4.1 and gather in two sets all elements of the two inverse images $(\pi_1, \pi_2)^{-1}(\{(Joe, no)\})$ and $(\pi_1, \pi_2)^{-1}(\{(Joe, no), (Ann, no)\})$, using RS-Equation (2.1).

▷ **Exercise 4-5** In the context of a probability space $(\Omega, \mathscr{A}, P)$, what is the substantive interpretation of the sets $\{\omega_1, \omega_2, \omega_5, \omega_6\}$ and $\{\omega_3, \omega_4, \omega_7, \omega_8\}$ occurring in the set $\mathscr{C}$ specified in Equation 4.3?

▷ **Exercise 4-6** Enumerate all elements of the σ-algebra generated by the map $X: \Omega \to \mathbb{R}$ specified in Table 4.1, using $\sigma(X) = X^{-1}(\mathscr{B})$ [see RS-Eqs. (2.11) and (2.12)], where $\mathscr{B}$ denotes the Borel σ-algebra on $\mathbb{R}$.

▷ **Exercise 4-7** The σ-algebra $\mathscr{C}$ in Equation (4.14) is identical to $\sigma(\pi_2) = \{\pi_2^{-1}(A_2): A_2 \in \mathscr{A}_2\}$, where $\mathscr{A}_2 = \mathscr{P}(\{a, b, c\})$ [see RS-Eqs. (2.11) and (2.12))]. Write down the power set $\mathscr{P}(\{a, b, c\})$, listing its elements in the sequence corresponding to the sequence of the elements in $\mathscr{C}$.

▷ **Exercise 4-8** Gather in a set of all elements of the σ-algebra $\sigma(1_A)$ generated by the indicator variable $1_A: \Omega \to \{0, 1\}$ specified in Equation (4.15). Use the power set of $\{0, 1\}$ as the σ-algebra on $\{0, 1\}$, and RS-Equations (2.1), (2.11), and (2.12).

▷ **Exercise 4-9** Check conditions (a) to (c) of RS-Definition 1.4 for the set in the last displayed line of Equation (4.18).

▷ **Exercise 4-10** Consider Table 4.3 and specify the two subsets of Ω for the restriction of the regular causality space $\big((\Omega, \mathscr{A}), (\mathscr{F}_t)_{t \in T}, \mathscr{C}_3, \mathscr{D}_{\mathscr{C}_3}\big)$ within which we can compare

(a) receiving only treatment *a* to receiving only treatment *b*

(b) receiving both treatments to receiving only one single treatment.

▷ **Exercise 4-11** Consider the example presented in Table 1.5 and specify the measurable map $W: \Omega \to \{male, female\}$ by assigning to each element of Ω the values *male* or *female* according to the sex of the person to be sampled. Then specify W once again as the composition of the projection $\pi_1: \Omega_1 \to \Omega_1$ and a map $g: \Omega_1 \to \{male, female\}$.

Solutions

▷ **Solution 4-1** The set of all values of (π_1, π_2) is $\{(Joe, no), (Joe, yes), (Ann, no), (Ann, yes)\}$.

▷ **Solution 4-2** In this example, $\Omega_1 = \{Joe, Ann\}$ and $\mathscr{A}_1 = \{\{Joe\}, \{Ann\}, \Omega_1, \emptyset\}$ is the power set of Ω_1 (see Example 4.10). Hence,

$$\begin{aligned}
\sigma(\pi_1) &= \{\pi_1^{-1}(A_1): A_1 \in \mathscr{A}_1\} && \text{[RS-(2.11), RS-(2.12)]} \\
&= \{\pi_1^{-1}(\{Joe\}), \pi_1^{-1}(\{Ann\}), \pi_1^{-1}(\Omega_1), \pi_1^{-1}(\emptyset)\} && [\mathscr{A}_1 = \mathscr{P}(\Omega_1)] \\
&= \{\{\omega_1, \ldots, \omega_4\}, \{\omega_5, \ldots, \omega_8\}, \Omega, \emptyset\}. && \text{[RS-(2.1), Table 4.1]}
\end{aligned}$$

▷ **Solution 4-3** In Example 4.10, $\Omega_1 = \{Joe, Ann\}$ and

$$\mathscr{A}_1 = \big\{\{Joe\}, \{Ann\}, \Omega_1, \emptyset\big\}$$

is the power set of Ω_1 . Similarly, $\Omega_2 = \{no, yes\}$ and

$$\mathscr{A}_2 = \big\{\{no\}, \{yes\}, \Omega_2, \emptyset\big\}$$

is the power set of Ω_2. Hence,

$$\begin{aligned}
\mathscr{A}_1 \otimes \mathscr{A}_2 &= \sigma\big(\{A_1 \times A_2 \colon A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\}\big) && \text{[RS-(1.10)]}\\
&= \sigma\Big(\big\{\{(Joe, no)\}, \{(Joe, yes)\}, \{(Ann, no)\}, \{(Ann, yes)\},\\
&\qquad \{(Joe, no), (Joe, yes)\}, \{(Ann, no), (Ann, yes)\},\\
&\qquad \{(Joe, no), (Ann, no)\}, \{(Joe, yes), (Ann, yes)\},\\
&\qquad \{(Joe, no), (Joe, yes), (Ann, no), (Ann, yes)\}\big\}\Big) && \text{[def. of } A_1 \times A_2]\\
&= \big\{\{(Joe, no)\}, \{(Joe, yes)\}, \{(Ann, no)\}, \{(Ann, yes)\},\\
&\quad \{(Joe, no), (Joe, yes)\}, \{(Joe, no), (Ann, no)\},\\
&\quad \{(Joe, no), (Ann, yes)\}, \{(Joe, yes), (Ann, no)\},\\
&\quad \{(Joe, yes), (Ann, yes)\}, \{(Ann, no), (Ann, yes)\},\\
&\quad \{(Joe, no), (Joe, yes), (Ann, no)\},\\
&\quad \{(Joe, no), (Joe, yes), (Ann, yes)\},\\
&\quad \{(Joe, yes), (Ann, no), (Ann, yes)\},\\
&\quad \{(Joe, no), (Ann, no), (Ann, yes)\}, \Omega_1 \times \Omega_2, \emptyset\big\}. && \text{[RS-Defs. 1.4, 1.7]}
\end{aligned}$$

▷ **Solution 4-4** The two inverse images are

$$\begin{aligned}
(\pi_1, \pi_2)^{-1}(\{(Joe, no)\}) &= \big\{\omega \in \Omega \colon (\pi_1, \pi_2)(\omega) \in \{(Joe, no)\}\big\}\\
&= \{\omega_1, \omega_2\}
\end{aligned}$$

and

$$\begin{aligned}
(\pi_1, \pi_2)^{-1}(\{(Joe, no), (Ann, no)\}) &= \big\{\omega \in \Omega \colon (\pi_1, \pi_2)(\omega) \in \{(Joe, no), (Ann, no)\}\big\}\\
&= \{\omega_1, \omega_2, \omega_5, \omega_6\}.
\end{aligned}$$

▷ **Solution 4-5** In the context of a probability space $(\Omega, \mathscr{A}, P)$, the set $\{\omega_1, \omega_2, \omega_5, \omega_6\}$ is the event that the sampled person receives treatment and $\{\omega_3, \omega_4, \omega_7, \omega_8\}$ is the event that it does not receive treatment.

▷ **Solution 4-6** Although there is an uncountably infinite number of elements B in the Borel σ-algebra $\mathscr{B}$ on the set $\mathbb{R}$ of real numbers (see RS-Rem. 1.14), there are only four different inverse images $X^{-1}(B)$ of sets $B \in \mathscr{B}$ under X because

$$\forall B \in \mathscr{B}\colon \quad X^{-1}(B) = \begin{cases} \{\omega_3, \omega_4, \omega_7, \omega_8\}, & \text{if } 0 \notin B \text{ and } 1 \in B\\ \{\omega_1, \omega_2, \omega_5, \omega_6\}, & \text{if } 0 \in B \text{ and } 1 \notin B\\ \Omega, & \text{if } 0 \in B \text{ and } 1 \in B\\ \emptyset, & \text{if } 0 \notin B \text{ and } 1 \notin B. \end{cases}$$

These four inverse images are the elements of the σ-algebra $\sigma(X) = X^{-1}(\mathscr{B})$ generated by X [see RS-Eqs. (2.11) and (2.12)]. Because, in this example, X is binary, $\sigma(X) = X^{-1}(\mathscr{B}) = X^{-1}\big(\mathscr{P}(\{0,1\})\big)$.

▷ **Solution 4-7** $\mathscr{A}_2 = \mathscr{P}(\{a, b, c\}) = \big\{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \Omega_2, \emptyset\big\}$.

▷ **Solution 4-8**

$$\begin{aligned}\sigma(1_A) &= \{1_A^{-1}(C) \colon C \in \mathscr{P}(\{0,1\})\} && [\text{RS-(2.11), RS-(2.12)}]\\ &= \{1_A^{-1}(\{0\}),\ 1_A^{-1}(\{1\}),\ 1_A^{-1}(\{0,1\}),\ 1_A^{-1}(\emptyset)\} && [\mathscr{P}(\{0,1\}) = \{\{0\},\{1\},\{0,1\},\emptyset\}]\\ &= \{\{\omega_1,\omega_2,\omega_7,\omega_8\}, \{\omega_3,\dots,\omega_6,\omega_9,\dots,\omega_{12}\}, \Omega, \emptyset\}. && [\text{RS-(2.1), Table 4.2}]\end{aligned}$$

▷ **Solution 4-9** Condition (a) of RS-Definition 1.4 holds for the set $\mathscr{C}|_{\Omega_{ab}}$ because $\Omega_{ab} \in \mathscr{C}|_{\Omega_{ab}}$. Condition (b) also holds because the complement $C^c = \Omega_{ab} \setminus C$ of each set $C \in \mathscr{C}|_{\Omega_{ab}}$ is an element of $\mathscr{C}|_{\Omega_{ab}}$ as well. Finally, Condition (c) of RS-Definition 1.4 holds too. For each sequence $C_1, C_2, \dots$ of elements of $\mathscr{C}|_{\Omega_{ab}}$, the union of these elements is an element of $\mathscr{C}|_{\Omega_{ab}}$ as well. Examples of such sequences are

$$\{\omega_1,\omega_2,\omega_7,\omega_8\}, \emptyset, \emptyset, \dots$$

$$\{\omega_1,\omega_2,\omega_7,\omega_8\}, \{\omega_3,\omega_4,\omega_9,\omega_{10}\}, \Omega_{ab}, \Omega_{ab}, \dots$$

and

$$\{\omega_1,\omega_2,\omega_7,\omega_8\}, \{\omega_3,\omega_4,\omega_9,\omega_{10}\}, \Omega_{ab}, \emptyset, \emptyset, \dots .$$

The union of the elements of such sequences is always itself an element of $\mathscr{C}|_{\Omega_{ab}}$.

▷ **Solution 4-10** For comparison (a) we have to use the restriction of the regular causality space $\big((\Omega,\mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}_3, \mathscr{D}_{\mathscr{C}_3}\big)$ to $\Omega_{(a)} = \{\omega_3,\dots,\omega_6, \omega_{11},\dots,\omega_{14}\}$ and for comparison (b) we have to use the rescriction of $\big((\Omega,\mathscr{A}), (\mathscr{F}_t)_{t\in T}, \mathscr{C}_3, \mathscr{D}_{\mathscr{C}_3}\big)$ to $\Omega_{(b)} = \{\omega_3,\dots,\omega_8, \omega_{11},\dots,\omega_{16}\}$.

▷ **Solution 4-11** The potential confounder $W\colon \Omega \to \{male, female\}$ can be defined by

$$W(\omega) = \begin{cases} male, & \text{if } \omega \in \{Tom,\dots, Jim\} \times \Omega_2 \times \Omega_3 \\ female, & \text{if } \omega \in \{Ann,\dots, Mia\} \times \Omega_2 \times \Omega_3. \end{cases}$$

Alternatively, we may define W as the composition $W = g(\pi_1)$ of the first projection π_1 and a map $g\colon \Omega_1 \to \{male, female\}$, where

$$\pi_1(\omega) = \pi_1(\omega_1,\omega_2,\omega_3) = \omega_1, \quad \forall\, (\omega_1,\omega_2,\omega_3) \in \Omega_1 \times \Omega_2 \times \Omega_3,$$

$\Omega_1 = \{Tom, Tim, \dots, Mia\}$ (see the first column of Table 1.5), and

$$g(\omega_1) = \begin{cases} male, \text{ if } \omega_1 \in \{Tom,\dots, Jim\} \\ female, \text{ if } \omega_1 \in \{Ann,\dots, Mia\}. \end{cases}$$

Chapter 5
True Outcome Variable and Causal Total Effects

In chapter 4, we introduced the notion of a *regular causality setup*, which is the formal framework in which we can discuss time order of measurable maps (see ch. 3), define putative cause variables, their potential confounders, outcome variables, and study their mathematical properties. We also defined the concept of a *global potential confounder of a putative cause variable* X, denoted D_X. A global potential confounder of X comprises all potential confounders of X in the sense that the σ-algebra generated by a potential confounder of X is a subset of the σ-algebra generated by D_X (for a detailed summary see Box 4.1).

In chapter 4, we also mentioned that a regular causality setup is also called a *regular probabilistic causality setup* if there is a probability measure P on the measurable space $(\Omega, \mathscr{A})$. In the framework of a probability space $(\Omega, \mathscr{A}, P)$, the measurable maps mentioned above, including the putative cause variables, their potential confounders, and the outcome variables are *random variables on* $(\Omega, \mathscr{A}, P)$ (see RS-Def. 2.2).

In this and all other chapters to come we assume that there is such a regular probabilistic causality setup $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$, which has all the properties of a regular causality setup studied in chapter 4. In the framework of such a regular probabilistic causality setup we can meaningfully define causal effects and dependencies among two random variables X and Y on a probability space $(\Omega, \mathscr{A}, P)$.

We begin this chapter defining a *true outcome variable* $\tau_x = E^{X=x}(Y \,|\, D_X)$ as a version of the D_X-conditional expectation of Y with respect to the $(X{=}x)$-conditional probability measure $P^{X=x}$ on $\mathscr{A}$ (see RS-ch. 5), where x denotes a value of X for which we assume $P(X{=}x) > 0$. As mentioned above, with D_X we condition on *all* potential confounders of X. Although, in empirical applications, the *values* of such a true outcome variable are rarely estimable, the expectation of τ_x as well as the conditional expectation of τ_x given a covariate of X, *can be estimated* under appropriate and realistic assumptions.

Based on the concept of a true outcome variable we then define a *true total effect variable*. The expectation of the true total effect variable is then defined to be the *causal average total effect* on Y comparing x to x', where x' denotes a second value of X. Then we turn to the definition of a *causal conditional total effect* of x compared to x' given the value z of a random variable Z, and a *causal Z-conditional total effect function* comparing treatment x to treatment x'. Each of these kinds of conditional total effects or effect functions provides specific information that might be of interest in empirical causal research and evaluation studies.

In the first place, these parameters and effect functions are of a purely theoretical nature. However, in the next chapter we study how and under which assumptions these various kinds of causal effects can be identified by empirically estimable parameters and how the causal effect functions can be identified by empirically estimable functions.

Requirements

Reading this chapter requires that the reader is familiar with the contents of the first five chapters of Steyer (2024), referred to as RS-chapters 1 to 5, the first two of which have already been required in chapters 3 and 4 of this book. RS-chapter 3 deals with the concepts *expectation*, *variance*, *covariance*, and *correlation*, and RS-chapter 4 with the concept of a *conditional expectation*. The most important one for the present chapter is RS-chapter 5, introducing the concept of a *conditional expectation with respect to the probability measure* $P^{X=x}$.

In this chapter, we will often refer to some of the following assumptions and notation.

Notation and Assumptions 5.1

(a) *Let* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *be a regular probabilistic causality setup and let* D_X *denote a global potential confounder of* X.

(b) *Let* $(\Omega'_X, \mathscr{A}'_X)$ *denote the value space of* X, *let* $x \in \Omega'_X$, *let* $\{x\} \in \mathscr{A}'_X$, *and let* $1_{X=x}$ *denote the indicator of the event* $\{X=x\} = \{\omega \in \Omega\colon X(\omega) = x\}$.

(c) *Let* Y *be real-valued with positive variance.*

(d) *Assume* $0 < P(X=x) < 1$, *define the probability measure* $P^{X=x}\colon \mathscr{A} \to [0,1]$ *by* $P^{X=x}(A) = P(A \,|\, X=x)$, *for all* $A \in \mathscr{A}$, *let* $E^{X=x}(Y \,|\, D_X)$ *denote a (version of the)* D_X*-conditional expectation of* Y *with respect to* $P^{X=x}$, *and* $\mathscr{E}^{X=x}(Y \,|\, D_X)$ *the set of all such versions.*

(e) *Let* Z *be a random variable on* $(\Omega, \mathscr{A}, P)$ *and let* $(\Omega'_Z, \mathscr{A}'_Z)$ *denote its value space.*

(f) *Let* $z \in \Omega'_Z$ *be a value of* Z, *let* $\{z\} \in \mathscr{A}'_Z$, *assume* $P(Z=z) > 0$, *and define the probability measure* $P^{Z=z}\colon \mathscr{A} \to [0,1]$ *by* $P^{Z=z}(A) = P(A \,|\, Z=z)$, *for all* $A \in \mathscr{A}$.

(g) *Let* $x' \in \Omega'_X$ *and* $\{x'\} \in \mathscr{A}'_X$, *let* $1_{X=x'}$ *denote the indicator of the event* $\{X=x'\} = \{\omega \in \Omega\colon X(\omega) = x'\}$, *assume* $0 < P(X=x') < 1$, *and let* $E^{X=x'}(Y \,|\, D_X)$ *denote a (version of the)* D_X*-conditional expectation of* Y *with respect to* $P^{X=x'}$, *which is defined analogously to* $P^{X=x}$ *in Assumption (d).*

5.1 True Outcome Variable

In this section, we introduce the concept of a *(total effects) true outcome variable of* Y *given the value* x *of* X. This concept can be defined in the framework of a regular probabilistic causality setup $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *if we assume* $P(X=x) > 0$. Although this concept is not mandatory for all causality conditions (see, e.g., chs. 8 and 9, it is fundamental for several causality conditions and useful for others.

In the definition of a true outcome variable we consider a D_X-conditional expectation of Y with respect to the conditional probability measure $P^{X=x}$, denoted $E^{X=x}(Y \,|\, D_X)$ (see RS-Def. 5.4), where D_X denotes a global potential confounder of X (see Def. 4.6). Considering the term $E^{X=x}(Y \,|\, D_X)$ is tantamount to conditioning on the event $\{X=x\} = \{\omega \in \Omega\colon X(\omega) = x\}$ and on *all* potential confounders of X, presuming $P(X=x) > 0$.

Remark 5.2 [Intuitive Background] Conditioning on a global potential confounder, we share John Stuart Mill's idea already described in the preface. However, we make a slight modification. Instead of comparing *values of* Y, we compare to each other the D_X-*conditional expectations of* Y given the values x and x' of the putative cause X.

The intuitive idea is most easily explained by an example. For instance, in Example 4.10 (see also RS-Tables 1.2 and 1.4), the projection π_1 (with values *Joe* and *Ann*) takes the role of a global potential confounder of X. Now, suppose a person u may receive treatment $(X{=}1)$ or control $(X{=}0)$ (e. g., no treatment or an alternative treatment). If there is a difference between the conditional expectation values $E(Y \mid X{=}1, U{=}u)$ and $E(Y \mid X{=}0, U{=}u)$, then this difference is due to (i. e., caused by) the treatment variable X.

In this example, conditioning on a person u means that 'everything else is invariant', for example, the severity of symptoms, the motivation for treatment, educational status, and so on, which all are attributes of the person. Hence, this is a probabilistic version of the *ceteris paribus clause.*[1] Note that the treatment effect can be different for different values of the global potential confounder, that is, in this example, it can be different for different persons. ◁

Remark 5.3 [Two Concepts] The intuitive idea of a true outcome variable outlined in Remark 5.2 can mathematically be specified by *conditional expectations* $E^{X=x}(Y \mid D_X)$ *with respect to a conditional probability measure* $P^{X=x}$ (see RS-section 5.1). If $P(X{=}x) > 0$, then this concept can be used to describe how a numerical random variable Y depends on a random variable D_X given the event $\{X{=}x\}$ that X takes on the value x. The difference

$$E^{X=x}(Y \mid D_X) \;-\; E^{X=x'}(Y \mid D_X)$$

then defines the true effect variable comparing the values x and x' to each other (see Def. 5.18), provided that we assume that $E^{X=x}(Y \mid D_X)$ and $E^{X=x'}(Y \mid D_X)$ are P-unique (see sect. 5.2 below).

Alternatively, we might use the *partial conditional expectations* $E(Y \mid X{=}x, D_X)$ and $E(Y \mid X{=}x', D_X)$ (see RS-Def. 5.34) comparing conditions x and x' to each other. According to RS-Theorem 5.40, if $P(X{=}x) > 0$, then $E(Y \mid X{=}x, D_X)$ is also a version of the conditional expectation $E^{X=x}(Y \mid D_X)$ with respect to the conditional probability measure $P^{X=x}$. Hence, we may use both concepts and their corresponding notation for the definition of a true outcome variable. However, $E^{X=x}(Y \mid D_X)$ is more convenient because we can immediately utilize the (at least since Kolmogorov, 1933/1977) well-known properties of conditional expectations, which are presented in some detail in RS-chapter 4. Reading the following definition, remember that $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$ denotes the image of Ω under the map X. ◁

Definition 5.4 [True Outcome Variable]
Let the Assumptions 5.1 (a) to (d) hold. Then

$$\tau_x := E^{X=x}(Y \mid D_X) \tag{5.1}$$

is called a *true outcome variable of Y given (the value) x (of X)*.

[1] The idea of comparing conditional expectation values on the individual level is already found in Splawa-Neyman (1923/1990). Later, it has been oversimplified by Rubin, who compares the values $Y_x(u)$ and $Y_{x'}(u)$ of his potential outcome variables between two treatment conditions x and x' (see, e. g., Rubin, 1974, 2005).

Note that τ_x is a random variable on the probability space $(\Omega, \mathscr{A}, P)$, just as the putative cause X, the outcome variable Y, and the global potential confounder D_X of X. Only Y has to be real-valued whereas X and D_X may take on their values in any set Ω'_X and Ω'_{D_X}, respectively.

Remark 5.5 [Uniqueness of a True Outcome Variable τ_x] According to Equation (5.1), a true outcome variable τ_x denotes any element of the set $\mathscr{E}^{X=x}(Y|D_X)$ of versions of the D_X-conditional expectation of Y with respect to the probability measure $P^{X=x}$ (see RS-Rem. 5.8). Note that, in general, there are many elements in the set $\mathscr{E}^{X=x}(Y|D_X)$. However, if $\tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y|D_X)$, then $\tau_x \underset{P^{X=x}}{=} \tau_x^*$, that is, the versions τ_x and τ_x^* are identical $P^{X=x}$-almost surely. Hence, a true outcome variable is uniquely defined up to almost sure identity with respect to the measure $P^{X=x}$. In this case we also say that τ_x is $P^{X=x}$-unique. $\triangleleft$

Remark 5.6 [Total Effects] To emphasize, controlling for a global potential confounder D_X, we do *not* condition on potential mediators. This is why we may explicitly refer to 'total effects' if the context is ambiguous. $\triangleleft$

Remark 5.7 [A Caveat on Notation] The shortcut τ_x for (a version of) the conditional expectation $E^{X=x}(Y|D_X)$ is meaningful only in a context in which the references to a specified outcome variable Y, a specified putative cause variable X with its value x, and a global potential confounder D_X of X are unambiguous. $\triangleleft$

Remark 5.8 [Value of a True Outcome Variable] A value of a true outcome variable τ_x is called the *true (or expected) outcome of Y given the value x of X and the value d of D_X* or *given the event* $\{X=x, D_X=d\}$. According to Equation (5.1) and RS-Equation (5.25),

$$\forall\, \omega \in \Omega: \quad \tau_x(\omega) \;=\; E^{X=x}(Y|D_X)(\omega) \;=\; E^{X=x}(Y|D_X=d), \quad \text{if } \omega \in \{D_X=d\}. \tag{5.2}$$

Furthermore, if $P(X=x, D_X=d) > 0$, then

$$E^{X=x}(Y|D_X=d) \;=\; E(Y|X=x, D_X=d), \tag{5.3}$$

[see RS-Eq. (5.26)]. That is, a value of a true outcome variable τ_x is identical to a conditional expectation value of Y given the value x of X and the value d of a global potential confounder D_X. Such a conditional expectation value is uniquely defined only if $P(X=x, D_X=d) > 0$ (see RS-Rem. 4.33). $\triangleleft$

Remark 5.9 [The Values of τ_x Cannot be Observed] Because the values of τ_x are the conditional expectation values $E^{X=x}(Y|D_X=d)$, they cannot be observed. In contrast, the values of X and Y can be observed, if the random experiment represented by $(\Omega, \mathscr{A}, P)$ is conducted. In most applications, the values of τ_x even cannot be estimated (for more details, see Rem. 5.13). $\triangleleft$

Remark 5.10 [τ_x is D_X-Measurable] Note that τ_x is a random variable on the probability space $(\Omega, \mathscr{A}, P)$ that is measurable with respect to D_X, that is, $\sigma(\tau_x) \subset \sigma(D_X)$. This immediately follows from RS-Definition 5.4 (a). $\triangleleft$

Remark 5.11 [Factorization of τ_x] According to RS-Corollary 4.18, there is a measurable function $g_x \colon (\Omega'_{D_X}, \mathscr{A}'_{D_X}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ such that $\tau_x = E^{X=x}(Y|D_X) = g_x(D_X)$ is the composition of D_X and the *factorization g_x of $E^{X=x}(Y|D_X)$* (see RS-sect. 5.29). Therefore, if $P(X=x, D_X=d) > 0$, then we may rewrite Equations (5.2) and (5.3) as follows:

Table 5.1. Joe and Ann with self-selection revisited

Possible outcomes ω_i (Unit, Treatment, Success)	Probability measures			Observables			Conditional expectations					
	$P(\{\omega_i\})$	$P^{X=0}(\{\omega_i\})$	$P^{X=1}(\{\omega_i\})$	Person variable U	Treatment variable X	Outcome variable Y	$P(X=1 \mid U)$	$E(Y \mid X)$	$E(Y \mid X, U)$	$\tau_0 = E^{X=0}(Y \mid U)$	$\tau_1 = E^{X=1}(Y \mid U)$	$\delta_{10} = \tau_1 - \tau_0$
$\omega_1 = (Joe, no, -)$	.144	.24	0	*Joe*	0	0	.04	.6	.7	.7	.8	.1
$\omega_2 = (Joe, no, +)$	.336	.56	0	*Joe*	0	1	.04	.6	.7	.7	.8	.1
$\omega_3 = (Joe, yes, -)$	.004	0	.01	*Joe*	1	0	.04	.42	.8	.7	.8	.1
$\omega_4 = (Joe, yes, +)$	.016	0	.04	*Joe*	1	1	.04	.42	.8	.7	.8	.1
$\omega_5 = (Ann, no, -)$	.096	.16	0	*Ann*	0	0	.76	.6	.2	.2	.4	.2
$\omega_6 = (Ann, no, +)$	.024	.04	0	*Ann*	0	1	.76	.6	.2	.2	.4	.2
$\omega_7 = (Ann, yes, -)$	.228	0	.57	*Ann*	1	0	.76	.42	.4	.2	.4	.2
$\omega_8 = (Ann, yes, +)$	.152	0	.38	*Ann*	1	1	.76	.42	.4	.2	.4	.2

$$\begin{aligned}\forall\, \omega \in \Omega: \quad \tau_x(\omega) &= E^{X=x}(Y|D_X)(\omega) = g_x(D_X)(\omega) = g_x\big(D_X(\omega)\big) = g_x(d) \\ &= E^{X=x}(Y|D_X{=}d) = E(Y|X{=}x, D_X{=}d), \quad \text{if } \omega \in \{D_X{=}d\},\end{aligned} \tag{5.4}$$

Hence, in order to assign a value of τ_x to an outcome $\omega \in \Omega$ of the random experiment, first we may assign to ω — via the global potential confounder D_X — a value d of this map, and then assign to d — via the factorization g_x — the corresponding conditional expectation value $E^{X=x}(Y|D_X{=}d)$, which is identical to $E(Y|X{=}x, D_X{=}d)$ [see Eq. (5.3) and Example 5.12]. $\triangleleft$

Example 5.12 [Joe and Ann With Self-Selection] Consider the random experiment presented in Table 5.1. It describes the same random experiment as Table 1.2 and it has the same structure as the experiment presented in Table 4.1. However, in Table 5.1 we added some terms that are important to illustrate the true outcome variables τ_0, τ_1 and their difference $\tau_1 - \tau_0$. Furthermore, we also used the more general notation for conditional expectations, which also apply if the outcome variable is not an indicator variable with values 0 and 1. In Example 4.10, we already specified the set $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$ of possible outcomes with its elements $\omega_i = (u, \omega_X, \omega_Y)$, the probability space $(\Omega, \mathscr{A}, P)$, the random variables $U = \pi_1$, X, Y, and the filtration $(\mathscr{F}_t)_{t \in T}$ in $\mathscr{A}$, where $T = \{1,2,3\}$. Furthermore, in Example 4.10 we also asserted that $U = \pi_1$ is a global potential confounder of the treatment variable X, playing the role of D_X in the general theory. Therefore, $\tau_x = E^{X=x}(Y|D_X) = E^{X=x}(Y|U)$ for both treatments $x = 0$ and $x = 1$.

Now we specify the values of the true outcome variables $\tau_x = E^{X=x}(Y|U)$ for the two treatment conditions $x = 0$ and $x = 1$. According to Remark 5.11,

$$\tau_0 = g_0(U), \tag{5.5}$$

where $U\colon \Omega \to \Omega_U$ with

$$U(\omega_i) \;=\; U\big((u,\omega_X,\omega_Y)\big) \;=\; u, \quad \text{for all } \omega_i \in \Omega, \tag{5.6}$$

and $g_0\colon \Omega_U \to \mathbb{R}$ with

$$g_0(u) \;=\; E(Y\,|\,X{=}0, U{=}u), \quad \text{for all } u \in \Omega_U \tag{5.7}$$

[see Eq. (5.4)]. Hence, in order to assign a value of τ_0 to an outcome $\omega_i \in \Omega$ of the random experiment, first we have to assign to ω_i a value u (Joe or Ann) of the person variable U, and then assign to u via g_0 the number $E(Y\,|\,X{=}0, U{=}u)$, that is, the corresponding conditional expectation value.

If, for instance, $\omega_3 = (Joe, yes, -)$ (see the third row in Table 5.1), then $U(\omega_3) = Joe$, and the value of τ_0 is

$$\tau_0(\omega_3) \;=\; g_0\big(U(\omega_3)\big) \;=\; g_0(Joe) \;=\; E(Y\,|\,X{=}0, U{=}Joe) \;=\; P(Y{=}1\,|\,X{=}0, U{=}Joe) \;=\; .7.$$

This is true *even though* $X(\omega_3) = 1$ and the value of the conditional expectation $E(Y\,|\,X, U)$ is

$$E(Y\,|\,X, U)(\omega_3) \;=\; E(Y\,|\,X{=}1, U{=}Joe) \;=\; P(Y{=}1\,|\,X{=}1, U{=}Joe) \;=\; .8.$$

Hence, the true outcome variable τ_0 given treatment 0 takes on a well-defined value for ω_3 *even though* the unit drawn receives treatment 1. This illustrates the distinction between the random variables $\tau_0 = E^{X=0}(Y\,|\,U)$ and $E(Y\,|\,X, U)$. While $\tau_0 = g_0(U)$ is solely a function of U [see Eq. (5.5)], the conditional expectation $E(Y\,|\,X, U)$ is a function of X and U, that is, there is a function $g\colon \Omega'_X \times \Omega_U \to \mathbb{R}$ such that $E(Y\,|\,X, U) = g(X, U)$, where $g(X, U)$ denotes the composition of (X, U) and g. Again, we simply say that *$E(Y\,|\,X, U)$ is a function of X and U*.

Because, in this example, the outcome variable Y is dichotomous with values 0 and 1, the conditional expectation value $E(Y\,|\,X{=}0, U{=}u)$ is also the conditional probability $P(Y{=}1\,|\,X{=}0, U{=}u)$ of success, and because in this example, U has only two values, Joe and Ann, the true outcome variable τ_0 also has only two different values, the two conditional probabilities $P(Y{=}1\,|\,X{=}0, U{=}Joe) = .7$ and $P(Y{=}1\,|\,X{=}0, U{=}Ann) = .2$ (see Table 5.1). Hence, if $\omega_i \in \{U{=}Ann\} = \{\omega \in \Omega\colon U(\omega) = Ann\}$, then the value of τ_0 is

$$\tau_0(\omega_i) \;=\; g_0\big(U(\omega_i)\big) \;=\; g_0(Ann) \;=\; E(Y\,|\,X{=}0, U{=}Ann) \;=\; P(Y{=}1\,|\,X{=}0, U{=}Ann) \;=\; .2.$$

Similarly, the *true outcome variable* $\tau_1 = E^{X=1}(Y\,|\,U)$ given treatment 1 is specified by

$$\tau_1 \;=\; g_1(U), \tag{5.8}$$

where $g_1\colon \Omega_U \to \mathbb{R}$ is defined by

$$g_1(u) \;=\; E(Y\,|\,X{=}1, U{=}u), \quad \text{for all } u \in \Omega_U. \tag{5.9}$$

Hence, if $\omega_i \in \{U{=}Joe\} = \{\omega \in \Omega\colon U(\omega) = Joe\}$, then the value of τ_1 is

$$\tau_1(\omega_i) \;=\; g_1\big(U(\omega_i)\big) \;=\; g_1(Joe) \;=\; E(Y\,|\,X{=}1, U{=}Joe) \;=\; P(Y{=}1\,|\,X{=}1, U{=}Joe) \;=\; .8,$$

and if $\omega_i \in \{U{=}Ann\}$, then the value of τ_1 is

$$\tau_1(\omega_i) \;=\; g_1\big(U(\omega_i)\big) \;=\; g_1(Ann) \;=\; E(Y\,|\,X{=}1, U{=}Ann) \;=\; P(Y{=}1\,|\,X{=}1, U{=}Ann) \;=\; .4.$$

The last two columns of Table 5.1 show which values the two random variables $\tau_0 = E^{X=0}(Y\,|\,D_X)$ and $\tau_1 = E^{X=1}(Y\,|\,D_X)$ assign to each of the eight possible outcomes ω_i of the random experiment. $\triangleleft$

Remark 5.13 [The Fundamental Problem of Causal Inference] In Remark 5.9 we already mentioned that the values of τ_x cannot be estimated. The reason is as follows: Consider the last but one column of Table 5.1 with the values $E^{X=1}(Y|U=u)$ of τ_1, namely .8 for $u = Joe$ and .4 for $u = Ann$. Imagine that we would conduct the random experiment presented in this table and that Joe is sampled and he selects treatment $(X=1)$. Then we can observe the value of Y for Joe under treatment, which is an estimate of $E^{X=1}(Y|U=Joe)$, although a bad one. However, in many applications, we cannot estimate $E^{X=0}(Y|U=Joe)$ at the same time because, if he selects treatment, then we cannot observe his value of Y under control, and vice versa. This has been called "the fundamental problem of causal inference" by Holland (1986). Observing Joe's value of the outcome variable Y, first under control and then under treatment may yield a value of Y for Joe under treatment that may *systematically* differ from the corresponding value that would be observed if Joe would not have been in control, in the first place. There are numerous reasons for this fact that are discussed in great detail by Campbell and Stanley (1966). From a formal point of view, it should be noted that observing Joe's value of the outcome variable Y under control and then again under treatment refers to a different random experiment (represented by a different probability space) than observing his outcome value under treatment without the preceding observation under control. $\lhd$

Example 5.14 [Nonorthogonal Factors] In Example 4.2.3 we already showed that U is a global potential confounder of X and that Z is a potential confounder of X. In the last three columns of Table 1.5 we specified the true outcome variables $\tau_0 = E^{X=0}(Y|U)$, $\tau_1 = E^{X=1}(Y|U)$, and $\tau_2 = E^{X=2}(Y|U)$. In this example, all three true outcome variables are uniquely defined because $P(X=1|U) > 0$ (see RS-Th. 5.27). $\lhd$

Remark 5.15 [True Outcomes vs. Potential Outcomes] Rubin (1974, 2005) assumes that, given an observational unit u and a treatment condition x, the values of his potential outcome variables Y_0 and Y_1 are fixed numbers. In the example presented in Table 5.1, this would mean to replace the two true outcome variables τ_0 and τ_1 by the two potential outcome variables Y_0 and Y_1 that can take on only the values 0 and 1. Substantively speaking, this would mean that, given a concrete treatment and a concrete observational unit, the outcome is fixed to 0 or 1. For example, if the outcome is being alive $(Y=1)$ or not $(Y=0)$ at the age of 80 and the treatment is receiving $(X=1)$ or no receiving $(X=0)$ an anti-smoking therapy before the age of 40, then this deterministic idea is not in line with our knowledge of causes for being alive or not being alive at the age of 80.

In contrast, the two true outcome variables τ_0 and τ_1 can take on any real number as values. In the example of Table 5.1, they can take on any value between 0 and 1, inclusively. In the smoking example, their values would be the person-specific probabilities of being alive at the age of 80 given treatment or given no treatment. [As we will see later on, it is irrelevant that we usually cannot estimate these conditional probabilities $P(Y=1|X=1, U=u)$.] Therefore, the true outcome variables can be considered to be a generalization of the potential outcome variables. Most important, in contrast to potential outcome variables, true outcome variables are in line with the idea that mediators and events that may occur between treatment and outcome might also affect the outcome variable Y.

Hence, in Rubin's potential outcome approach the observational unit u and treatment x determine *the value of the outcome variable* Y. In contrast, according to true outcome theory, u and x only determine *the conditional distribution* $P_{Y|X=x,U=u}$, and with it, the conditional expectation value $E(Y|X=x,U=u)$ of the outcome variable Y. ◁

5.2 True Total Effect Variable

Now we introduce the concept of a *true* $(D_X=d)$*-conditional total effect comparing x to x'*, the latter being two different values of the focused putative cause variable X. The basic idea of this concept is to hold constant all those other possible causes of Y that are prior or simultaneous to X at one combination of their values and then compare the conditional expectation values of Y between the two values x and x' of X.

As mentioned before, the intuitive version of this basic idea goes back at least to John Stuart Mill (1843/1865). It is often referred to as the *ceteris paribus clause* (all other things equal). Remember that a global potential confounder can be a multivariate random variable, consisting of several univariate random variables (potential confounders). Also note that a global potential confounder of X does not comprise potential mediators, that is, variables that might be affected by X and might have an effect on Y. The term 'does not comprise' is used here in the sense that the σ-algebra generated by such a potential mediator is not a subset of the σ-algebra generated by D_X.

If we assume $P(X=x, D_X=d) > 0$ and $P(X=x', D_X=d) > 0$, then a true $(D_X=d)$-conditional total effect

$$E(Y\,|\,X=x, D_X=d) \;-\; E(Y\,|\,X=x', D_X=d)$$

is a uniquely defined number [see RS-Eq. (3.23)]. This difference may also be called *the true total effect on Y comparing x to x' given (the event)* $\{D_X=d\}$.

For different values d and d' of the global potential confounder D_X, the true conditional total effects can differ from each other. This necessitates a second concept, a *true total effect variable of Y comparing x to x'*, that is,

$$E^{X=x}(Y|D_X) \;-\; E^{X=x'}(Y|D_X) \;=\; \tau_x \;-\; \tau_{x'}.$$

In general, this random variable is not uniquely defined so that there can be many versions of such a true total effect variable. However, assuming P-uniqueness of $\tau_x = E^{X=x}(Y|D_X)$ and of $\tau_{x'} = E^{X=x'}(Y|D_X)$ implies that the difference $\tau_x - \tau_{x'}$ is P-unique as well [see RS-Box 5.1 (viii)].

Remark 5.16 [True Outcome Variables Are $P^{X=x}$-Unique] In Remark 5.5 we already mentioned that, according to its definition as a D_X-conditional expectation of Y with respect to the probability measure $P^{X=x}$, a true outcome variable τ_x is $P^{X=x}$-unique. That is, if $\mathscr{E}^{X=x}(Y|D_X)$ denotes the set of all versions of the D_X-conditional expectation of Y with respect to the measure $P^{X=x}$, then

$$\forall\, \tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y|D_X)\colon\; \tau_x \underset{P^{X=x}}{=} \tau_x^* \tag{5.10}$$

holds, where

$$\tau_x \underset{P^{X=x}}{=} \tau_x^* \quad :\Leftrightarrow \quad P^{X=x}\big(\{\omega \in \Omega\colon \tau_x(\omega) = \tau_x^*(\omega)\}\big) = 1. \tag{5.11}$$

That is, whenever τ_x and τ_x^* are two versions of the D_X-conditional expectation of Y with respect to $P^{X=x}$, then they are identical $P^{X=x}$-almost surely. ◁

Remark 5.17 [Assuming P-Uniqueness of True Outcome Variables] However, in the definition of a *true total effect variable* (see Def. 5.18) we assume that the true outcome variables τ_x and $\tau_{x'}$ are P-unique. Remember,

$$\tau_x \text{ is } P\text{-unique} \quad :\Leftrightarrow \quad \forall\, \tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y \mid D_X)\colon\ \tau_x \underset{P}{=} \tau_x^*, \tag{5.12}$$

where

$$\tau_x \underset{P}{=} \tau_x^* \quad :\Leftrightarrow \quad P\big(\{\omega \in \Omega\colon \tau_x(\omega) = \tau_x^*(\omega)\}\big) = 1. \tag{5.13}$$

According to RS-Theorem 5.27, P-uniqueness of τ_x is equivalent to $P(X{=}x \mid D_X) \underset{P}{>} 0$, which is defined by

$$P(X{=}x \mid D_X) \underset{P}{>} 0 \quad :\Leftrightarrow \quad P\big(\{\omega \in \Omega\colon P(X{=}x \mid D_X)(\omega) > 0\}\big) = 1. \tag{5.14}$$

That is, P-uniqueness of τ_x is equivalent to $P(X{=}x \mid D_X)$ being positive, P-almost surely. ◁

Definition 5.18 [True Total Effect Variable and True Total Effect]
Let the Assumptions 5.1 (a) to (d) and (g) hold. Furthermore, let τ_x and $\tau_{x'}$ denote true outcome variables of Y given the values x and x' of X, respectively.

(i) If τ_x and $\tau_{x'}$ are P-unique, then we call $CTE_{D_X;xx'}\colon \Omega'_{D_X} \to \mathbb{R}$ a version of the true total effect function comparing x to x' (with respect to Y), if

$$CTE_{D_X;xx'}(D_X) \underset{P}{=} \tau_x - \tau_{x'} \tag{5.15}$$

holds for the composition $CTE_{D_X;xx'}(D_X)$ of D_X and $CTE_{D_X;xx'}$.

(ii) If τ_x and $\tau_{x'}$ are P-unique, then the composition $CTE_{D_X;xx'}(D_X)$ is called a version of the true total effect variable comparing x to x' (with respect to Y).

(iii) If d is a value of D_X such that $P(X{=}x, D_X{=}d), P(X{=}x', D_X{=}d) > 0$, then we call

$$CTE_{D_X;xx'}(d) \ := \ E(Y \mid X{=}x, D_X{=}d) - E(Y \mid X{=}x', D_X{=}d) \tag{5.16}$$

the true total effect on Y given the value d of D_X comparing x to x'.

Of course, the concepts of a true total effect and a true total effect variable are of a theoretical nature and can be estimated only under rather restrictive assumptions. However, other causal total effects such as the expectation of $\tau_x - \tau_{x'}$ (see sect. 5.3) can be estimated under realistic assumptions.

Remark 5.19 [$CTE_{D_X;xx'}(D_X)$ Versus $CTE_{D_X;xx'}$] While $CTE_{D_X;xx'}(D_X)$ is a random variable on the probability space $(\Omega, \mathscr{A}, P)$ assigning values to all $\omega \in \Omega$, the function $CTE_{D_X;xx'}$ is a random variable on the probability space $(\Omega'_{D_X}, \mathscr{A}'_{D_X}, P_{D_X})$ that assigns values to all $d \in \Omega'_{D_X}$ (see RS-Rem. 2.39). From a substantive point of view the two maps contain the same information. ◁

Example 5.20 [Joe and Ann With Self-Selection] In Example 5.12, we already specified the true outcome variables τ_0 and τ_1. Their difference $\tau_1 - \tau_0$ is a version of the true total effect variable. Because, in this example, the person variable U is a global potential confounder of X, the true total effect variable may also be written

$$CTE_{U;10}(U) \;=\; E^{X=1}(Y|U) \;-\; E^{X=0}(Y|U) \;=\; \tau_1 - \tau_0 .$$

It is a random variable on the probability space $(\Omega, \mathscr{A}, P)$. In contrast, the true total effect function $CTE_{U;10}\colon \Omega_U \to \mathbb{R}$ is a random variable on the probability space $\big(\Omega_U, \mathscr{P}(\Omega_U), P_U\big)$.

Using the functions g_0 and g_1 defined by Equations (5.7) and (5.9),

$$CTE_{U;10} \;:=\; g_1 - g_0.$$

Its values are

$$CTE_{U;10}(Joe) \;=\; g_1(Joe) - g_0(Joe) \;=\; .8 - .7 \;=\; .1,$$

the treatment effect of Joe, and

$$CTE_{U;10}(Ann) \;=\; g_1(Ann) - g_0(Ann) \;=\; .4 - .2 \;=\; .2,$$

the treatment effect of Ann. Hence, $CTE_{U;10}$ assigns to each person $u \in \Omega_U$ the person-specific true effect on the outcome variable Y (*success*) comparing $x = 1$ to $x = 0$. ◁

Remark 5.21 [Uniqueness of a True Total Effect Variable] For simplicity, we denote a true total effect variable $CTE_{D_X;xx'}(D_X)$ also by

$$\delta_{xx'} \;=\; \tau_x - \tau_{x'}. \tag{5.17}$$

It is a random variable on $(\Omega, \mathscr{A}, P)$ that is not necessarily unique. However, if $P(X{=}x)$ and $P(X{=}x')$ are positive and $\tau_x, \tau_{x'}$ are P-unique, then the difference variable $\delta_{xx'}$ is also P-unique [see RS-Box 5.1 (viii)]. That is, two versions of such a difference variable are identical P-almost surely. Hence, if $\delta_{xx'}$ and $\delta^*_{xx'}$ are two such versions, then $E(\delta_{xx'}) = E(\delta^*_{xx'})$ [see RS-Box 3.1 (vi) with $A = \Omega$]. Other implications are that $\delta_{xx'}$ and $\delta^*_{xx'}$ have identical distributions, variances, and covariances with other random variables (see RS-Rem. 4.11). ◁

Remark 5.22 [Uniqueness of a True Total Effect Function] In contrast to $\delta_{xx'}$, which is a function on Ω, a true total effect *function* $CTE_{D_X;xx'}$ is a function on Ω'_{D_X}. It is also not necessarily unique. However, if there are two versions of $CTE_{D_X;xx'}$, then they are identical P_{D_X}-almost surely. This follows from $\delta_{xx'} \underset{P}{=} \delta^*_{xx'}$ (see Rem. 5.21) and RS-Theorem 2.54. ◁

Remark 5.23 [Values of a True Total Effect Variable] As mentioned above, $\delta_{xx'}$ is a random variable on $(\Omega, \mathscr{A}, P)$. According to Equation (5.1), and RS-Equations (5.25) and (5.26),

$$\begin{aligned}
\forall\, \omega \in \Omega\colon \quad \delta_{xx'}(\omega) \;&=\; E^{X=x}(Y|D_X)(\omega) - E^{X=x'}(Y|D_X)(\omega)\\
&=\; E^{X=x}(Y\,|\,D_X{=}d) - E^{X=x'}(Y\,|\,D_X{=}d)\\
&=\; E(Y|X{=}x, D_X{=}d) - E(Y|X{=}x', D_X{=}d), \qquad \text{if } \omega \in \{D_X{=}d\}.
\end{aligned} \tag{5.18}$$

Hence, a value of a true total effect variable $\delta_{xx'}$ is the difference between the conditional expectation values of Y given the values (x, d) and (x', d) of (X, D_X). If d is a value of D_X such that $P(X{=}x, D_X{=}d), P(X{=}x', D_X{=}d) > 0$, then

$$\forall\, \omega \in \Omega\colon \quad \delta_{xx'}(\omega) \;=\; CTE_{D_X;xx'}(d), \qquad \text{if } \omega \in \{D_X{=}d\} \tag{5.19}$$

[see Def. 5.18 (iii) and RS-Def. 5.30]. If $P(X{=}x, D_X{=}d), P(X{=}x', D_X{=}d) > 0$, then this value is identical for all versions of the true total effect function $CTE_{D_X;xx'}$ and for all versions of the true total effect variable $CTE_{D_X;xx'}(D_X) = \delta_{xx'}$.

◁

Remark 5.24 [Probabilistic Ceteris Paribus Clause] Considering a value $CTE_{D_X;xx'}(d)$ is tantamount to comparing x to x' (with respect to the outcome variable Y) keeping constant the global potential confounder D_X, and with it, *keeping constant all potential confounders of* X. Keeping constant D_X, is the translation of the *ceteris paribus clause* for total effects into probability theory. $\triangleleft$

Example 5.25 [Joe and Ann With Self-Selection] In Example 5.20, we already specified the true total effect function $CTE_{U;10}\colon \Omega_U \to \mathbb{R}$. The corresponding true total effect variable $CTE_{U;10}(U)$ is the composition of U and $CTE_{U;10}$. It is a random variable on $(\Omega, \mathscr{A}, P)$. For all $\omega_i \in \Omega$, its values are assigned by

$$\delta_{xx'}(\omega_i) = CTE_{U;10}(U)(\omega_i) = CTE_{U;10}\big(U(\omega_i)\big) = \begin{cases} CTE_{U;10}(Joe) = .1, & \text{if } \omega_i \in \{U{=}Joe\} \\ CTE_{U;10}(Ann) = .2, & \text{if } \omega_i \in \{U{=}Ann\} \end{cases}$$

(see the last column of Table 5.1). $\triangleleft$

5.3 Causal Average Total Effect

Now we define a *causal average total effect* by the expectation

$$E(\delta_{xx'}) \ = \ E(\tau_x - \tau_{x'}) \ = \ E(\tau_x) \ - \ E(\tau_{x'}) \tag{5.20}$$

of a true total effect variable [see Eq. (5.17) and RS-Box 3.1 (vii)]. Hence, the causal average total effect is the *expectation* of a true total effect variable $\delta_{xx'}$; it is not an unweighted average as the name might suggest.

As mentioned before, this expectation can be estimated under assumptions that are less restrictive than those that allow estimating the total effect variable $\delta_{xx'}$ itself. This will be detailed in the chapters on unbiasedness and its sufficient conditions.

Definition 5.26 [Causal Average Total Effect]
Let the Assumptions 5.1 (a) to (d) and (g) hold, let τ_x and $\tau_{x'}$ denote true outcome variables of Y given the values x and x' of X, respectively, and let $\delta_{xx'} = \tau_x - \tau_{x'}$. Furthermore, assume that τ_x and $\tau_{x'}$ are P-unique and that the expectations $E(\tau_x)$ and $E(\tau_{x'})$ are finite. Then

$$ATE_{xx'} \ := \ E(\delta_{xx'}), \tag{5.21}$$

is called the causal average total effect on Y comparing x to x' (with respect to P).

Taking the expectation of $\delta_{xx'}$ (with respect to P) means that we consider the average total effect *with respect to the measure* P, that is,

$$ATE_{xx'} \ = \ E(\delta_{xx'}) \ = \ E\big(CTE_{D_X;xx'}(D_X)\big) \ = \ E\big(E^{X=x}(Y \mid D_X) - E^{X=x'}(Y \mid D_X)\big). \tag{5.22}$$

Note again that $CTE_{D_X;xx'}(D_X)$ is a random variable on $(\Omega, \mathscr{A}, P)$. In principle, we can also consider the average total effect with respect to another measure than P (for more details see Rem. 5.45).

Remark 5.27 [Expectation of $CTE_{D_X;xx'}$ With Respect to the Distribution of D_X] According to RS-Theorem 3.11, $ATE_{xx'}$ is identical to the expectation of a true total effect function $CTE_{D_X;xx'}$ with respect to the distribution of the global potential confounder D_X, that is,

$$ATE_{xx'} \;=\; E(\delta_{xx'}) \;=\; E_{D_X}(CTE_{D_X;xx'}). \tag{5.23}$$

Note that $CTE_{D_X;xx'}$ is a random variable on $(\Omega'_{D_X}, \mathscr{A}'_{D_X}, P_{D_X})$, not on $(\Omega, \mathscr{A}, P)$ (see again RS-Rem. 2.39). ◁

Remark 5.28 [P-Uniqueness and Expectations of τ_x and $\tau_{x'}$] According to RS-Box 5.1 (vii), P-uniqueness of τ_x also implies

$$\forall\, \tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y|D_X): \quad E(\tau_x) = E(\tau_x^*). \tag{5.24}$$

Hence, under the assumptions of Definition 5.26, the expectation of τ_x is a uniquely defined number, and the same applies to the expectation of $\tau_{x'}$. [2] ◁

Remark 5.29 [Expectation of $\delta_{xx'}$] Under the assumptions of Definition 5.26, $E(\delta_{xx'})$ is a uniquely defined real number because

$$\begin{aligned} E(\delta_{xx'}) \;&= E(\tau_x - \tau_{x'}) \;=\; E(\tau_x) - E(\tau_{x'}) && [(5.20)] \\ &= E\big(E^{X=x}(Y|D_X)\big) - E\big(E^{X=x'}(Y|D_X)\big). && [(5.1)] \end{aligned}$$

If the true outcome variables τ_x and $\tau_{x'}$ are not P-unique or the expectations $E(\tau_x)$ and $E(\tau_{x'})$ are not finite, then the causal average total effect $ATE_{xx'}$ is not defined. ◁

Now we gather some sufficient conditions for $E(\tau_x)$ to be finite. For more details and proofs see SN-Remark 14.47, from which the following remark is adapted.

Remark 5.30 [Sufficient Conditions for Finiteness of $E(\tau_x)$] Remember, the expectation of a random variable Y on a probability space $(\Omega, \mathscr{A}, P)$ exists if the integral $\int Y^+ \, dP$ of the positive part of Y or the integral $\int Y^- \, dP$ of the negative part of Y is finite (see RS-Def. 3.1). Hence, under the assumptions of Definition 5.26, some sufficient conditions for the expectation $E(\tau_x)$ of any version $\tau_x \in \mathscr{E}^{X=x}(Y|D_X)$ to exists and to be finite are:

(a) $\sigma(D_X)$ is a finite set and $E^{X=x}(Y)$ is finite
(b) τ_x has only a finite number of real values
(c) τ_x is P-almost surely bounded on both sides, that is, $\exists\, \alpha \in \mathbb{R}: -\alpha \underset{P}{\le} \tau_x \underset{P}{\le} \alpha$
(d) Y is P-almost surely bounded on both sides, that is, $\exists\, \alpha \in \mathbb{R}: -\alpha \underset{P}{\le} Y \underset{P}{\le} \alpha$.

Note that $0 \underset{P}{\le} Y \underset{P}{\le} \alpha$, for $0 < \alpha \in \mathbb{R}$, as well as $Y = 1_A$, for $A \in \mathscr{A}$, are special cases of (d). ◁

Remark 5.31 [Causal Average Effect] If X represents a treatment variable, then $ATE_{xx'}$ is also called the 'average causal effect' or the 'causal average treatment effect' comparing treatment x to treatment x', which is unambiguous as long as no direct and/or indirect treatment effects are discussed in the same context. ◁

---

[2] The expectation $E(\tau_x)$ corresponds to the term $E[Y|do(x)]$ in Pearl's and to $E(Y_x)$ in Rubin's terminologies (see, e. g., Pearl, 2009, p. 108 and Rubin, 2005, p. 323).

Remark 5.32 [Hypothesis of a t-Test in a Randomized Experiment] The causal average total effect to be zero is what is tested by a t-test of the hypothesis $\mu_0 = \mu_1$ [i. e., $E(Y|X{=}0) = E(Y|X{=}1)$] *in an experiment with randomized assignment of an observational unit to a treatment condition*. The expectations $E(\tau_0)$ and $E(\tau_1)$ occurring in Equation (5.20) are estimated, for example, by the sample group means in such a randomized experiment with treatment conditions 0 and 1. In this case, the expectations $E(\tau_0)$ and $E(\tau_1)$ are identical to the conditional expectation values $E(Y|X{=}0)$ and $E(Y|X{=}1)$, respectively. (For more details see ch. 8, in particular Th. 8.43.) $\lhd$

Example 5.33 [Joe and Ann With Self-Assignment] In Tables 1.2 and 5.1 we presented an example in which the person sampled assigns him- or herself to one of the two treatment conditions, 0 and 1. In Example 5.12, we specified the true outcome variables $\tau_0 = g_0(U)$ and $\tau_1 = g_1(U)$ and their values $g_0(u)$ and $g_1(u)$ for this example.

Now we illustrate the *causal average total effect* comparing treatment to control, which can be computed as follows:

$$\begin{aligned}
&ATE_{10} \;=\; E\big(E^{X=1}(Y|U) - E^{X=0}(Y|U)\big) \;=\; E(\tau_1 - \tau_0) \;=\; E(\tau_1) - E(\tau_0)\\
&= \sum_u g_1(u)\cdot P(U{=}u) - \sum_u g_0(u)\cdot P(U{=}u)\\
&= P(Y{=}1|X{=}1, U{=}\mathit{Joe})\cdot P(U{=}\mathit{Joe}) \;+\; P(Y{=}1|X{=}1, U{=}\mathit{Ann})\cdot P(U{=}\mathit{Ann})\\
&\quad -\; \big(P(Y{=}1|X{=}0, U{=}\mathit{Joe})\cdot P(U{=}\mathit{Joe}) \;+\; P(Y{=}1|X{=}0, U{=}\mathit{Ann})\cdot P(U{=}\mathit{Ann})\big)\\
&= .80\cdot .50 \;+\; .40\cdot .50 \;-\; (.70\cdot .50 \;+\; .20\cdot .50) \;=\; .15,
\end{aligned}$$

using the transformation theorem (see RS-Th. 3.11 and RS-Rem. 3.13).

Figure 1.3 visualizes the causal average total effect, which, in this example, is not identical to the difference

$$E(Y|X{=}1) - E(Y|X{=}0) \;=\; P(Y{=}1|X{=}1) - P(Y{=}0|X{=}0) \;=\; .48 - .6 \;=\; -.18.$$

Figure 1.3 also illustrates various conditional probabilities of success. The points marked by the dashed line are the probabilities $P(Y{=}1|X{=}1, U{=}\mathit{Joe})$ and $P(Y{=}1|X{=}0, U{=}\mathit{Joe})$ of success for Joe given that he is treated and given that he is not treated, respectively. Similarly, the two points marked by the solid line show the probabilities $P(Y{=}1|X{=}1, U{=}\mathit{Ann})$ and $P(Y{=}1|X{=}0, U{=}\mathit{Ann})$ of success for Ann given that she is treated and given that she is not treated, respectively. The points marked by the dotted line represent the conditional probabilities $P(Y{=}1|X{=}1)$ and $P(Y{=}1|X{=}0)$ of success given treatment and given control, respectively. The size of the area of the dotted circles is proportional to the conditional probabilities $P(U{=}u|X{=}x)$ that are used in the computation of the conditional expectation values

$$E(Y|X{=}x) \;=\; \sum_u E(Y|X{=}x, U{=}u)\cdot P(U{=}u|X{=}x) \tag{5.25}$$

[see RS-Box 3.2 (ii)]. $\lhd$

5.4 Causal Conditional Total Effect and Total Effect Function

So far we considered the random variables X, Y, and D_X, a global potential confounder of X. Now we bring into play an additional random variable Z on $(\Omega, \mathscr{A}, P)$ and introduce the concepts of the *causal conditional total effect* given the value z of Z and of a *causal* Z-*conditional total effect function*. Often Z is a pretest that assesses the 'same' attribute as the outcome variable Y, only prior to treatment. In other examples, Z could be X itself. First, we explain the assumptions based on which we can introduce these concepts, present their definitions, and then turn to re-aggregating conditional effects.

5.4.1 Notation, Assumptions and Definitions

One of the assumptions in the definition of a causal conditional total effect is $P^{Z=z}$-uniqueness of the true outcome variables τ_x and $\tau_{x'}$, which is defined in the following remark.

Remark 5.34 [$P^{Z=z}$-Uniqueness of a True Total Effect Variable] Let the Assumptions 5.1 (a) to (f) hold. Then

$$\tau_x \textit{ is } P^{Z=z}\textit{-unique} \quad :\Leftrightarrow \quad \forall\, \tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y|D_X)\colon P^{Z=z}\big(\{\omega \in \Omega\colon \tau_x(\omega) = \tau_x^*(\omega)\}\big) = 1, \tag{5.26}$$

That is, if τ_x is $P^{Z=z}$-unique, then all versions of the D_X-conditional expectation of Y with respect to the probability measure $P^{X=x}$ are identical, $P^{Z=z}$-almost surely.

Also note that $P^{Z=z}$-uniqueness of τ_x is equivalent to $P(X{=}x|D_X) \underset{P^{Z=z}}{>} 0$ (see RS-Th. 5.27), which is defined by

$$P(X{=}x|D_X) \underset{P^{Z=z}}{>} 0 \quad :\Leftrightarrow \quad P^{Z=z}\big(\{\omega \in \Omega\colon P(X{=}x|D_X)(\omega) > 0\}\big) = 1. \tag{5.27}$$

Hence, $P^{Z=z}$-uniqueness of τ_x is equivalent to $P(X{=}x|D_X)$ being positive, $P^{Z=z}$-almost surely. ◁

Remark 5.35 [An Implication of $P^{Z=z}$-Uniqueness of τ_x and $\tau_{x'}$] The conditional expectation value $E(\tau_x - \tau_{x'}|Z{=}z)$ is a uniquely defined number if $P(Z{=}z) > 0$ and we assume $P^{Z=z}$-uniqueness of τ_x and $\tau_{x'}$. In other words, if τ_x, $\tau_{x'}$ are $P^{Z=z}$-unique and $\delta_{xx'}$, $\delta^*_{xx'}$ are different versions of the true total effect variable $CTE_{D_X;xx'}(D_X)$ [see Def. 5.18 (ii)], then

$$E^{Z=z}(\delta_{xx'}) \;=\; E^{Z=z}(\delta^*_{xx'}) \;=\; E(\delta_{xx'}|Z{=}z) \;=\; E(\delta^*_{xx'}|Z{=}z) \tag{5.28}$$

[see RS-Box 5.1 (viii)]. ◁

Remark 5.36 [P-Uniqueness Implies $P^{Z=z}$-Uniqueness] If $P(Z{=}z) > 0$, then P-uniqueness of $\tau_x = E^{X=x}(Y|D_X)$ implies that τ_x is also $P^{Z=z}$-unique [see RS-Box 5.1 (v)]. In contrast, $P^{Z=z}$-uniqueness of τ_x does not imply that it is also P-unique. Hence, $P^{Z=z}$-uniqueness of τ_x is a weaker assumption than P-uniqueness of τ_x. However, there are conditions under which $P^{Z=z}$-uniqueness implies P-uniqueness, so that they are equivalent to each other. ◁

Lemma 5.37 [A Condition Under Which $P^{Z=z}$-Uniqueness Implies P-Uniqueness] *Let the Assumptions 5.1 (a) to (e) hold. Furthermore, assume 5.1 (f) for all $z \in Z(\Omega)$. Finally, let τ_x denote a true outcome variable of Y given the value x. Then*

$$\big(\forall\, z \in Z(\Omega)\colon \tau_x \textit{ is } P^{Z=z}\textit{-unique}\big) \quad \Leftrightarrow \quad \tau_x \textit{ is } P\textit{-unique}. \tag{5.29}$$

(Proof p. 147)

Hence, under the assumptions of Lemma 5.37, τ_x is $P^{Z=z}$-unique for *all values z of Z* if and only if τ_x is P-unique. However, note that P-uniqueness of τ_x is also defined if Z is continuous and $P(Z{=}z) = 0$, for all $z \in Z(\Omega)$. In this case, Proposition (5.29) does not hold.

Definition 5.38 [Causal Conditional Total Effect Function]
Let the Assumptions 5.1 (a) to (e) and (g) hold, let τ_x and $\tau_{x'}$ denote true outcome variables of Y given the value x and x', respectively, and let $CTE_{Z;xx'}: (\Omega'_Z, \mathscr{A}'_Z) \to (\mathbb{R}, \mathscr{B})$ be a measurable function.

(i) If τ_x and $\tau_{x'}$ are P-unique, then the composition

$$CTE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'} \,|\, Z) \tag{5.30}$$

is called a version of the causal Z-conditional total effect variable comparing x to x' (with respect to Y), and $CTE_{Z;xx'}$ is called a version of the causal Z-conditional total effect function comparing x to x' (with respect to Y).

(ii) Let the Assumptions 5.1 (a) to (g) hold and assume that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique. Then

$$CTE_{Z;xx'}(z) \;=\; E(\tau_x - \tau_{x'} \,|\, Z{=}z). \tag{5.31}$$

is called the causal $(Z{=}z)$-conditional total effect on Y comparing x to x'.

Under the assumptions of Definition 5.38 (ii), $CTE_{Z;xx'}(z)$ is uniquely defined. Furthermore, if τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique *for all $z \in Z(\Omega)$*, then, according to Lemma 5.37, they are also P-unique and

$$\forall\, z \in Z(\Omega)\colon\; CTE_{Z;xx'}(Z)(\omega) = CTE_{Z;xx'}(z), \quad \text{if } \omega \in \{Z{=}z\}. \tag{5.32}$$

That is, if τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique for all $z \in Z(\Omega)$, then the causal $(Z{=}z)$-conditional total effects $CTE_{Z;xx'}(z)$ are the uniquely defined values of the causal Z-conditional total effect function $CTE_{Z;xx'}$.

Remark 5.39 [A Characterization of $CTE_{Z;xx'}(Z)$] Lemma 3.16 and RS-Proposition (4.8) imply that $CTE_{Z;xx'}(Z)$ is Z-measurable and a version of the Z-conditional expectation of $\tau_x - \tau_{x'}$, that is,

$$CTE_{Z;xx'}(Z) \;\in\; \mathscr{E}(\tau_x - \tau_{x'} \,|\, Z). \tag{5.33}$$

$\lhd$

Remark 5.40 [$CTE_{Z;xx'}(Z)$ **Versus** $CTE_{Z;xx'}$] While the composition $CTE_{Z;xx'}(Z)$ is a random variable on the probability space $(\Omega, \mathscr{A}, P)$, which assigns values to all $\omega \in \Omega$, the function $CTE_{Z;xx'}$ is a random variable on $(\Omega'_Z, \mathscr{A}'_Z, P_Z)$, which assigns values to all $z \in \Omega'_Z$. It is the *factorization* of the conditional expectation $CTE_{Z;xx'}(Z)$ (see RS-sect. 4.3). $\lhd$

Remark 5.41 [Conditioning on D_X Versus Conditioning on Z] So far, we considered two kinds of random variables on which we condition. The first is D_X, a global potential confounder of X. Such a variable is essential for the theory and in particular for translating the ceteris paribus clause into the language of probability theory. It has been used to define a true outcome variable $\tau_x = E^{X=x}(Y|D_X)$.

Estimating the true outcome variables τ_x and $\tau_{x'}$ or their difference $\delta_{xx'}$ requires strong assumptions. Considering a causal Z-conditional total effect variable, we re-aggregate the true total effect variable $\tau_x - \tau_{x'}$ [see Eq. (5.30)]. This yields a less fine-grained or coarsened total effect variable, but it is still a *causal* conditional total effect variable. Considering $E(\tau_x - \tau_{x'}|Z)$ instead of $\tau_x - \tau_{x'}$ itself, we may loose information, but we do not loose causal interpretability. In contrast to a true total effect variable $\tau_x - \tau_{x'}$, a Z-conditional total effect variable $E(\tau_x - \tau_{x'}|Z)$ can often be identified under realistic assumptions by empirically estimable conditional expectations (see, e. g., chs. 6 to 11). ◁

Remark 5.42 [$E(\tau_x|Z)$ Versus $E^{X=x}(Y|Z)$] Note the distinction between the two Z-conditional expectations $E(\tau_x|Z)$ and $E(\tau_{x'}|Z)$ on one side and the two Z-conditional expectations $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ on the other side. The difference between the first two conditional expectations is a causal Z-conditional total effect variable, that is,

$$CTE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'}|Z) \underset{P}{=} E(\tau_x|Z) - E(\tau_{x'}|Z). \tag{5.34}$$

In contrast, the conditional expectations $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$, and their difference have no causal meaning unless they are unbiased (see Def. 6.13). The difference $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$ is just a *Z-conditional prima facie effect variable*, which can be seriously misleading if erroneously interpreted as a causal conditional total effect variable. ◁

Remark 5.43 [Coarsening the True Total Effect Variable $\tau_x - \tau_{x'}$] With $E(\tau_x|Z)$ we coarsen (or re-aggregate) the true outcome variable $\tau_x = E^{X=x}(Y|D_X)$. Conditioning on the global potential confounder D_X we control for all potential confounders of X. Therefore the conditional expectations $E^{X=x}(Y|D_X)$ and $E^{X=x'}(Y|D_X)$ inform us how Y depends on the values x and x' *controlling for all potential confounders of X*. Hence, considering the Z-conditional expectation of the difference variable $\tau_x - \tau_{x'}$ does not introduce bias. As already stated in Remark 5.41, it just coarsens the most fine-grained total effects to causal total effects that are less fine-grained. In contrast, considering the conditional expectations $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ and their difference, we only control for Z, possibly neglecting important potential confounders. In chapter 6 we will define $E^{X=x}(Y|Z)$ to be *unbiased* or *biased* depending on whether or not $E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z)$ (see again Def. 6.13). ◁

Remark 5.44 [Values of a Causal Conditional Total Effect Variable] Note that the composition $CTE_{Z;xx'}(Z)$ of the conditional total effect function $CTE_{Z;xx'}$ and Z is a random variable on $(\Omega, \mathscr{A}, P)$, and according to RS-Equation (4.18),

$$\begin{aligned} \forall\, \omega \in \Omega: \quad CTE_{Z;xx'}(Z)(\omega) &= E(\tau_x - \tau_{x'}|Z)(\omega) \\ &= E(\tau_x - \tau_{x'}|Z{=}z) \;=\; CTE_{Z;xx'}(z), \quad \text{if } \omega \in \{Z{=}z\}. \end{aligned} \tag{5.35}$$

Hence, if z is a value of Z such that the assumptions of Definition 5.38 (ii) are satisfied, then $CTE_{Z;xx'}(z)$ is uniquely defined and it is identical to the $(Z{=}z)$-conditional expectation value of $\tau_x - \tau_{x'}$ given the event $\{Z{=}z\} = \{\omega \in \Omega: Z(\omega) = z\}$. Furthermore, according to RS-Equation (3.35),

$$CTE_{Z;xx'}(z) \;=\; E(\tau_x - \tau_{x'}|Z{=}z) \;=\; E(\tau_x|Z{=}z) - E(\tau_{x'}|Z{=}z). \tag{5.36}$$

That is, a value of the causal conditional total effect variable $CTE_{Z;xx'}(Z)$ is the difference between the $(Z{=}z)$-conditional expectation values of τ_x and $\tau_{x'}$. $\triangleleft$

Remark 5.45 [Average Total Effect With Respect to $P^{Z=z}$] In Definition 5.38 (ii) it is assumed that $P(Z{=}z) > 0$ and that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique. Therefore, the causal $(Z{=}z)$-conditional total effect $CTE_{Z;xx'}(z)$ on Y comparing x to x' is identical to the causal average total effect on Y comparing x to x' with respect to the measure $P^{Z=z}$. That is,

$$CTE_{Z;xx'}(z) \;=\; E(\tau_x - \tau_{x'} | Z{=}z) \;=\; E^{Z=z}(\tau_x - \tau_{x'}). \tag{5.37}$$

$\triangleleft$

Remark 5.46 [Causal Conditional Versus Causal Average Total Effects] A causal conditional total effect variable is more informative than the causal average total effect. If the values z of Z are pretest scores that assess the 'same' attribute (e. g., *life satisfaction*) as the outcome variable Y (the post-test), but prior to the onset of the treatment, then comparing the conditional total effects $CTE_{Z;xx'}(z)$ and $CTE_{Z;xx'}(z')$ shows if these conditional total effects are different for different values z and z' of this pretest. If they are, then the numbers $CTE_{Z;xx'}(z)$ and $CTE_{Z;xx'}(z')$ may inform us about the differential indication of the treatment. That is, they answer questions such as "Which treatment is good for which kind of persons?" $\triangleleft$

5.4.2 *Causal* $(X{=}x^*)$*-Conditional Total Effect*

A special case of a $(Z{=}z)$-conditional effect is the $(X{=}x^*)$-conditional effect comparing x to x'. In this case, the random variable$X$ does not only play the role of the focused putative cause variable, but also of the variable on whose values we condition. Note that x^* can be identical to x, x', or to a third value of X. For $Z{=}X$ and $z{=}x^*$, Definition 5.38 (ii) yields

$$CTE_{X;xx'}(x^*) \;=\; E(\tau_x - \tau_{x'} | X{=}x^*), \tag{5.38}$$

the *causal* $(X{=}x^*)$*-conditional total effect* on Y comparing x to x'.

Remark 5.47 [Substantive Meaning] Suppose X represents a treatment variable in an experiment or in a quasi-experiment. If there are two treatment conditions, treatment $(X{=}1)$ and control $(X{=}0)$, then we may consider $CTE_{X;10}(1)$, the $(X{=}1)$-conditional total effect comparing treatment $(X{=}1)$ to control $(X{=}0)$, and $CTE_{X;10}(0)$, the $(X{=}0)$-conditional total effect comparing treatment $(X{=}1)$ to control $(X{=}0)$. These effects are also known as the 'average effect on the treated' and the 'average effect on the untreated', respectively. $\triangleleft$

Remark 5.48 [Pre-Facto Perspective] At first sight, the concept of an $(X{=}x^*)$*-conditional total effect* comparing x to x' seems strange. If, for example, x' represents 'no treatment', how can we talk about the causal average (or conditional) total treatment effect on the *untreated*? However, remember that we are not talking about data that resulted from an experiment — an interpretation that is suggested by the term 'treatment effect on the untreated'. Instead, we are considering a random experiment *that is still to be conducted*, that is, we look at the random experiment from the *pre facto perspective*. This is what probabilistic theories are about: a random experiment that is not yet conducted. Talking about the probability of an event does not make sense for an event that already occurred, unless we *do as if* it did not yet occur, that is, unless we take the pre facto perspective. Hence, we can talk about a causal individual total effect although the individual is not yet treated and even if it will never be treated, just in the same way as we can talk about the probability of flipping 'heads', even if the coin is never flipped. ◁

Remark 5.49 [Causal $(X{=}x^*)$-Conditional Total Treatment Effects] Causal conditional total effects given a specific value x^* of the treatment variable X are often more informative than the causal average total effect, especially, if the X-conditional expectations of the true outcome variables τ_x and $\tau_{x'}$ actually depend on the values of X. However, if τ_x and $\tau_{x'}$ are mean-independent from X, that is, if

$$E(\tau_x|X) \underset{P}{=} E(\tau_x) \quad \text{and} \quad E(\tau_{x'}|X) \underset{P}{=} E(\tau_{x'}), \tag{5.39}$$

then

$$CTE_{X;xx'}(X) \underset{P}{=} E(\tau_x - \tau_{x'}|X) \underset{P}{=} E(\tau_x|X) - E(\tau_{x'}|X) \underset{P}{=} E(\tau_x) - E(\tau_{x'}) = ATE_{xx'}. \tag{5.40}$$

Hence, if Proposition (5.39) holds, then the causal $(X{=}x^*)$-conditional total treatment effects $CTE_{X;xx'}(x^*)$ are identical for all values x^* of X for which $P(X{=}x^*) > 0$. A sufficient condition of Proposition (5.39) is stochastic independence of X and the global potential confounder D_X (see Exercise 5-12), a condition that is created in the randomized experiment. (For more details see ch. 8.) ◁

Remark 5.50 [$CTE_{X;xx'}(x)$ Versus $CTE_{X;xx'}(x')$] Suppose we are interested in the effects of a treatment (represented by $X{=}x$) compared to a control (represented by $X{=}x'$) with respect to the outcome variable Y, say *well-being*, and assume that there is no random assignment of persons to treatments. In this case, the persons that tend to take the treatment may differ in their well-being before treatment and in other pre-treatment variables from those who tend to be in the control condition. In this case, there might be large differences between the causal $(X{=}x)$-conditional total effect $CTE_{X;xx'}(x)$ compared to the causal $(X{=}x')$-conditional total effect $CTE_{X;xx'}(x')$. In this scenario the causal average total effect $ATE_{xx'}$ would not be of much interest. The causal $(X{=}x)$-conditional effect $CTE_{X;xx'}(x)$ helps us evaluating how good the treatment is on average for those who tend to take this treatment. In contrast, $CTE_{X;xx'}(x')$ informs us about the average effect of the treatment on those who tend not to take the treatment — under the side conditions under which the random experiment is to be conducted. Hence, if the causal conditional total effect $CTE_{X;xx'}(x)$ is smaller than the causal conditional total effect $CTE_{X;xx'}(x')$, one may raise the question whether or not it would be worthwhile to change the regime of assigning units to treatment, provided, of course, that this regime is under our control. ◁

Remark 5.51 [Multivariate Random Variable Z] Also note that the concept of a causal $(Z{=}z)$-conditional total effect is not restricted to a univariate random variable Z. Instead, $Z = (Z_1, \dots, Z_m)$ may also be an m-variate random variable on $(\Omega, \mathscr{A}, P)$ such that a value $z = (z_1, \dots, z_m)$ of Z is an m-tupel of values of the random variables $Z_1, \dots, Z_m$. ◁

5.4.3 Complete Re-Aggregation

In this section we consider *complete re-aggregation* of conditional effects, that is, we consider the *expectation* of a causal total effect variable $CTE_{Z;xx'}$. According to Theorem 5.52, this expectation is identical to the causal average total effect on Y comparing x to x'.

Theorem 5.52 [Complete Re-Aggregation of a Conditional Total Effect Variable] *Let the Assumptions 5.1 (a) to (e) and (g) hold, and let $CTE_{Z;xx'}(Z)$ denote a causal Z-conditional total effect variable comparing x to x′ with respect to Y. Then*

$$E\big(CTE_{Z;xx'}(Z)\big) \;=\; ATE_{xx'}\,. \tag{5.41}$$

(Proof p. 147)

Remark 5.53 [Expectation With Respect to the Distribution of Z] According to Theorem 5.52, the expectation of a causal Z-conditional total effect variable is identical to the average total effect. To emphasize, the expectation is with respect to the distribution of Z, that is,

$$E\big(CTE_{Z;xx'}(Z)\big) \;=\; E_Z(CTE_{Z;xx'}) \;=\; ATE_{xx'}, \tag{5.42}$$

which immediately follows from RS-Equation (3.11). Taking this expectation means to re-aggregate the $(Z{=}z)$-conditional total effects to a single number, the causal average total effect. ◁

Remark 5.54 [The Proper Way of Re-Aggregation] Inserting the definition of $CTE_{Z;xx'}(Z)$ [see Eq. (5.30)] into Equation (5.41) yields

$$\begin{aligned}
ATE_{xx'} &= E\big(CTE_{Z;xx'}(Z)\big) && [(5.41)]\\
&= E\big(E(\tau_x - \tau_{x'} | Z)\big) && [(5.30)]\\
&= E(\tau_x - \tau_{x'}) && [\text{RS-Box 4.1 (iv)}]\\
&= E(\tau_x) - E(\tau_{x'}). && [\text{RS-Box 3.1 (ix)}]
\end{aligned} \tag{5.43}$$

Hence, complete re-aggregation actually yields the average causal effect, that is, the expectation of the true effect function $\tau_x - \tau_{x'}$ [see Def. 5.18 (i)].

If Y is binary, then this implies that we take the expectation of the D_X-conditional probabilities

$$P^{X=x}(Y{=}1|D_X) = E^{X=x}(Y|D_X) = \tau_x \quad \text{and} \quad P^{X=x'}(Y{=}1|D_X) = E^{X=x'}(Y|D_X) = \tau_{x'},$$

and not of their log odds ratios or other transformations of these probabilities. In general, a re-aggregation of such transformed probabilities does not yield the causal average total effect. In fact, the expectation of log odds ratios may result in negative 'effects' although the causal average effect is positive and vice versa (see Exercise 5-13). To emphasize, if Y is an indicator variable with values 0 and 1, then Equations (5.43) yield

$$ATE_{xx'} \;=\; E\big(P^{X=x}(Y{=}1|D_X)\big) \;-\; E\big(P^{X=x'}(Y{=}1|D_X)\big). \tag{5.44}$$

◁

5.4.4 Partial Re-Aggregation

Now we turn to a less rigorous re-aggregation of a causal total effect variable $CTE_{Z;xx'}(Z)$, considering a W-conditional expectation of $CTE_{Z;xx'}(Z)$, which, according to Theorem 5.55, is a causal W-conditional total effect variable $CTE_{W;xx'}(W)$, provided that W is Z-measurable, that is, provided that $\sigma(W) \subset \sigma(Z)$. Furthermore, the conditional expectation value $E\big(CTE_{Z;xx'}(Z)\,\big|\,W{=}w\big)$ is identical to the causal $(W{=}w)$-conditional total effect on Y comparing x to x', if we assume $P(W{=}w) > 0$.

Theorem 5.55 [Partial Re-Aggregation of a Conditional Total Effect Variable]
Let the assumptions of Theorem 5.52 hold and assume that W is a Z-measurable random variable on $(\Omega, \mathscr{A}, P)$.

(i) *Then*

$$E\big(CTE_{Z;xx'}(Z)\,\big|\,W\big) \underset{P}{=} CTE_{W;xx'}(W). \tag{5.45}$$

(ii) *If* $w \in W(\Omega)$ *is a value of W such that* $P(W{=}w) > 0$, *then*

$$E\big(CTE_{Z;xx'}(Z)\,\big|\,W{=}w\big) = CTE_{W;xx'}(w). \tag{5.46}$$

(Proof p. 147)

Remark 5.56 [Partial Re-Aggregation] According to Theorem 5.55 (i), the W-conditional expectation of a causal Z-conditional total effect variable is P-almost surely identical to a causal W-conditional total effect variable, provided that W is Z-measurable. If W is Z-measurable, then $\sigma(W) \subset \sigma(Z)$, and if also $\sigma(W) \neq \sigma(Z)$, then the conditional expectation $E\big(CTE_{Z;xx'}(Z)\,\big|\,W\big)$ may be called a *partial re-aggregation* of the original causal total effect variable $CTE_{Z;xx'}(Z)$. It is tantamount to *coarsening* the original causal total effect variable $CTE_{Z;xx'}(Z)$ to a less fine-grained causal total effect variable $CTE_{W;xx'}(W)$. ◁

Remark 5.57 [The Proper Way of Partial Re-Aggregation] Inserting the definition of the causal conditional total effect variable $CTE_{Z;xx'}(Z)$ [see Eq. (5.30)] into Equation (5.45) yields

$$\begin{aligned} CTE_{W;xx'}(W) &\underset{P}{=} E\big(CTE_{Z;xx'}(Z)\,\big|\,W\big) && [(5.45)]\\ &\underset{P}{=} E\big(E(\tau_x - \tau_{x'}|Z)\,\big|\,W\big) && [(5.30)]\\ &\underset{P}{=} E(\tau_x - \tau_{x'}|W) && [\text{RS-Box 4.1 (xiii)}]\\ &\underset{P}{=} E(\tau_x|W) - E(\tau_{x'}|W). && [\text{RS-Box 4.1 (xviii)}] \end{aligned} \tag{5.47}$$

Hence, partial re-aggregation actually yields the causal W-conditional effect variable, that is, the W-conditional expectation of the true effect variable $\tau_x - \tau_{x'}$ [see Def. 5.18 (i)].

If Y is binary, then we take a W-conditional expectation of Z-conditional probabilities

$$P^{X=x}(Y{=}1|D_X) = E^{X=x}(Y|D_X) = \tau_x \quad \text{and} \quad P^{X=x'}(Y{=}1|D_X) = E^{X=x'}(Y|D_X) = \tau_{x'},$$

and *not of their log odds ratios* or other transformations of these probabilities. In general, a partial re-aggregation of such transformed probabilities does not yield a causal W-conditional total effect variable. ◁

5.5 Example: Joe and Ann With Bias at the Individual Level

Now we illustrate the various causal total effects by an example in which there are two treatment variables. Such a two-factorial experiment has already been discussed at an informal level in section 2.3 and at a structural level in section 4.2.2. In contrast to the example in section 4.2.2, now the outcome variable is not binary any more. In this example, we also exemplify that bias can occur at the individual level if we just control for the person variable but not for the second treatment variable that is simultaneous to the focused putative cause variable.

The Random Experiment and the Probability Measure

The parameters displayed in Table 5.2 refer to a random experiment in which

(a) a person is sampled from the set $\Omega_U = \{Joe, Ann\}$ of persons,
(b) the sampled person may (yes) or may not (no) receive treatment a (represented by Z) and, at the same time, receive (yes) or not (no) receive treatment b (represented by X).
(c) After an appropriate time a real-valued outcome variable Y (e. g., *well-being*, *life satisfaction*, or a score on a *symptom checklist*) is observed.

The *set of possible outcomes* of this random experiment is

$$\Omega \;=\; \Omega_1 \times \Omega_2 \times \Omega_3 \;=\; \Omega_U \times (\Omega_Z \times \Omega_X) \times \Omega_Y,$$

where

$$\Omega_1 \;=\; \Omega_U \;:=\; \{Joe, Ann\},$$

$$\Omega_2 \;=\; \Omega_Z \times \Omega_X \;:=\; \{(no, no), (no, yes), (yes, no), (yes, yes)\},$$

and Ω_Y is the set of possible observations based on which the score of the outcome variable Y is computed.

If $\Omega_3 = \Omega_Y$ is finite or countably infinite, then we may choose $\mathscr{A} = \mathscr{P}(\Omega)$ as the σ-algebra on Ω, where $\mathscr{P}(\Omega)$ denotes the power set of Ω. However, if $\Omega_Y = \mathbb{R}$, then $\mathscr{A}$ is the product σ-algebra $\mathscr{A} = \mathscr{P}(\Omega_U) \otimes \mathscr{P}(\Omega_Z \times \Omega_X) \otimes \mathscr{B}$, where $\mathscr{B}$ denotes the Borel σ-algebra on $\mathbb{R}$ (see RS-Rem. 1.14).

The *probability measure P on* $(\Omega, \mathscr{A})$ is known only in that part which can be computed from the parameters displayed in Table 5.2. For example, the conditional probabilities for the two kinds of treatments are as follows: Joe receives treatment a $(Z{=}1)$ with probability $P(Z{=}1 \mid U{=}Joe) = 1/2$ and he receives treatment b $(X{=}1)$ with probability $P(X{=}1 \mid U{=}Joe, Z{=}0) = 3/4$ if he does not receive treatment a $(Z{=}0)$, and with probability $P(X{=}1 \mid U{=}Joe, Z{=}1) = 1/4$ if he receives treatment a $(Z{=}1)$ as well. Similarly, Ann receives treatment b $(X{=}1)$ with probability $P(X{=}1 \mid U{=}Ann, Z{=}0) = 3/4$ if she does not receive treatment a $(Z{=}0)$, and with probability $P(X{=}1 \mid U{=}Ann, Z{=}1) = 1/4$ if she also receives treatment a $(Z{=}1)$. This is a realistic scenario if the probability of a person getting treatment b is fixed by design depending on whether or not this person receives treatment a. (Availability of resources might be a reason for such a design.)

Specifying the Filtration in $\mathscr{A}$

We specify the filtration $(\mathscr{F}_t)_{t \in T}$, $T = \{1, 2, 3\}$, as follows: $\mathscr{F}_1 = \sigma(\pi_1) = \sigma(U)$, $\mathscr{F}_2 = \sigma(\pi_1, \pi_2) = \sigma(U, Z, X)$, and $\mathscr{F}_3 = \sigma(\pi_1, \pi_2, \pi_3) = \sigma(U, Z, X, Y)$, presuming that the two treatment variables X and Z are simultaneous. Of course, treatments are processes in time and they could be represented by a more fine-grained filtration that would allow us to study dependencies within the treatment process. However, for our present purpose the filtration $(\mathscr{F}_t)_{t \in T}$ specified above will suffice.

Table 5.2. Joe and Ann with bias at the individual level

Person u	Group therapy z	$P(U=u)$	$P(Z=z\,|\,U=u)$	$P(X=1\,|\,U=u, Z=z)$	$E^{X=0}(Y\,|\,U=u, Z=z)$	$E^{X=1}(Y\,|\,U=u, Z=z)$	$P^{X=0}(Z=z\,|\,U=u)$	$P^{X=1}(Z=z\,|\,U=u)$
Joe	0	1/2	1/2	3/4	68	82	1/4	3/4
	1		1/2	1/4	96	100	3/4	1/4
Ann	0	1/2	1/2	3/4	80	98	1/4	3/4
	1		1/2	1/4	104	106	3/4	1/4

Focused Treatment Variable and Potential Confounders

There are *several* true total causal effects we might look at. In principle, we might be interested in

(a_1) the causal individual total effect on Y of *treatment a* ($Z=1$) compared to *not treatment a* ($Z=0$) given that Joe (Ann) also receives treatment b ($X=1$),
(b_1) the causal individual total effect on Y of treatment a ($Z=1$) compared to *not treatment a* ($Z=0$) given that Joe (Ann) does *not* receive treatment b ($X=0$), and
(c_1) the average of these causal individual total effects, averaging over the two values of X (representing *treatment b* and *not treatment b*, respectively).

Of course, the causal average effect is certainly less informative than the two conditional effects.

Similarly, we may also be interested in

(a_2) the causal individual total effect on Y of *treatment b* ($X=1$) compared to *not treatment b* ($X=0$) given that Joe (Ann) also receives treatment a ($Z=1$),
(b_2) the causal individual total effect of *treatment b* ($X=1$) compared to *not treatment b* ($X=0$) on Y given that Joe (Ann) does *not* receive *treatment a* ($Z=0$), and
(c_2) the average of these causal individual total effects, averaging over the two values of Z (representing *treatment a* and *not treatment a*, respectively).

Looking at the effects (a_1) to (c_1), we consider *treatment b* to be a (qualitative) covariate and *treatment a* to be the putative cause variable asking for the causal conditional effects of *treatment a* given *treatment b* and their average, the 'main effect' of *treatment a*. In contrast, looking at the effects (a_2) to (c_2), we consider *treatment a* to be a (qualitative) covariate and *treatment b* to be the putative cause variable.

In principle, both treatment variables, X and Z, may take the role of a covariate (and potential confounder), *depending on which treatment effects we are studying*, the causal effects of X on Y or the causal effects of Z on Y. In this example, we focus on X as a putative cause variable of Y, that is, we consider the regular probabilistic causality setup

$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$. In this setup, the index sets in Definition 4.11 are $J = \{1,2\}$ and $K = \{2\}$, implying

$$\mathscr{C} \;=\; \sigma(\pi_{2j}, j \in K) \;=\; \sigma(\pi_{22}) \;=\; \sigma(X)$$

and

$$\mathscr{D}_{\mathscr{C}} \;=\; \sigma(\pi_{2j}, j \in J \setminus K) \;=\; \sigma(\pi_1, \pi_{21}) \;=\; \sigma(U, Z)\,.$$

In this setup, Z is a *potential confounder of* X because $\sigma(Z) \subset \mathscr{D}_{\mathscr{C}}$ [see Def. 4.11 (iv)]. Furthermore, the bivariate random variable (U, Z) is a *global potential confounder of* X because $\sigma(U, Z) = \sigma(\pi_1, \pi_{21})$ [see Def. 4.11 (iii)].

True Outcome Variables and True Total Effects

As stated above, $D_X = (U, Z)$ is a global potential confounder of X. Hence, choosing X as putative cause, the *true outcome variables* are

$$\tau_0 \;=\; E^{X=0}(Y \,|\, D_X) \;=\; E^{X=0}(Y \,|\, U, Z) \quad \text{and} \quad \tau_1 \;=\; E^{X=1}(Y \,|\, D_X) \;=\; E^{X=1}(Y \,|\, U, Z). \tag{5.48}$$

The values of τ_0 and τ_1, denoted $E^{X=0}(Y|U{=}u, Z{=}z)$ and $E^{X=1}(Y|U{=}u, Z{=}z)$, are displayed in Table 5.2. According to RS-Equation (5.26),

$$E^{X=0}(Y|U{=}u, Z{=}z) \;=\; E(Y|X{=}0, U{=}u, Z{=}z)$$

and

$$E^{X=1}(Y|U{=}u, Z{=}z) \;=\; E(Y|X{=}1, U{=}u, Z{=}z).$$

The *true total effect variable* is

$$CTE_{(U,Z);10}(U, Z) \;=\; \tau_1 - \tau_0\,.$$

It is (U, Z)-measurable. According to Table 5.2, its values are as follows: *If Joe does not receive treatment a* $(Z{=}0)$, then his causal total effect of treatment b compared to $\neg b$ is

$$\begin{aligned} CTE_{(U,Z);10}(\mathit{Joe}, 0) \;&=\; E(Y|X{=}1, U{=}\mathit{Joe}, Z{=}0) \;-\; E(Y|X{=}0, U{=}\mathit{Joe}, Z{=}0) \\ &= 82 \;-\; 68 \;=\; 14. \end{aligned} \tag{5.49}$$

In contrast, *if he does receive treatment a* $(Z{=}1)$, then it is

$$\begin{aligned} CTE_{(U,Z);10}(\mathit{Joe}, 1) \;&=\; E(Y|X{=}1, U{=}\mathit{Joe}, Z{=}1) \;-\; E(Y|X{=}0, U{=}\mathit{Joe}, Z{=}1) \\ &= 100 \;-\; 96 \;=\; 4. \end{aligned} \tag{5.50}$$

Similarly, Ann's total effect of treatment b compared to $\neg b$ is

$$\begin{aligned} CTE_{(U,Z);10}(\mathit{Ann}, 0) \;&=\; E(Y|X{=}1, U{=}\mathit{Ann}, Z{=}0) \;-\; E(Y|X{=}0, U{=}\mathit{Ann}, Z{=}0) \\ &= 98 \;-\; 80 \;=\; 18, \end{aligned} \tag{5.51}$$

if she does not receive treatment a $(Z{=}0)$, whereas it is

$$\begin{aligned} CTE_{(U,Z);10}(\mathit{Ann}, 1) \;&=\; E(Y|X{=}1, U{=}\mathit{Ann}, Z{=}1) \;-\; E(Y|X{=}0, U{=}\mathit{Ann}, Z{=}1) \\ &= 106 \;-\; 104 \;=\; 2, \end{aligned} \tag{5.52}$$

if she does $(Z=1)$. Hence, all these $(U=u, Z=z)$-conditional total effects are positive, and they are the true total effects on Y comparing treatment b to treatment $\neg b$ [see Def. 5.18 (iii)]. However, in this example, these effects *are not* identical to the individual effects $CTE_{U;10}(u)$, as will be shown later on in this section. The individual total effect $CTE_{U;10}(u)$ is an attribute of the person u, whereas the true total effect $CTE_{(U,Z);10}(u,z)$ is an attribute of the person u in treatment z.

Causal Average Total Effect

Using the true total effects [see Eqs. (5.49) to (5.52)], the *causal average total effect* of *treatment b* $(X=1)$ compared to $\neg b$ $(X=0)$ can be computed by

$$E\big(CTE_{(U,Z);10}(U,Z)\big) = \sum_u \sum_z CTE_{(U,Z);10}(u,z) \cdot P(U=u, Z=z)$$
$$= 14 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} + 18 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} = 9.5.$$

Because $CTE_{(U,Z);10}(U,Z)$ denotes the composition of (U,Z) and the true total effect function $CTE_{(U,Z);10}$ [see Def. 5.38 (i)], in these computations we used RS-Equation (3.13), and

$$P(U=u, Z=z) = P(Z=z \mid U=u) \cdot P(U=u) = 1/2 \cdot 1/2 = 1/4,$$

for all values (u,z) of the global potential confounder (U,Z). Note that in other examples the probabilities $P(U=u, Z=z)$ may not be identical for all pairs (u,z) of values of U and Z.

$(U=u)$-Conditional Prima Facie Effects

Now we compute the $(U=u)$-conditional prima facie effects, that is, the differences

$$E(Y \mid X=1, U=Joe) - E(Y \mid X=0, U=Joe)$$

and

$$E(Y \mid X=1, U=Ann) - E(Y \mid X=0, U=Ann).$$

These differences may also be called the *individual* or *person-specific prima facie effects*. They are not identical to the *causal* individual total effects, that is, they are not the values of the causal U-conditional total effect function, which will be computed in the next subsection. Hence, in this example, these individual prima facie effects are *biased* (see ch. 6 for more details).

In order to compute the conditional expectation values of Y given treatment and unit, we use the equation

$$E(Y \mid X=x, U=u) = \sum_z E(Y \mid X=x, U=u, Z=z) \cdot P(Z=z \mid X=x, U=u), \tag{5.53}$$

which is always true if Z is discrete with $P(X=x, U=u, Z=z) > 0$ for all values of Z [see RS-Box 3.2 (ii)]. Both kinds of parameters occurring on the right-hand side of this equation are displayed in Table 5.2. This does not only include the conditional expectation values $E(Y \mid X=x, U=u, Z=z) = E^{X=x}(Y \mid U=u, Z=z)$, but also the conditional probabilities $P(Z=z \mid X=x, U=u) = P^{X=x}(Z=z \mid U=u)$, which have been computed via:

$$P(Z{=}z\,|\,X{=}x, U{=}u) \;=\; \frac{P(X{=}x\,|\,U{=}u, Z{=}z)\cdot P(Z{=}z\,|\,U{=}u)}{\sum_z P(X{=}x\,|\,U{=}u, Z{=}z)\cdot P(Z{=}z\,|\,U{=}u)} \tag{5.54}$$

(see Exercises 5-14 and 5-15).

Hence, using Equation (5.53), the $(X{=}x, U{=}Joe)$-conditional expectation values for Joe are

$$E(Y\,|\,X{=}0, U{=}Joe) \;=\; 68\cdot\frac{1}{4} + 96\cdot\frac{3}{4} \;=\; 89\,,$$

$$E(Y\,|\,X{=}1, U{=}Joe) \;=\; 82\cdot\frac{3}{4} + 100\cdot\frac{1}{4} \;=\; 86.5\,,$$

and his individual prima facie effect is

$$E(Y\,|\,X{=}1, U{=}Joe) \;-\; E(Y\,|\,X{=}0, U{=}Joe) \;=\; 86.5 \;-\; 89 \;=\; -2.5\,. \tag{5.55}$$

In this example, the individual prima facie effect of *treatment b* compared to *not treatment b* is negative, namely −2.5, although all $(U{=}Joe, Z{=}z)$-conditional effects are positive, namely 14 for $U{=}Joe$ and $Z{=}0$ (i. e., given *Joe and not treatment a*) and 4 for $U{=}Joe$ and $Z{=}1$ (i. e., given *Joe and treatment a*).

Similarly, using Equation (5.53), the $(X{=}x, U{=}Ann)$-conditional expectation values for Ann are

$$E(Y\,|\,X{=}0, U{=}Ann) \;=\; 80\cdot\frac{1}{4} + 104\cdot\frac{3}{4} \;=\; 98\,,$$

$$E(Y\,|\,X{=}1, U{=}Ann) \;=\; 98\cdot\frac{3}{4} + 106\cdot\frac{1}{4} \;=\; 100\,,$$

and her individual prima facie effect of X on Y is

$$E(Y\,|\,X{=}1, U{=}Ann) \;-\; E(Y\,|\,X{=}0, U{=}Ann) \;=\; 100 \;-\; 98 \;=\; 2\,. \tag{5.56}$$

This prima facie effect does not have a causal interpretation. It is not identical to the causal $(U{=}Ann)$-conditional total effect of X on Y, which is computed in the following subsection.

Causal $(U{=}u)$-Conditional Total Effects

Because $D_X = (U, Z)$ is a global potential confounder in the example presented in Table 5.2, for Joe the *causal $(U{=}u)$-conditional (or individual) total effects* of *treatment b* $(X{=}1)$ compared to *not treatment b* $(X{=}0)$ can be computed by

$$\begin{aligned}
&CTE_{U;10}(Joe) \;=\; E\big(CTE_{(U,Z);10}(U,Z)\,|\,U{=}Joe\big)\\
&= E\big(E^{X=1}(Y\,|\,U,Z) - E^{X=0}(Y\,|\,U,Z)\,\big|\,U{=}Joe\big)\\
&= \sum_u\sum_z \big(E(Y\,|\,X{=}1, U{=}u, Z{=}z) - E(Y\,|\,X{=}0, U{=}u, Z{=}z)\big)\cdot P(U{=}u, Z{=}z\,|\,U{=}Joe)\\
&= (82-68)\cdot\frac{1}{2} + (100-96)\cdot\frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot 0 \;=\; 9
\end{aligned}$$

[see Eqs. (5.31), (5.48), and RS-Eq. (3.28)]. This is an example of partial re-aggregation (see Th. 5.55). For Ann, the corresponding causal total individual effect is

$$
\begin{aligned}
&CTE_{U;10}(Ann) \;=\; E\big(CTE_{(U,Z);10}(U,Z)\,|\,U{=}Ann\big)\\
&= E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\,\big|\,U{=}Ann\big)\\
&= \sum_u \sum_z \big(E(Y|X{=}1,U{=}u,Z{=}z) - E(Y|X{=}0,U{=}u,Z{=}z)\big)\cdot P(U{=}u,Z{=}z\,|\,U{=}Ann)\\
&= (82-68)\cdot 0 + (100-96)\cdot 0 + (98-80)\cdot \frac{1}{2} + (106-104)\cdot \frac{1}{2} \;=\; 10.
\end{aligned}
$$

Hence, in this example, the two causal individual total effects for Joe and Ann are both positive.

Comparing the causal individual effect $CTE_{U;10}(Joe) = 9$ to the corresponding prima facie effect -2.5 [see Eq. (5.55)] shows that the individual prima facie effect $E(Y\,|\,X{=}1, U{=}\,Joe)$ $-\,E(Y\,|\,X{=}0, U{=}\,Joe)$ strongly differs from its causal counterpart, and the same applies to the individual prima facie effect of Ann. This is evident if we compare her prima facie effect $E(Y\,|\,X{=}1, U{=}Ann) - E(Y\,|\,X{=}0, U{=}Ann) = 2$ to her causal individual effect $CTE_{U;10}(Ann)$ $= 10$.

According to Equation (5.41), the expectation of these causal individual effects,

$$
\begin{aligned}
E\big(CTE_{U;10}(U)\big) \;&=\; E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\,\big|\,U\big)\\
&= \sum_u CTE_{U;10}(u)\cdot P(U{=}u) \;=\; 9\cdot\frac{1}{2} \;+\; 10\cdot\frac{1}{2} \;=\; 9.5,
\end{aligned}
$$

is the causal average total effect ATE_{10} of *treatment b* compared to *not treatment b*.

Causal individual total effects are more informative than the causal average total effect and usually more informative than causal conditional total effects given a value of a pre-test or a second treatment variable. However, note again that causal individual [i. e., $(U{=}u)$-conditional] total effects are not necessarily the most fine-grained causal total effects. In this example, there is a second treatment variable, denoted Z, that contributes to the variation of the outcome variable Y beyond the individual level. This is exemplified comparing the causal $(U{=}u, Z{=}z)$-conditional total effects to the causal $(U{=}u)$-conditional total effects.

Causal $(Z{=}z)$-Conditional Total Effects

In the example presented in Table 5.2, the causal $(Z{=}0)$-conditional (i. e., given *not treatment a*) total effect of *treatment b* $(X{=}1)$ compared to *not treatment b* $(X{=}0)$ can be computed by

$$
\begin{aligned}
&CTE_{Z;10}(0) \;=\; E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,Z{=}0\big)\\
&= E\big(E^{X=1}(Y|U,Z) - E^{X=0}(Y|U,Z)\,\big|\,Z{=}0\big)\\
&= \sum_u \sum_z \big(E(Y|X{=}1,U{=}u,Z{=}z) - E(Y|X{=}0,U{=}u,Z{=}z)\big)\cdot P(U{=}u,Z{=}z\,|\,Z{=}0)\\
&= (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16
\end{aligned}
$$

[see Eqs. (5.31), (5.48), and RS-Eq. (3.28)]. The corresponding causal $(Z{=}1)$-conditional (i. e., given *treatment a*) total effect is

$$CTE_{Z;10}(1) \;=\; E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,Z{=}1\big)$$

$$= E\big(E^{X=1}(Y\,|\,U,Z) - E^{X=0}(Y\,|\,U,Z)\,\big|\,Z{=}1\big)$$

$$= \sum_u \sum_z \big(E(Y\,|\,X{=}1,U{=}u,Z{=}z) - E(Y\,|\,X{=}0,U{=}u,Z{=}z)\big)\cdot P(U{=}u,Z{=}z\,|\,Z{=}1)$$

$$= (82-68)\cdot 0 + (100-96)\cdot \frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot \frac{1}{2} \;=\; 3.$$

According to Equation (5.41), taking the expectation

$$E\big(CTE_{Z;10}(Z)\big) \;=\; \sum_z E\big(CTE_{Z;10}(z)\,\big|\,Z{=}z\big)\cdot P(Z{=}z) \;=\; 16\cdot\frac{1}{2} \;+\; 3\cdot\frac{1}{2} \;=\; 9.5 \tag{5.57}$$

again yields the average total effect. In this equation, we used the theorem of total probability in order to compute $P(Z{=}z) = \sum_u P(Z{=}z\,|\,U{=}u)\cdot P(U{=}u)$ (see RS-Th. 1.38), which yields $P(Z{=}0) = P(Z{=}1) = 1/2$.

Causal $(X{=}x)$-Conditional Total Effects

Consider again the example presented in Table 5.2. Because $D_X = (U,Z)$, the causal $(X{=}0)$-conditional total effect of *treatment b* $(X{=}1)$ compared to *not treatment b* $(X{=}0)$ can be computed by

$$CTE_{X;10}(0) \;=\; E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}0\big)$$

$$= \sum_u \sum_z \big(E(Y\,|\,X{=}1,U{=}u,Z{=}z) - E(Y\,|\,X{=}0,U{=}u,Z{=}z)\big)\cdot P(U{=}u,Z{=}z\,|\,X{=}0)$$

$$= (82-68)\cdot\frac{1}{8} + (100-96)\cdot\frac{3}{8} + (98-80)\cdot\frac{1}{8} + (106-104)\cdot\frac{3}{8} \;=\; 6.25$$

[see again Eqs. (5.31), (5.48), and RS-Eq. (3.28)].

In contrast,

$$CTE_{X;10}(1) \;=\; E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}1\big)$$

$$= \sum_u \sum_z \big(E(Y\,|\,X{=}1,U{=}u,Z{=}z) - E(Y\,|\,X{=}0,U{=}u,Z{=}z)\big)\cdot P(U{=}u,Z{=}z\,|\,X{=}1)$$

$$= (82-68)\cdot\frac{3}{8} + (100-96)\cdot\frac{1}{8} + (98-80)\cdot\frac{3}{8} + (106-104)\cdot\frac{1}{8} \;=\; 12.75$$

yields the $(X{=}1)$-conditional total effect of *treatment b* $(X{=}1)$ compared to *not treatment b* $(X{=}0)$. In these equations, we used

$$P(U{=}u,Z{=}z\,|\,X{=}x) \;=\; \frac{P(X{=}x\,|\,U{=}u,Z{=}z)\cdot P(U{=}u,Z{=}z)}{P(X{=}x)}. \tag{5.58}$$

According to Equation (5.41), taking the expectation

$$E\big(CTE_{X;10}(X)\big) \;=\; \sum_x CTE_{X;10}(x)\cdot P(X{=}x) \;=\; 6.25\cdot\frac{1}{2} \;+\; 12.75\cdot\frac{1}{2} \;=\; 9.5$$

yields the causal average total effect. In this equation, we again used the theorem of total probability, that is, $P(X{=}x) = \sum_u \sum_z P(X{=}x\,|\,U{=}u,Z{=}z)\cdot P(U{=}u,Z{=}z)$ (see RS-Th. 1.38), which yields $P(X{=}0) = P(X{=}1) = 1/2$.

Causal $(X{=}x, Z{=}z)$-Conditional Total Effects

Because, in the example presented in Table 5.2, (U, Z) is a global potential confounder, the causal conditional total effect $CTE_{(X,Z);10}(x,z)$ of *treatment b* $(X{=}1)$ compared to *not treatment b* $(X{=}0)$ *given* $X{=}x$ and $Z{=}z$ can be computed by

$$
\begin{aligned}
&CTE_{(X,Z);10}(x,z)\\
&= E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}x, Z{=}z\big) \;=\; E(\delta_{10}\,|\,X{=}x, Z{=}z) \qquad (5.59)\\
&= \sum_u \sum_{z^*} \big(E^{X=1}(Y\,|\,U{=}u, Z{=}z^*) - E^{X=0}(Y\,|\,U{=}u, Z{=}z^*)\big)\cdot P(U{=}u, Z{=}z^*\,|\,X{=}x, Z{=}z)
\end{aligned}
$$

[see Eqs. (5.31), (5.48), and RS-Eq. (3.28)]. In this equation, the summation is over all values u of U and all values z^* of Z. In the term $P(U{=}u, Z{=}z^*|X{=}x, Z{=}z)$ we condition on a fixed value x of X and a fixed value z of Z. The conditional probabilities $P(U{=}u, Z{=}z^*|X{=}x, Z{=}z)$ can be computed via

$$
P(U{=}u, Z{=}z^*|X{=}x, Z{=}z) \;=\; \begin{cases} P(U{=}u|X{=}x, Z{=}z), & \text{if } z^*{=}\,z \\ 0, & \text{if } z^*{\neq}\, z, \end{cases} \qquad (5.60)
$$

because, if $z = z^*$, then

$$
\begin{aligned}
P(U{=}u, Z{=}z^*|X{=}x, Z{=}z) &= \frac{P(U{=}u, Z{=}z^*, X{=}x, Z{=}z)}{P(X{=}x, Z{=}z)}\\
&= \frac{P(U{=}u, X{=}x, Z{=}z)}{P(X{=}x, Z{=}z)} \;=\; P(U{=}u\,|\,X{=}x, Z{=}z) \qquad (5.61)\\
&= \frac{P(X{=}x\,|\,U{=}u, Z{=}z)\cdot P(U{=}u, Z{=}z)}{P(X{=}x\,|\,Z{=}z)\cdot P(Z{=}z)}\,,
\end{aligned}
$$

where

$$
P(X{=}x\,|\,Z{=}z) \;=\; \sum_u P(X{=}x|U{=}u, Z{=}z)\cdot P(U{=}u\,|\,Z{=}z)\,, \qquad (5.62)
$$

with $P(U{=}u\,|\,Z{=}z) = P(Z{=}z\,|\,U{=}u)\cdot P(U{=}u)/P(Z{=}z)$. In this example, $P(U{=}u\,|\,Z{=}z) = 1/2$, for all values of U and Z. Hence, Equation (5.62) yields

$$
P(X{=}0\,|\,Z{=}0) \;=\; 1/4\cdot 1/2 + 1/4\cdot 1/2 \;=\; 1/4,
$$

and using Equation (5.60) we receive:

$$
P(U{=}u, Z{=}0\,|\,X{=}0, Z{=}0) \;=\; \frac{1/4\cdot 1/4}{1/4\cdot 1/2} \;=\; \frac{1}{2},
$$

for $u{=}$*Joe* and for $u{=}$*Ann*. Hence, the equation for $E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}x, Z{=}z\big)$ yields

$$
\begin{aligned}
&E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}0, Z{=}0\big) \;=\; E(\delta_{10}\,|\,X{=}0, Z{=}0)\\
&= (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16.
\end{aligned}
$$

For $X{=}1$ and $Z{=}0$, we receive

$$E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}1,Z{=}0\big) \;=\; E(\delta_{10}\,|\,X{=}1,Z{=}0)$$
$$= (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16,$$

for $X{=}0$ and $Z{=}1$, we receive

$$E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}0,Z{=}1\big) \;=\; E(\delta_{10}\,|\,X{=}0,Z{=}1)$$
$$= (82-68)\cdot 0 + (100-96)\cdot\frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot\frac{1}{2} \;=\; 3,$$

and for $X{=}1$ and $Z{=}1$,

$$E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X{=}1,Z{=}1\big) \;=\; E(\delta_{10}\,|\,X{=}1,Z{=}1)$$
$$= (82-68)\cdot 0 + (100-96)\cdot\frac{1}{2} + (98-80)\cdot 0 + (106-104)\cdot\frac{1}{2} \;=\; 3.$$

Hence, in this example, the conditional total effects $E(\delta_{10}\,|\,Z{=}z)$ and $E(\delta_{10}\,|\,X{=}x, Z{=}z)$ are identical.

According to Equation (5.41), taking the expectation

$$E\Big(E\big(CTE_{(U,Z);10}(U,Z)\,\big|\,X,Z\big)\Big) \;=\; E\big(E(\delta_{10}\,|\,X,Z)\big)$$
$$= \sum_x \sum_z E(\delta_{10}\,|\,X{=}x, Z{=}z)\cdot P(X{=}x, Z{=}z)$$
$$= 16\cdot\frac{1}{4} \;+\; 3\cdot\frac{1}{4} \;+\; 16\cdot\frac{1}{4} \;+\; 3\cdot\frac{1}{4} \;=\; 9.5\,,$$

yields again the causal average total effect ATE_{10}. In this equation, we used $P(X{=}x, Z{=}z) = P(X{=}x\,|\,Z{=}z)\cdot P(Z{=}z)$, where the conditional probabilities $P(X{=}x\,|\,Z{=}z)$ are obtained via Equation (5.62).

5.6 Summary and Conclusions

In this chapter, we introduced the concept of a true outcome variable of the value x of a putative cause variable X, which was then used to define various causal total effects. Assuming $P(X{=}x) > 0$, a true outcome variable τ_x has been defined such that its values are the conditional expectation values $E(Y\,|\,X{=}x, D_X{=}d)$ of the outcome variable Y holding constant X at the value x and D_X at a value d, where D_X denotes a global potential confounder of X. These conditional expectation values are uniquely defined if $P(X{=}x, D_X{=}d) > 0$. Note that this requirement is not necessary in the definition of a true outcome variable itself. Also note, in this definition we only consider *total effects*.

Based on the concept of a true outcome variable $\tau_x = E^{X=x}(Y\,|\,D_X)$, we defined several kinds of causal total effects of treatment x compared to another treatment x' using the *true total effect variable*

$$\delta_{xx'} \;=\; \tau_x \;-\; \tau_{x'}\,.$$

The definitions of a *causal average total effect* $ATE_{xx'}$ and of a *causal Z-conditional total effect variable* $CTE_{Z;xx'}(Z)$ (see Box 5.1) are based on the assumption that the two true

Box 5.1 Glossary of new concepts

Let $\big((\Omega,\mathscr{A},P),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big)$ be a regular probabilistic causality setup, let D_X denote a global potential confounder of X, let Y be real-valued with $Var(Y)>0$, let $x\in X(\Omega)$ be a value of X, and assume $P(X{=}x)>0$.

τ_x	*True outcome variable of Y given the value x of X*. It is a version of the D_X-conditional expectation of Y with respect to the probability measure $P^{X=x}$, that is, $\tau_x = E^{X=x}(Y\,\vert\,D_X)$. This is a random variable on $(\Omega,\mathscr{A},P)$ that is defined only if $P(X{=}x)>0$. With a global potential confounder D_X, we condition on *all* potential confounders of X.

Additionally, let also $x'\in\Omega'_X$, assume $P(X{=}x')>0$, and let τ_x and $\tau_{x'}$ denote true outcome variables of Y given the values x and x' of X, respectively.

$ATE_{xx'}$	*Causal average total effect comparing x to x′*. If τ_x and $\tau_{x'}$ are P-unique, then it is defined by $$ATE_{xx'} \;:=\; E(\tau_x-\tau_{x'}).$$

Additionally, let also Z be a random variable on $(\Omega,\mathscr{A},P)$.

$CTE_{Z;xx'}(Z)$	*Causal Z-conditional total effect variable comparing x to x′*. If τ_x and $\tau_{x'}$ are P-unique, then $$CTE_{Z;xx'}(Z) \;:=\; E(\tau_x-\tau_{x'}\,\vert\,Z).$$ It is the composition of Z and $CTE_{Z;xx'}$, and therefore a random variable on $(\Omega,\mathscr{A},P)$.
$CTE_{Z;xx'}$	*Causal Z-conditional total effect function comparing x to x′*. It is a factorization of $E(\tau_x-\tau_{x'}\,\vert\,Z)$. Hence, it is a random variable on $(\Omega'_Z,\mathscr{A}'_Z,P_Z)$.
$CTE_{Z;xx'}(z)$	*Causal $(Z{=}z)$-conditional total effect comparing x to x′*. If $P(Z{=}z)>0$ and τ_x, $\tau_{x'}$ are $P^{Z=z}$-unique, then $$CTE_{Z;xx'}(z) \;:=\; E(\tau_x-\tau_{x'}\,\vert\,Z{=}z).$$
$CTE_{U;xx'}(u)$	*Causal individual total effect comparing x to x′ for unit u*.
$CTE_{X;xx'}(x^*)$	*Causal $(X{=}x^*)$-conditional total effect comparing x to x′*.
$CTE_{(X,Z);xx'}(x^*,z)$	*Causal $(X{=}x^*,Z{=}z)$-conditional total effect comparing x to x′*.

outcome variables τ_x and $\tau_{x'}$ are P-unique. Defining the *causal* $(Z{=}z)$*-conditional total effect* $CTE_{Z;xx'}(z)$ we only assume that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique. The term 'total' is used in order to distinguish these effects from direct and indirect effects, which are not considered in this volume.

While D_X is a global potential confounder of X on which we condition in order to control for all potential confounders of X, the variable Z may be used to re-aggregate the $(D_X{=}d)$-conditional total effects in order to consider less fine-grained causal conditional total effects. Examples of Z are the observational-unit variable U, a pre-treatment variable Z, and a treatment variable X.

Causal Average Total Effect

Often we have to content ourselves with the causal average total effect or with causal conditional total effects. Note however, that there might be cases in which half of the units have positive causal individual total effects and the other half negative ones. The causal average total effect can then be zero. This is not a paradox but the nature of an average. Also remember that a causal average total effect is informative for causal inference, whereas an ordinary true mean difference $E(Y \mid X=1) - E(Y \mid X=0)$, the prima facie effect, is not. These prima facie effects have no causal interpretation at all, unless they are identical to the causal average total effect. This will be detailed in chapter 6.

Main Effects Versus Conditional Effects in Analysis of Variance

Conceptually, the *causal average total effect* is what is tested in a t-test for two independent groups, *provided that the data are sampled in a perfect randomized experiment.* Similarly, in this case, a test of the *main effect* of the 'treatment factor' in orthogonal analysis of variance is a simultaneous test of several causal average total effects if there are more than two treatment conditions. Furthermore, if Z is a qualitative covariate of X, then it is considered a second 'factor' in analysis of variance. In this case, the $(Z=z)$-conditional total effects are often called the 'simple main effects' (see, e. g., Woodward & Bonett, 1991).

Note that causal average total effects are uniquely defined even if there are inter-individual differences in the causal individual total effects, and even if there is interaction between X and a covariate Z of X in the sense that the effect of X depends on the values of Z. However, only in the randomized experiment can we be sure that, with the main effects in analysis of variance, we test the causal average total effects.

Of course, the causal conditional effects given the values of a covariate are usually more informative than their average, that is, than the causal average total effect; but sometimes averaging is useful in order to avoid information overload, and sometimes we may be able to estimate precisely enough only the causal average effect, but not the causal conditional effects, for example, because of small sample sizes.

Pre-Facto Versus Post-Facto or 'Counterfactual' Perspective

Note that our definitions of the various kinds of total effects solely use concepts of probability theory. No concepts have to be borrowed from philosophy or any other science, although the basic idea goes back at least to Mill (1843/1865). We do not take a counterfactual but a *pre-facto perspective*, which is the perspective taken in *every* application of probability theory. Causal total effects are parameters, just in the same way as the probability of flipping 'heads' is a parameter about which we can talk *before* the coin is flipped and even if the coin is *never* flipped. It is even meaningful to talk about the *causal conditional total effect of a treatment given control* [see Eq. (5.38) and Rem. 5.47].

Theoretical Parameters Versus Data

Note that all concepts introduced in this chapter such as the *causal average total effects*, *causal conditional total effects*, *causal individual total effects*, and so on, are of a purely theoretical nature. This does not mean that they are irrelevant for practical research. On the contrary, they explicate what exactly we are looking for when we ask for the causal total effects, for example, comparing two values of a treatment variable with respect to an outcome variable Y. It is these effects that we have to estimate if we want to evaluate (a) if and how dangerous an infection is, and (b) the overall effects of medical, psychological, social, or political interventions on a specified criterion. And this includes their (undesired) side effects.

The examples treated in chapters 4 and 5 exemplify what we mean by 'purely theoretical nature'. For example, Table 5.2 does not show data that might be obtained in a data sample. Instead, it contains the theoretical parameters we would like to estimate from sample data. *Data serve to estimate theoretical parameters*, including the various causal effects. Defining these parameters is necessary if we want to study the conditions under which these parameters can in fact be estimated.

Causal Versus Noncausal Theoretical Parameters

Not all theoretical parameters have a causal meaning. In terms of the metaphor presented in the preface, the causal total effects are the *size* of the invisible man. In contrast, in chapter 1, we only dealt with (a) ordinary conditional expectation values $E(Y|X=x)$ of an outcome variable Y given treatment x, (b) conditional expectation values $E(Y|X=x, Z=z)$ of the outcome variable given treatment x and value z of another random variable Z, (c) differences between these (conditional) expectation values, the (conditional) *prima facie effects*, and (d) averages over these conditional prima facie effects.

The conditional expectation values $E(Y|X=x)$ and $E(Y|X=x, Z=z)$ are easily estimated under the usual assumptions made for a sample, such as the assumption of independent and identically distributed observations. However, they are only like the length of the invisible man's shadow; depending on the angle of the sun, they can be seriously biased if mistaken for the size of the invisible man itself.

Limitations

A limitation of the concept of a true outcome variable and the definitions of causal effects based thereon is that they are defined only for values x of a putative cause variable X that have a positive probability. This is not restrictive as long as we confine ourselves to considering only experiments and quasi-experiments. True outcome variables are restrictive, however, if we study causal dependencies among *continuous* random variables and causal dependencies on latent variables. In these cases, the concept of a true outcome variable does not apply. In chapter 8 we will treat a class of causality conditions that do also apply if the putative cause variable X is a continuous random variable. Furthermore, the class of causality conditions presented in chapter 9 also applies if X is continuous or a latent variable. In these cases causal effects and causal dependencies have to be defined without true outcome variables. Nevertheless, true outcome variables are important for many applications.

Another limitation is that we did not consider potential mediators, that is, variables that are between X and Y. However, focussing only on causal total effects does not imply that we denie that there are variables mediating these effects. This is one of the virtues of true outcome theory: Given treatment x and person u, only the *expectation* of the outcome variable Y is fixed, not the value of Y itself. In contrast, in Rubin's potential outcome approach it is assumed that the value of a potential outcome variable is fixed if we condition on a treatment x and a person u. Such a determinism is at odds with the idea that potential mediators might also affect the outcome variable Y. True outcome theory remedies this deficiency. Nevertheless, defining potential mediators, direct, and indirect causal effects would necessitate a filtration $(\mathscr{F}_t)_{t \in T}$ with more than three σ-algebras $\mathscr{F}_t$.

5.7 Proofs

Proof of Lemma 5.37

Define the event $A := \{\omega \in \Omega\colon P(X{=}x \mid D_X)(\omega) > 0\}$. Then

$$
\begin{aligned}
& \forall\, z \in Z(\Omega)\colon \tau_x \text{ is } P^{Z=z}\text{-unique} \\
\Leftrightarrow\quad & \forall\, z \in Z(\Omega)\colon P(X{=}x \mid D_X) \underset{P^{Z=z}}{>} 0 && \text{[RS-Th. 5.27]} \\
\Leftrightarrow\quad & \forall\, z \in Z(\Omega)\colon P^{Z=z}(A) = 1 && \text{[def. of } A\text{, (5.27)]} \\
\Rightarrow\quad & \sum_{z \in Z(\Omega)} P(A \mid Z{=}z) \cdot P(Z{=}z) \;=\; \sum_{z \in Z(\Omega)} 1 \cdot P(Z{=}z) && \text{[RS-(5.1)]} \\
\Rightarrow\quad & P(A) \;=\; 1 && \Big[\text{RS-(1.38)}, \textstyle\sum_z P(Z{=}z) = 1\Big] \\
\Leftrightarrow\quad & P(X{=}x \mid D_X) \underset{P}{>} 0 && \text{[def. of } A\text{, (5.14)]} \\
\Leftrightarrow\quad & \tau_x \text{ is } P\text{-unique.} && \text{[RS-Th. 5.27]}
\end{aligned}
$$

The reverse implication immediately follows from RS-Box 5.1 (v).

Proof of Theorem 5.52

$$
\begin{aligned}
E\big(CTE_{Z;xx'}(Z)\big) &= E\big(E(\tau_x - \tau_{x'} \mid Z)\big) && \text{[(5.30)]} \\
&= E(\tau_x - \tau_{x'}). && \text{[RS-Box 4.1 (iv)]} \\
&= ATE_{xx'}\,. && \text{[Def. 5.26]}
\end{aligned}
$$

Proof of Theorem 5.55

Proposition (i).

$$
\begin{aligned}
E\big(CTE_{Z;xx'}(Z) \,\big|\, W\big) &\underset{P}{=} E\big(E(\tau_x - \tau_{x'} \mid Z) \,\big|\, W\big) && \text{[(5.30)]} \\
&\underset{P}{=} E(\tau_x - \tau_{x'} \mid W) && \text{[RS-Box 4.1 (xiii)]} \\
&\underset{P}{=} CTE_{W;xx'}(W). && \text{[(5.30)]}
\end{aligned}
$$

Proposition (ii).

$$
E\big(CTE_{Z;xx'}(Z) \,\big|\, W\big) \underset{P}{=} CTE_{W;xx'}(W) \qquad \text{[(5.45)]}
$$

$$\Leftrightarrow \quad E\big(CTE_{Z;xx'}(Z)\,\big|\,W\big) \underset{P}{=} E(\tau_x - \tau_{x'}|W) \qquad \text{[Def. 5.38 (i)]}$$
$$\Rightarrow \quad E\big(CTE_{Z;xx'}(Z)\,\big|\,W{=}w\big) = E(\tau_x - \tau_{x'}|W{=}w) \qquad [RS\text{-}(2.68),\ RS\text{-}(4.17)]$$
$$\Leftrightarrow \quad E\big(CTE_{Z;xx'}(Z)\,\big|\,W{=}w\big) = CTE_{W;xx'}(w). \qquad \text{[Def. 5.38 (ii)]}$$

5.8 Exercises

▷ **Exercise 5-1** What is the conceptual framework in which we can define a true outcome variable?

▷ **Exercise 5-2** What does it mean that a true outcome variable τ_x is P-unique.

▷ **Exercise 5-3** Which are the values of the true outcome variable τ_0 and of the conditional expectation $E(Y|X,U)$ for $\omega_4 = (Joe, yes, +)$ in the example presented in Table 5.1?

▷ **Exercise 5-4** Compute the values of the true outcome variable $\tau_0 = E^{X=0}(Y|U)$ in the example presented in RS-Table 2.1.

▷ **Exercise 5-5** Suppose that X is a binary treatment variable and Y an outcome variable. Why are the conditional expectation values $E(Y|X{=}0)$, $E(Y|X{=}1)$, and their difference, the prima facie effect $E(Y|X{=}1) - E(Y|X{=}0)$, often useless in the evaluation of the causal total effect?

▷ **Exercise 5-6** Suppose that X is a treatment variable and Y an outcome variable. If the conditional expectation values $E(Y|X{=}x)$ and their differences $E(Y|X{=}x) - E(Y|X{=}x')$ do not represent the treatment effects we are interested in, then what *are* the treatment effects we would like to study?

▷ **Exercise 5-7** What is the difference between the causal average total effect $ATE_{xx'}$ and the prima facie effect $PFE_{xx'}$?

▷ **Exercise 5-8** What is the *causal conditional total effect* $CTE_{Z;xx'}(z)$ on Y comparing x to x' *given the value* z *of a random variable* Z?

▷ **Exercise 5-9** What is the *causal conditional total effect* $CTE_{X;xx'}(x^*)$ on Y comparing x to x' *given the value* x^* of the putative cause variable X?

▷ **Exercise 5-10** What is the causal conditional total effect $CTE_{(X,Z);xx'}(x^*, z)$ on Y comparing x to x' given treatment x^* and the value z of a random variable Z?

▷ **Exercise 5-11** Use RS-Theorem 1.38 to compute the probability $P(X{=}1)$ for the example displayed in Table 5.2.

▷ **Exercise 5-12** Show that Proposition (5.39) follows from independence of X and D_X.

▷ **Exercise 5-13** Open your RStudio, go to www.causal-effects.de, tools, *Aggregation*[0,1]*Xplorer*, shinyapplication.r, download, open the file, Run App. Then choose the example 'Treatment effect, X and Z are dependent [Inversion of SME and MEM]', and click on 'Compute aggregated effects and visualizations'.

▷ **Exercise 5-14** Compute the probability $P^{X=1}(U{=}Ann, Z{=}0)$ in the example of Table 5.2.

▷ **Exercise 5-15** Compute the conditional probabilities $P^{X=x}(Z{=}z\,|\,U{=}u)$ in Table 5.2.

▷ **Exercise 5-16** Compute the causal average total effect ATE_{10} for the random experiment presented in Table 5.2.

▷ **Exercise 5-17** Compute the causal conditional total effect $CTE_{Z;10}(0)$ given *no group therapy* for the random experiment presented in Table 5.2.

▷ **Exercise 5-18** Let Z represent *sex* with values m (males) and f (females). Furthermore, suppose $CTE_{Z;10}(m) = 11$, $CTE_{Z;10}(f) = 5$, $P(Z{=}m) = 1/3$, and $P(Z{=}f) = 2/3$. Which is the causal average total effect ATE_{10}?

Solutions

▷ **Solution 5-1** *First*, we assume that there is a probability space $(\Omega, \mathscr{A}, P)$. (In an empirical application, this probability space represents the concrete random experiment considered.) *Second*, there are two random variables on $(\Omega, \mathscr{A}, P)$, say X and Y, where X represents the putative cause variable and Y the outcome variable. *Third*, there is a filtration $(\mathscr{F}_t)_{t \in T}$ in $\mathscr{A}$ in which X is prior to Y (see Box 3.1). *Fourth*, we assume that $P(X{=}x) > 0$. *Fifth*, we assume that D_X is a global potential confounder of X. By definition, $\tau_x = E^{X=x}(Y|D_X)$ is $P^{X=x}$-unique. In this definition, we do not assume that the true outcome variable $\tau_x = E^{X=x}(Y|D_X)$ is P-unique (see Exercise 5-2).

▷ **Solution 5-2** By definition of a true outcome variable $\tau_x = E^{X=x}(Y|D_X)$, there may be different versions of a true outcome variable. In general, two such versions τ_x and τ_x^* are identical almost surely with respect to the probability measure $P^{X=x}$. This is what we mean saying that τ_x *is* $P^{X=x}$-*unique* (see Rem. 5.16). If we assume that τ_x is P-unique, then two versions τ_x and τ_x^* are identical almost surely with respect to the probability measure P (see Rem. 5.17).

▷ **Solution 5-3** The value $\tau_0(\omega_4)$ of the true outcome variable τ_0 is $E(Y|X{=}0, U{=}Joe) = .7$. In contrast, the value $E(Y|X,U)(\omega_4)$ of the conditional expectation $E(Y|X,U)$ is $E(Y|X{=}1, U{=}Joe) = .8$ (see the fourth row in Table 5.1).

▷ **Solution 5-4** The values of the true outcome variable $\tau_0 = E^{X=0}(Y|U)$ are the two conditional expectation values $E(Y|X{=}0, U{=}Joe)$ and $E(Y|X{=}0, U{=}Ann)$. Because Y is binary, $E(Y|X{=}0, U{=}u)$ $= P(Y{=}1|X{=}0, U{=}u)$. These conditional probabilities can be computed from the probabilities of the elementary events presented in RS-Table 2.1 as follows:

$$P(Y{=}1|X{=}0, U{=}Joe) \;=\; \frac{P(Y{=}1, X{=}0, U{=}Joe)}{P(X{=}0, U{=}Joe)} \;=\; \frac{.35}{.15+.35} \;=\; .7$$

and

$$P(Y{=}1|X{=}0, U{=}Ann) \;=\; \frac{P(Y{=}1, X{=}0, U{=}Ann)}{P(X{=}0, U{=}Ann)} \;=\; \frac{.06}{.24+.06} \;=\; .2.$$

▷ **Solution 5-5** A potential confounder W of X may determine the probabilities of being treated [i. e., $P(X{=}x|W) \neq P(X{=}x)$, $x \in \{0,1\}$] *and* the $(X{=}x)$-conditional expectation values of the outcome variable Y [i. e., $E^{X=x}(Y|W) \neq E^{X=x}(Y)$, $x \in \{0,1\}$]. In this case, there are examples in which the difference $E(Y|X{=}1) - E(Y|X{=}0)$ is not identical to the treatment effects to be studied. Simpson's paradox presented in chapter 1 is such an example. Another example of such a potential confounder is W = *severity of symptoms*. If there is self-selection or if there is systematic selection to treatment by experts that is also determined by the severity of the symptoms, then W will affect the treatment probability *and* the conditional expectation values of the outcome variable (e. g., *severity of symptoms after treatment*).

▷ **Solution 5-6** The basic idea is to consider the true total effect variable, that is, the difference $\tau_x - \tau_{x'} = E^{X=x}(Y|D_X) - E^{X=x'}(Y|D_X)$, where we condition on a global potential confounder D_X of X. This means controlling for *all* potential confounders. If we take the expectation of the difference $\tau_x - \tau_{x'}$ (over the distribution of D_X), then this yields the causal average total effect on Y comparing x to x'. Note that taking this expectation necessitates to assume that τ_x and $\tau_{x'}$ are P-unique.

This assumption implies that the expectation $E(\tau_x - \tau_{x'})$ is identical for all versions of τ_x and $\tau_{x'}$ (see Remarks. 5.27 to 5.29).

▷ **Solution 5-7** The causal average total effect $ATE_{xx'}$ comparing treatment x to treatment x' has been defined by Equation (5.21) (see also the solution to Exercise 5-6). It is this causal average total effect that is of interest in the empirical sciences if our goal is to evaluate the treatment conditions x and x' with respect to the outcome variable Y by a single number. In contrast, the prima facie effect $PFE_{xx'}$ comparing x to x' is usually not of interest for the evaluation of such a treatment effect because it can be biased. Both terms differ from each other because $PFE_{xx'} = E(Y|X{=}x) - E(Y|X{=}x')$ is not neccesarily identical to $ATE_{xx'} = E(\tau_x) - E(\tau_{x'})$. Note, however, that *there are* conditions under which $PFE_{xx'} = ATE_{xx'}$. Such conditions, which are called *causality conditions*, are studied in some detail in the next chapters.

▷ **Solution 5-8** The causal conditional total effect (on the outcome variable Y) comparing x to x' given the value z of a random variable Z is the $(Z{=}z)$-conditional expectation value of the true total effect variable $\delta_{xx'} = \tau_x - \tau_{x'}$, that is,

$$CTE_{Z;xx'}(z) \;=\; E(\delta_{xx'}|Z{=}z).$$

It is presumed that $P(Z{=}z) > 0$ and that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique. These assumptions imply that $CTE_{Z;xx'}(z)$ is a uniquely defined number.

▷ **Solution 5-9** The causal conditional total effect (on Y) comparing x to x' given the value x^* of the putative cause variable X is the $(X{=}x^*)$-conditional expectation value of $\delta_{xx'} = \tau_x - \tau_{x'}$, that is,

$$CTE_{X;xx'}(x^*) \;=\; E(\delta_{xx'}|X{=}x^*),$$

where we presume $P(X{=}x^*) > 0$ and that τ_x and $\tau_{x'}$ are $P^{X=x^*}$-unique. If X represents a treatment variable, $x^*{=}x$, and $x'{=}0$ represents a control group, then $CTE_{X;x0}(x)$ is the causal conditional total effect comparing treatment x to control *given treatment* x. If $x^*{=}x'$ and $x'{=}0$ represents a control group, then $CTE_{X;x0}(0)$ is the causal conditional total effect comparing treatment x to control *given control*. Although this sounds paradoxical, the term $CTE_{X;x0}(0)$ is meaningful and well-defined. Like all other causal effects it refers to a random experiment to be conducted *in the future*. This means that these concepts are well-defined even if the experiment is not yet conducted, or will never be conducted (see sect. 5.4.2 for more details).

▷ **Solution 5-10** The causal conditional total effect $CTE_{(X,Z);xx'}(x^*, z)$ (on Y) comparing x to x' given treatment x^* and value z of Z is the $(X{=}x^*, Z{=}z)$-conditional expectation value of the true-effect variable, that is,

$$CTE_{(X,Z);xx'}(x^*, z) \;=\; E(\delta_{xx'}|X{=}x^*, Z{=}z),$$

where we presume $P\big((X,Z){=}(x^*,z)\big) > 0$ and that τ_x and $\tau_{x'}$ are $P^{X=x^*,Z=z}$-unique. If X represents a treatment variable, $x^*{=}x$, the value 0 of X represents a control group, and m (male) the value of Z = *sex*, then $CTE_{(X,Z);x0}(x, m)$ is the causal conditional total effect comparing treatment x to control given treatment x and the person to be sampled is male. If $x^*{=}0$, then $CTE_{(X,Z);x0}(0, m)$ is the causal conditional total effect comparing treatment x to control given control and the person to be sampled is male.

▷ **Solution 5-11** Note that the four pairs (u, z) of values of U and Z are disjoint and all these pairs of values have positive probabilities. Hence, we can apply the theorem of total probability [see RS-Eq. (1.38)]:

$$\begin{aligned} P(X{=}1) \;&=\; \sum_u \sum_z P(X{=}1|U{=}u, Z{=}z) \cdot P(U{=}u, Z{=}z) \\ &=\; \left(\frac{3}{4} + \frac{1}{4} + \frac{3}{4} + \frac{1}{4}\right) \cdot \frac{1}{4} \;=\; \frac{1}{2}. \end{aligned}$$

Also note that $P(X{=}1) = E(1_{X=1}) = E[E(1_{X=1}\,|\,U,Z)]$ [see RS-Box 4.1 (iv)], and that the probabilities $P(X{=}1\,|\,U{=}u, Z{=}z)$ are the values of the conditional expectation $E(1_{X=1}\,|\,U,Z)$. Then using RS-Equation (3.13) yields the same formula. This second way makes clear that the unconditional probability $P(X{=}1)$ is the expectation of the conditional probability $P(X{=}1|U,Z)$ [see again RS-Box 4.1 (iv)].

▷ **Solution 5-12** Independence of X and D_X implies that also X and $E^{X=x}(Y|D_X)$ are independent because $E^{X=x}(Y|D_X)$ is D_X-measurable [see RS-Box 2.1 (iv)]. Hence,

$$\begin{aligned}
E(\tau_x|X) &\underset{P}{=} E\big(E^{X=x}(Y|D_X)\,\big|\,X\big) && \big[\tau_x \underset{P}{=} E^{X=x}(Y|D_X),\ \text{RS-Box 4.1 (xiv)}\big]\\
&\underset{P}{=} E\big(E^{X=x}(Y|D_X)\big) && [\text{RS-Box 4.1 (v)}]\\
&= E(\tau_x). && \big[\tau_x \underset{P}{=} E^{X=x}(Y|D_X),\ \text{RS-Box 3.1 (v)}\big]
\end{aligned}$$

The prove for $\tau_{x'}$ is analog.

▷ **Solution 5-13** No solution provided. See what happens with the various techniques of re-aggregating conditional effects, for example, re-aggregating the log odds ratios and compare it to re-aggregating conditional effects according to Equation (5.44). Play with other parameter constellations.

▷ **Solution 5-14** We have to use the equation

$$P^{X=x}(U{=}u, Z{=}z) = \frac{P(X{=}x, U{=}u, Z{=}z)}{P(X{=}x)} = \frac{P(X{=}x\,|\,U{=}u, Z{=}z)\cdot P(U{=}u, Z{=}z)}{P(X{=}x)}$$

[see RS-Eq. (5.1)]. For $U{=}Ann$, $Z{=}0$, and $X{=}1$, this equation yields

$$\begin{aligned}
P^{X=1}(U{=}Ann, Z{=}0) &= \frac{P(X{=}1|U{=}Ann, Z{=}0)\cdot P(U{=}Ann, Z{=}0)}{P(X{=}1)}\\
&= \frac{3/4\cdot 1/4}{1/2} = \frac{3}{8}.
\end{aligned}$$

▷ **Solution 5-15** If $P(X{=}x) > 0$, then, applying RS-Equation (5.1) several times, yields

$$\begin{aligned}
P^{X=x}(Z{=}z\,|\,U{=}u) &= \frac{P^{X=x}(Z{=}z, U{=}u)}{P^{X=x}(U{=}u)}\\
&= \frac{P(Z{=}z, U{=}u, X{=}x)/P(X{=}x)}{P(U{=}u, X{=}x)/P(X{=}x)}\\
&= P(Z{=}z|X{=}x, U{=}u).
\end{aligned}$$

For both units u, the complementary conditional probability to $P(X{=}1|U{=}u, Z{=}0) = 3/4$ (see Table 5.2) is $P(X{=}0\,|\,U{=}u, Z{=}0) = 1/4$. Similarly the complementary conditional probability to $P(X{=}1|U{=}u, Z{=}1) = 1/4$ (see again Table 5.2) is $P(X{=}0\,|\,U{=}u, Z{=}1) = 3/4$. Now we can use the equation

$$P(Z{=}z\,|\,X{=}x, U{=}u) = \frac{P(X{=}x\,|\,U{=}u, Z{=}z)\cdot P(Z{=}z\,|\,U{=}u)}{\sum_z P(X{=}x\,|\,U{=}u, Z{=}z)\cdot P(Z{=}z\,|\,U{=}u)},$$

which follows from Bayes' Theorem (see RS-Th. 1.39) using $P(Z{=}z\,|\,X{=}x, U{=}u) = P^{U=u}(Z{=}z\,|\,X{=}x)$ [see RS-Eq. (5.26)]. For example, the probability of not receiving group therapy ($Z{=}0$), if Ann is drawn ($U{=}Ann$) and does not receive individual therapy ($X{=}0$), is

$$P(Z{=}0\,|\,X{=}0, U{=}Ann) = \frac{P(X{=}0\,|\,U{=}Ann, Z{=}0)\cdot P(Z{=}0\,|\,U{=}Ann)}{\sum_z P(X{=}0\,|\,U{=}Ann, Z{=}z)\cdot P(Z{=}z\,|\,U{=}Ann)}, \tag{5.63}$$

where $P(X{=}0\,|\,U{=}Ann, Z{=}0) = 1/4$, $P(Z{=}0\,|\,U{=}Ann) = 1/2$, and

$$\begin{aligned}
&\sum_z P(X{=}0\,|\,U{=}Ann, Z{=}z)\cdot P(Z{=}z\,|\,U{=}Ann)\\
&= P(X{=}0\,|\,U{=}Ann, Z{=}0)\cdot P(Z{=}0\,|\,U{=}Ann) \;+\; P(X{=}0\,|\,U{=}Ann, Z{=}1)\cdot P(Z{=}1|\,U{=}Ann)\\
&= \frac{1}{4}\cdot\frac{1}{2} \;+\; \frac{3}{4}\cdot\frac{1}{2} \;=\; \frac{1}{2}.
\end{aligned}$$

Inserting this result into Equation (5.63) yields the conditional probability

$$P(Z{=}0\,|\,X{=}0, U{=}Ann) \;=\; P^{X=0}(Z{=}0\,|\,U{=}Ann) \;=\; 1/4.$$

Using the same procedure for all values of U, X, and Z leads to the other conditional probabilities displayed in the last two columns of Table 5.2.

▹ **Solution 5-16** In this example, (U,Z) is a global potential confounder of X. Hence, according to RS-Equation (3.14), Equations (5.20) and (5.21), using the parameters shown in Table 5.2 results in

$$\begin{aligned}
ATE_{10} &= \sum_u\sum_z \big(E(Y\,|\,X{=}1, U{=}u, Z{=}z) - E(Y\,|\,X{=}0, U{=}u, Z{=}z)\big)\cdot P(U{=}u, Z{=}z)\\
&= \big((82-68)+(100-96)+(98-80)+(106-104)\big)\cdot\frac{1}{4} \;=\; 9.5.
\end{aligned}$$

▹ **Solution 5-17** In this example, (U,Z) is a global potential confounder of X. Hence, according to RS-Equation (3.14) and Equation (5.31), $E(\delta_{10}\,|\,Z{=}0) = E^{Z=0}(\delta_{10})$, using the parameters displayed in Table 5.2 results in

$$\begin{aligned}
CTE_{Z;10}(0) &= \sum_u\sum_z \big(E(Y\,|\,X{=}1, U{=}u, Z{=}z) - E(Y\,|\,X{=}0, U{=}u, Z{=}z)\big)\cdot P(U{=}u, Z{=}z\,|\,Z{=}0)\\
&= \sum_u\sum_z \big(E(Y\,|\,X{=}1, U{=}u, Z{=}z) - E(Y\,|\,X{=}0, U{=}u, Z{=}z)\big)\cdot P^{Z=0}(U{=}u, Z{=}z)\\
&= (82-68)\cdot\frac{1}{2} + (100-96)\cdot 0 + (98-80)\cdot\frac{1}{2} + (106-104)\cdot 0 \;=\; 16.
\end{aligned}$$

▹ **Solution 5-18** Using Equation (5.41), we can compute the causal average total effect as follows:

$$ATE_{10} \;=\; CTE_{Z;10}(m)\cdot\frac{1}{3} + CTE_{Z;10}(f)\cdot\frac{2}{3} \;=\; 11\cdot\frac{1}{3} + 5\cdot\frac{2}{3} \;=\; 7.$$

Part III

Causality Conditions

Chapter 6

Unbiasedness and Identification of Causal Effects

In chapter 4, we introduced the concepts of a *regular probabilistic causality setup*, a *potential confounder* or *covariate of* X, and a *global potential confounder of* X. In chapter 5, we turned to the concepts of a *true outcome variable*, a *true total effect variable*, a *causal average total effect*, a *causal conditional total effect variable*, and a *causal conditional total effect*. All these parameters and variables are of a theoretical nature. It is not evident how they can be computed (identified) from parameters of the joint distribution of observable random variables such as X, Y, and a (possibly multivariate) covariate Z. That is, it is not yet evident how causal effects can be computed from those parameters that can be estimated in a data sample.

Tackling this problem, in this chapter we introduce and study *unbiasedness* of various conditional expectation values, conditional expectations, prima facie effects, and prima facie effect functions. In particular, we study how and under which conditions these parameters and random variables can be used to identify the corresponding causal effects and effect functions. Hence, in this chapter we provide the link between causal effects and causal effect functions on one side and parameters and functions that can empirically be estimated on the other side. The *unbiasedness conditions* are the first and logically weakest kind of causality conditions, which, together with the structural components listed in a regular probabilistic causality setup, distinguish *causal* stochastic dependencies from *ordinary* stochastic dependencies that have no causal meaning.

We start with the concept of *unbiasedness* of a conditional expectation value $E(Y|X{=}x)$, unbiasedness of the corresponding prima facie effects, that is, of the differences $E(Y|X{=}x)$ $-$ $E(Y|X{=}x')$, and unbiasedness of a conditional expectation $E(Y|X)$. We also treat a first way to identify the causal average total effect $ATE_{xx'}$. Then we consider an additional random variable Z that is assumed to be a covariate of X, and define unbiasedness of a conditional expectation value $E(Y|X{=}x, Z{=}z)$, unbiasedness of the conditional expectations $E^{X=x}(Y|Z)$, $E^{Z=z}(Y|X)$, and $E(Y|X, Z)$, as well as the prima facie effect variables $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$. We also show how to identify the causal average total effect $ATE_{xx'}$ as well as a causal conditional total effect function $CTE_{Z;xx'}$ and a causal conditional total effect $CTE_{Z;xx'}(z)$. Next, we illustrate these concepts by some numerical examples. Finally, we show that unbiasedness can be accidental, presenting an example in which the conditional expectation values $E(Y|X{=}x)$ are unbiased, whereas the conditional expectation values $E(Y|X{=}x, Z{=}z)$ are not, although Z is a covariate of X. This example emphasizes the limitations of unbiasedness, in particular if compared to the causality conditions presented in chapters 8 and 10.

Requirements

Reading this chapter we assume again that the reader is familiar with the contents of the first five chapters of Steyer (2024), referred to as RS-chapters 1 to 5. Furthermore, we assume familiarity with chapters 4 and 5 of the present book.

In this chapter we often will refer to the following notation and assumptions.

Notation and Assumptions 6.1

(a) *Let* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *be a regular probabilistic causality setup and let* D_X *denote a global potential confounder of* X.

(b) *Let* $(\Omega'_X, \mathscr{A}'_X)$ *denote the value space of* X, *let* $x \in \Omega'_X$, *let* $\{x\} \in \mathscr{A}'_X$, *and let* $1_{X=x}$ *denote the indicator of the event* $\{X=x\} = \{\omega \in \Omega\colon X(\omega) = x\}$.

(c) *Let* Y *be real-valued with positive variance.*

(d) *Assume* $P(X=x) > 0$, *define the probability measure* $P^{X=x}\colon \mathscr{A} \to [0,1]$ *by* $P^{X=x}(A) = P(A \,|\, X=x)$, *for all* $A \in \mathscr{A}$, *let* $\tau_x = E^{X=x}(Y \,|\, D_X)$ *denote a (version of the) true outcome variable of* Y *given* x, *and* $\mathscr{E}^{X=x}(Y \,|\, D_X)$ *the set of all such versions.*

(e) *Let* Z *be a covariate of* X, *that is, let* $\sigma(Z) \subset \sigma(D_X)$, *and let* $(\Omega'_Z, \mathscr{A}'_Z)$ *denote its value space.*

(f) *Let* $z \in \Omega'_Z$ *be a value of* Z, *let* $\{z\} \in \mathscr{A}'_Z$, *assume* $P(Z=z) > 0$, *and define the probability measure* $P^{Z=z}\colon \mathscr{A} \to [0,1]$ *by* $P^{Z=z}(A) = P(A \,|\, Z=z)$, *for all* $A \in \mathscr{A}$.

(g) *Let* $x' \in \Omega'_X$ *and* $\{x'\} \in \mathscr{A}'_X$, *let* $1_{X=x'}$ *denote the indicator of the event* $\{X=x'\} = \{\omega \in \Omega\colon X(\omega) = x'\}$, *assume* $0 < P(X=x') < 1$, *define the measure* $P^{X=x'}$ *analogously to* $P^{X=x}$ *in Assumption (d), let* $\tau_{x'} = E^{X=x'}(Y \,|\, D_X)$ *denote a (version of the) true outcome variable given* x', *and define* $\delta_{xx'} := \tau_x - \tau_{x'}$.

(h) *Let* $X(\Omega) = \{0, 1, \ldots, J\}$ *denote the image of* Ω *under* X, *for all* $x \in X(\Omega)$, *let* $\{x\} \in \mathscr{A}'_X$ *and assume* $0 < P(X=x) < 1$. *Finally, let* $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$ *denote the* $(J+1)$*-variate random variable of the true outcome variables* τ_x, $x \in X(\Omega)$.

6.1 Unbiasedness of $E(Y|X)$ and Its Values $E(Y|X=x)$

Under the Assumptions 6.1 (a) to (d) and (g), and assuming that $\tau_x = E^{X=x}(Y|D_X)$ and $\tau_{x'} = E^{X=x'}(Y|D_X)$ are P-unique, we defined the average total effect by $ATE_{xx'} = E(\delta_{xx'})$ (see sect. 5.3). Inserting the definition of a true total effect variable $\delta_{xx'} = \tau_x - \tau_{x'}$ and the definition of a true outcome variable yields

$$\begin{aligned} ATE_{xx'} &= E(\delta_{xx'}) = E(\tau_x - \tau_{x'}) \\ &= E(\tau_x) - E(\tau_{x'}) = E\big(E^{X=x}(Y \,|\, D_X)\big) - E\big(E^{X=x'}(Y \,|\, D_X)\big). \end{aligned} \tag{6.1}$$

Under the assumptions mentioned above, including P-uniqueness of τ_x and $\tau_{x'}$, all terms in these equations are uniquely defined numbers [see RS-Th. 5.27 (v)].

Remark 6.2 [A Condition Equivalent to P-Uniqueness] Remember, according to RS-Theorem 5.27 (ii), P-uniqueness of τ_x is equivalent to $P(X=x \,|\, D_X) \underset{P}{>} 0$, which is defined by

$$P(X=x \,|\, D_X) \underset{P}{>} 0 \quad :\Leftrightarrow \quad P\big(\{\omega \in \Omega\colon P(X=x \,|\, D_X)(\omega) > 0\}\big) = 1. \tag{6.2}$$

Hence, assuming $P(X=x|D_X) \underset{P}{>} 0$ means that the probability of $P(X=x|D_X)$ taking on a value greater than zero is one. ◁

Definition 6.3 [Unbiasedness of $E(Y|X=x)$ and $E(Y|X)$]
Let the Assumptions 6.1 (a) to (d) hold.

(i) *We call the conditional expectation value $E(Y|X=x)$ unbiased and denote it by $E(Y|X=x) \vdash D_X$ if the following two conditions hold:*

  (a) *τ_x is P-unique*
  (b) *$E(Y|X=x) = E(\tau_x)$.*

(ii) *Let the Assumptions 6.1 (a) to (d) and (h) hold. Then we call the conditional expectation $E(Y|X)$ unbiased, denoted $E(Y|X) \vdash D_X$ if, for all $x \in X(\Omega)$, the conditional expectation values $E(Y|X=x)$ are unbiased.*

Remark 6.4 [Unbiased With Respect to $\mathscr{D}_{\mathscr{C}}$] Note that Definition 6.3 refers to *total effects* true outcome variables. Considering these true outcome variables $\tau_x = E^{X=x}(Y|D_X)$, we condition on a global potential confounder D_X of X, and with it on its generated σ-algebra $\sigma(D_X) = \mathscr{D}_{\mathscr{C}}$ (see RS-Def. 5.4). This is also the reason why $\mathscr{D}_{\mathscr{C}}$ occurs in the shortcuts for unbiasedness. ◁

Example 6.5 [No Treatment for Joe] In the example displayed in RS-Table 2.1, we assume that the regular causality space is the same as specified in Example 4.10, and again, X, Y, and U take the roles of the putative cause variable, the outcome variable, and the global potential confounder D_X. Furthermore, the co-domain Ω'_X of X is any subset of $\mathbb{R}$ (including $\mathbb{R}$ itself) containing the elements 0 and 1. The values of the two true outcome variables

$$\tau_x \;=\; P^{X=x}(Y=1|U) \;=\; E^{X=x}(Y|U), \quad x = 0, 1,$$

are displayed in the table. Whereas τ_0 is the only element of the set $\mathscr{E}^{X=0}(Y|U)$ (which means that τ_0 is uniquely defined and therefore P-unique), τ_1 is not the only element in the set $\mathscr{E}^{X=1}(Y|U)$ because the random variable $\tau_1^* = P^{X=1}(Y=1|U)^*$ displayed in the last column of RS-Table 2.1 is also an element of $\mathscr{E}^{X=1}(Y|U)$. Furthermore, $\tau_1 \underset{P}{=} \tau_1^*$ does not hold. Instead, $\tau_1(\omega) \neq \tau_1^*(\omega)$ for $\omega \in \{\omega_1, \dots, \omega_4\}$ and $P(\{\omega_1, \dots \omega_4\}) = .5$. Hence, in this example, the conditional expectation value $E(Y|X=0)$ is unbiased, whereas the conditions of the definition of unbiasedness neither hold for $E(Y|X=1)$ nor for the conditional expectation $E(Y|X)$ because τ_1 is not P-unique (see Def. 6.3). In fact, in this example, $E(\tau_1) = E\big(P^{X=1}(Y=1|U)\big) \neq E(\tau_1^*) = E\big(P^{X=1}(Y=1|U)^*\big)$, which is easily seen looking at the last two columns of RS-Table 2.1 and the probabilities $P(\{\omega_i\})$ displayed in the first numerical column. ◁

Remark 6.6 [Unbiased Estimators Versus Unbiased Parameters] Unbiasedness in statistics usually refers to *estimators* of a parameter. In the framework of the theory of causal effects, we say that the conditional expectation value $E(Y|X=x)$ is unbiased if it is identical to $E(\tau_x)$, which is the parameter of interest in causal inference. Of course, this parameter has to be uniquely defined. This is secured by requiring P-uniqueness of τ_x [see Def. 6.3 and RS-Th. 5.27 (v)]. ◁

Remark 6.7 [Expectation With Respect to the Measure $P^{X=x}$] If $P(X{=}x) > 0$, then

$$E(Y|X{=}x) \;=\; E^{X=x}(Y) \tag{6.3}$$

[see RS-Eq. (3.24)]. Hence, if $P(X{=}x) > 0$, then the conditional expectation value $E(Y|X{=}x)$ is unbiased if and only if the expectation $E^{X=x}(Y)$ of Y with respect to the conditional probability measure $P^{X=x}$ [see again RS-Eq. (5.1)] is unbiased, that is, if and only if τ_x is P-unique and

$$E^{X=x}(Y) \;=\; E(\tau_x) \;=\; E\big(E^{X=x}(Y|D_X)\big). \tag{6.4}$$

$\lhd$

Remark 6.8 [Identification of $E(\tau_x)$] Unbiasedness of $E(Y|X{=}x)$ is important because it gives us access to the expectation $E(\tau_x)$ of the true outcome variable τ_x. If $E(Y|X{=}x)$ is unbiased, then, according to Definition 6.3 (i), an estimate of $E(Y|X{=}x)$ is also an estimate of $E(\tau_x)$. In contrast to $E(\tau_x)$, the conditional expectation value $E(Y|X{=}x)$ can often be estimated from a data sample of the random variables X and Y. The sample mean of the values of Y that are observed together with the value x of the putative cause variable X is such an estimate of $E(Y|X{=}x)$ (see RS-Exercise 3-4). If X is a treatment variable, then this is the sample mean of the observed values of Y in treatment x. $\lhd$

Equivalent Conditions of Unbiasedness of $E(Y|X)$ and $E(Y|X{=}x)$

In the following theorem, we present three conditions that are equivalent to unbiasedness of $E(Y|X{=}x)$. Each of these conditions involves a true outcome variable τ_x. Reading this theorem, remember

$$\tau_x \vDash 1_{X=x} \quad\Leftrightarrow\quad E(\tau_x|\,1_{X=x}) \underset{P}{=} E(\tau_x) \tag{6.5}$$

[see RS-Prop. (4.35)]. Hence, we read $\tau_x \vDash 1_{X=x}$ as *τ_x is mean-independent from $1_{X=x}$*, the indicator variable of the event $\{X{=}x\} = \{\omega \in \Omega\colon X(\omega) = x\}$ that X takes on the value x.

Theorem 6.9 [Equivalent Conditions of Unbiasedness of $E(Y|X{=}x)$]
Let the Assumptions 6.1 (a) to (d) hold and assume that τ_x is P-unique.

(i) Then each of the following two equations is equivalent to $E(Y|X{=}x) \vdash D_X$.

$$E^{X=x}(\tau_x) \;=\; E(\tau_x) \tag{6.6}$$

$$E(\tau_x|\,1_{X=x}) \underset{P}{=} E(\tau_x). \tag{6.7}$$

(ii) If, additionally,

$$\varepsilon_x \;:=\; \tau_x \;-\; E(\tau_x|\,1_{X=x}), \tag{6.8}$$

then each of the following two equations is also equivalent to $E(Y|X{=}x) \vdash D_X$

$$E^{X=x}(\varepsilon_x) \;=\; E(\varepsilon_x) \tag{6.9}$$

$$E(\varepsilon_x|\,1_{X=x}) \underset{P}{=} E(\varepsilon_x). \tag{6.10}$$

(Proof p. 187)

Remark 6.10 [Sufficient Conditions of Unbiasedness] Later it will be shown that $E^{X=x}(\tau_x) = E(\tau_x)$ as well as $\tau_x \vDash 1_{X=x}$ follow from independence of τ_x and $1_{X=x}$ (see Th. 7.7), which itself follows from $D_X \perp\!\!\!\perp 1_{X=x}$, that is, from independence of a global potential confounder D_X of X and the indicator $1_{X=x}$ [see Th. 8.22 (i) for $X = 1_{X=x}$]. Note that, in an experiment in which X is the treatment variable and the person variable U takes the role of a global potential confounder of X, the independence condition $D_X \perp\!\!\!\perp 1_{X=x}$ can be created by randomized assignment of the observational unit to treatment x (see the examples in RS-Table 1.2 and in Table 6.3). ◁

Remark 6.11 [Dichotomous X] If X is dichotomous and x is one of the two values, then

$$E(\tau_x|X) \underset{P}{=} E(\tau_x|\, 1_{X=x}) \underset{P}{=} E(\tau_x|\, 1_{X\neq x}), \tag{6.11}$$

because, if X is dichotomous, then the σ-algebras generated by X, $1_{X=x}$, and $1_{X\neq x}$ are identical. Remember, the σ-algebras $\sigma(X)$, $\sigma(1_{X=x})$, and $\sigma(1_{X\neq x})$ play a crucial role in the definition of the conditional expectations $E(\tau_x|X)$, $E(\tau_x|\, 1_{X=x})$, and $E(\tau_x|\, 1_{X\neq x})$ (see RS-Def. 4.4). Hence, if X is dichotomous and the conditional expectations $E^{X=x}(Y|D_X)$ and $E^{X\neq x}(Y|D_X)$ are P-unique, then

$$\tau_x \vDash X \quad\Leftrightarrow\quad \tau_x \vDash 1_{X=x} \quad\Leftrightarrow\quad E(Y|X) \vdash D_X, \tag{6.12}$$

where $E(Y|X) \vdash D_X$ denotes *unbiasedness of* $E(Y|X)$. Remember, if X is dichotomous, then unbiasedness of $E(Y|X)$ is defined by unbiasedness of $E(Y|X{=}x)$ and $E(Y|X{\neq}x)$ [see Def. 6.3 (ii)], that is, in this case

$$E(Y|X) \vdash D_X \quad\Leftrightarrow\quad E(Y|X{=}x) \vdash D_X \wedge E(Y|X{\neq}x) \vdash D_X. \tag{6.13}$$

◁

Remark 6.12 [Unbiasedness of $E(Y|X)$ in Quasi-Experiments] In empirical applications in which there is no randomized assignment of the observational unit to one of the treatment conditions, unbiasedness of $E(Y|X)$ or $E(Y|X{=}x)$ is not very likely. However, if we additionally consider a (uni- or multivariate) covariate Z of X and the conditional expectation values $E(Y|X{=}x, Z{=}z)$, then unbiasedness of these parameters and of the conditional expectation $E(Y|X,Z)$ is much more realistic, even beyond experiments with randomized assignment of the unit to a treatment condition. This motivates the following section. ◁

6.2 Unbiasedness of $E(Y|X,Z)$ and Related Terms

Now we extend the concept of unbiasedness to conditioning on a (possibly multivariate) covariate Z of X or on one of its values z. That is, we assume $\sigma(Z) \subset \sigma(D_X)$. Note that $Z := (Z_1,\dots,Z_m)$ is a covariate of X if and only if $Z_1,\dots,Z_m$ are covariates of X (see RS-sect. 2.1).

6.2.1 Definition and First Properties

Remember that a true outcome variable τ_x denotes a version of the D_X-conditional expectation of Y with respect to the probability measure $P^{X=x}$. That is, τ_x denotes an element of the set $\mathscr{E}^{X=x}(Y|D_X)$ (see Def. 5.4 and Rem. 5.5). Furthermore, remember

$$P^{X=x}(Z=z) > 0 \quad \Leftrightarrow \quad P^{Z=z}(X=x) > 0 \quad \Leftrightarrow \quad P(X=x, Z=z) > 0, \tag{6.14}$$

and that $P(X=x, Z=z) > 0$ does not only imply $P^{X=x}(Z=z) > 0$ and $P(Z=z) > 0$, but also

$$E^{X=x}(Y|Z=z) = E^{Z=z}(Y|X=x) = E(Y|X=x, Z=z) \tag{6.15}$$

[see RS-Eq. (5.26)]. All terms appearing in Proposition (6.14) and Equation (6.15) have been introduced in RS-section 5.1.

Furthermore, remember, under the Assumptions 6.1 (a) to (f),

$$\tau_x \textit{ is } P^{Z=z}\textit{-unique} \quad :\Leftrightarrow \quad \forall\, \tau_x, \tau_x^* \in \mathscr{E}^{X=x}(Y|D_X)\colon P^{Z=z}\big(\{\omega \in \Omega\colon \tau_x(\omega) = \tau_x^*(\omega)\}\big) = 1, \tag{6.16}$$

and that this property follows from P-uniqueness of τ_x [see RS-Box 5.1 (v)]. Also note that $P^{Z=z}$-uniqueness of τ_x is equivalent to $P(X=x|D_X) \underset{P^{Z=z}}{>} 0$ (see RS-Th. 5.27), which is defined by

$$P(X=x|D_X) \underset{P^{Z=z}}{>} 0 \quad :\Leftrightarrow \quad P^{Z=z}\big(\{\omega \in \Omega\colon P(X=x|D_X)(\omega) > 0\}\big) = 1. \tag{6.17}$$

If $P(X=x|D_X) \underset{P^{Z=z}}{>} 0$, then we also say that $P(X=x|D_X)$ is positive, $P^{Z=z}$-almost surely.

Definition 6.13 [Unbiasedness of $E^{X=x}(Y|Z)$ and $E^{X=x}(Y|Z=z)$]
Let the Assumptions 6.1 (a) to (e) hold.

(i) *Then $E^{X=x}(Y|Z)$ is called unbiased, denoted $E^{X=x}(Y|Z) \vdash D_X$, if*

  (a) *τ_x is P-unique*
  (b) *$E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z)$.*

(ii) *If we additionally assume 6.1 (f) and $P^{X=x}(Z=z) > 0$, then $E^{X=x}(Y|Z=z)$ is called unbiased, denoted $E^{X=x}(Y|Z=z) \vdash D_X$, if*

  (a) *τ_x is $P^{Z=z}$-unique*
  (b) *$E^{X=x}(Y|Z=z) = E(\tau_x|Z=z)$.*

Under the assumptions of Definition 6.13 (ii), the $(Z=z)$-conditional expectation value $E(\tau_x|Z=z)$ is uniquely defined. Furthermore, if τ_x is $P^{Z=z}$-unique *for all* $z \in Z(\Omega)$, then, according to Lemma 5.37, it is also P-unique and

$$\forall\, z \in Z(\Omega)\colon\; E(\tau_x|Z)(\omega) = E(\tau_x|Z=z), \quad \text{if } \omega \in \{Z=z\}. \tag{6.18}$$

That is, if τ_x is $P^{Z=z}$-unique for all $z \in Z(\Omega)$, then the $(Z=z)$-conditional expectation values $E(\tau_x|Z=z)$ are the uniquely defined values of the conditional expectation $E(\tau_x|Z)$.

Remark 6.14 [Unbiasedness of $E^{Z=z}(Y|X=x)$ and $E(Y|X=x, Z=z)$] Let the assumptions of Definition 6.13 (ii) hold. Then Proposition (6.15) allows us to define

$$E^{Z=z}(Y|X=x) \vdash D_X \quad :\Leftrightarrow \quad E^{X=x}(Y|Z=z) \vdash D_X \tag{6.19}$$

and

$$E(Y|X=x, Z=z) \vdash D_X \quad :\Leftrightarrow \quad E^{X=x}(Y|Z=z) \vdash D_X. \tag{6.20}$$

Hence, unbiasedness of $E^{Z=z}(Y|X=x)$ and of $E(Y|X=x, Z=z)$ are equivalent to unbiasedness of $E^{X=x}(Y|Z=z)$ and to each other. $\triangleleft$

Remark 6.15 [Identification of $E(\tau_x|Z{=}z)$ and $E(\tau_x|Z)$] According to Defniition 6.13 (i), if $E^{X=x}(Y|Z)$ is unbiased, then we can identify the Z-conditional expectation $E(\tau_x|Z)$ by $E^{X=x}(Y|Z)$. Hence, an estimate of $E^{X=x}(Y|Z)$ is also an estimate of $E(\tau_x|Z)$ provided that $E^{X=x}(Y|Z)$ is unbiased. Analogously, if $E^{X=x}(Y|Z{=}z)$ is unbiased, then, according to Definition 6.13 (i), we can identify the $(Z{=}z)$-conditional expectation $E(\tau_x|Z{=}z)$ of a true outcome variable τ_x by the $(Z{=}z)$-conditional expectation $E^{X=x}(Y|Z{=}z)$ of Y with respect to the conditional probability measure $P^{X=x}$. Hence, an estimate of $E^{X=x}(Y|Z{=}z)$ is also an estimate of $E(\tau_x|Z{=}z)$, provided that $E^{X=x}(Y|Z{=}z)$ is unbiased. In contrast to $E(\tau_x|Z)$, the conditional expectation $E^{X=x}(Y|Z)$ can often be estimated from a data sample of the random variables X, Y, and Z using the values of Y and Z observed together with the value x of the putative cause variable X (see Exercise 6-10). ◁

Theorem 6.16 [An Implication of Unbiasedness of $E^{X=x}(Y|Z)$ For $E(Y|X{=}x,Z{=}z)$]
Let the Assumptions 6.1 (a) to (f) hold and assume $P(X{=}x,Z{=}z)>0$. Then

$$E^{X=x}(Y|Z) \vdash D_X \quad \Rightarrow \quad E(Y|X{=}x,Z{=}z) \vdash D_X. \tag{6.21}$$

(Proof p. 188)

In the following theorem we treat a sufficient condition for unbiasedness of a conditional expectation $E^{X=x}(Y|Z)$ with respect to the probability measure $P^{X=x}$.

Theorem 6.17 [Sufficient Condition for Unbiasedness of $E^{X=x}(Y|Z)$]
Let the Assumptions 6.1 (a) to (f) hold and for all $z\in Z(\Omega)$ assume $P^{X=x}(Z{=}z)>0$. Then

$$\forall\, z\in Z(\Omega)\colon\ E^{X=x}(Y|Z{=}z)\vdash D_X \quad \Rightarrow \quad E^{X=x}(Y|Z)\vdash D_X. \tag{6.22}$$

(Proof p. 188)

In the following definition we introduce unbiasedness of $E(Y|X,Z)$ and of $E^{Z=z}(Y|X)$, building on the terms introduced in Definition 6.13. The latter term can be defined only if we assume $P(X{=}x,Z{=}z)>0$ for all values $x\in X(\Omega)$ of X. Among other things, this assumption implies $P(Z{=}z)>0$ and $P(X{=}x)>0$ for all $x\in X(\Omega)$.

Definition 6.18 [Unbiasedness of $E(Y|X,Z)$ and $E^{Z=z}(Y|X)$]
Let the Assumptions 6.1 (a) to (e) and (h) hold.

(i) *Then $E(Y|X,Z)$ is called unbiased, denoted $E(Y|X,Z)\vdash D_X$, if*

$$\forall\, x\in X(\Omega)\colon\ E^{X=x}(Y|Z)\vdash D_X. \tag{6.23}$$

(ii) *If we additionally assume 6.1 (f) and $P^{Z=z}(X{=}x)>0$, for all $x\in X(\Omega)$, then $E^{Z=z}(Y|X)$ is called unbiased, denoted $E^{Z=z}(Y|X)\vdash D_X$, if*

$$\forall\, x\in X(\Omega)\colon\ E^{Z=z}(Y|X{=}x)\vdash D_X. \tag{6.24}$$

Under the assumptions of Definition 6.18 (ii),

$$\forall\, x \in X(\Omega){:}\;\; E^{Z=z}(Y \mid X{=}x) \vdash D_X \quad \Leftrightarrow \quad \forall\, x \in X(\Omega){:}\;\; E^{X=x}(Y \mid Z{=}z) \vdash D_X \tag{6.25}$$

$$\Leftrightarrow \quad \forall\, x \in X(\Omega){:}\;\; E(Y \mid X{=}x, Z{=}z) \vdash D_X \tag{6.26}$$

[see Props. (6.19) and (6.20)]. Hence, we can replace Proposition (6.24) by each of the propositions on the right-hand sides of (6.25) and (6.26).

In the next theorem we treat a condition that is equivalent to unbiasedness of a conditional expectation $E^{Z=z}(Y \mid X)$.

Theorem 6.19 [[A Condition Equivalent to Unbiasedness of $E^{Z=z}(Y \mid X)$]
Let the Assumptions 6.1 (a) to (f) and (h) hold. Furthermore, for all $x \in X(\Omega)$, assume $P^{Z=z}(X{=}x) > 0$ and that τ_x is $P^{Z=z}$-unique. Then

$$\big(\forall\, x \in X(\Omega){:}\;\; E^{Z=z}(Y \mid X{=}x) = E^{Z=z}(\tau_x)\big) \quad \Leftrightarrow \quad E^{Z=z}(Y \mid X) \vdash D_X. \tag{6.27}$$

(Proof p. 188)

6.2.2 Equivalent Conditions of Z-Conditional Unbiasedness

Now we turn to some conditions that are equivalent to unbiasedness of a conditional expectation $E^{X=x}(Y \mid Z)$. These conditions are also used in the proofs of sufficient conditions of unbiasedness (see chs. 8 to 10). Note again, in this chapter we assume that Z is a covariate of X. That is, in contrast to chapter 5 where we defined a causal Z-conditional total effect variable, now we exclude that a putative cause variable X can take the role of Z.

Theorem 6.20 [Conditions Equivalent to Unbiasedness of $E^{X=x}(Y \mid Z)$]
Let the Assumptions 6.1 (a) to (e) hold and that τ_x is P-unique.

(i) Then

$$E^{X=x}(Y \mid Z) \vdash D_X \quad \Leftrightarrow \quad E^{X=x}(\tau_x \mid Z) \underset{P}{=} E(\tau_x \mid Z) \tag{6.28}$$

$$\Leftrightarrow \quad E(\tau_x \mid 1_{X=x}, Z) \underset{P}{=} E(\tau_x \mid Z). \tag{6.29}$$

(ii) If, additionally, for all $\tau_x \in \mathscr{E}^{X=x}(Y \mid D_X)$, we define

$$\varepsilon_x \;:=\; \tau_x \;-\; E(\tau_x \mid 1_{X=x}, Z), \tag{6.30}$$

then

$$E^{X=x}(Y \mid Z) \vdash D_X \quad \Leftrightarrow \quad E^{X=x}(\varepsilon_x \mid Z) \underset{P}{=} E(\varepsilon_x \mid Z) \tag{6.31}$$

$$\Leftrightarrow \quad E(\varepsilon_x \mid 1_{X=x}, Z) \underset{P}{=} E(\varepsilon_x \mid Z). \tag{6.32}$$

(Proof p. 189)

Hence, if we assume that Z is a covariate of X and τ_x is P-unique, then Theorem 6.20 provides four equations that are equivalent to $E^{X=x}(Y \mid Z) \underset{P}{=} E(\tau_x \mid Z)$. The first two use the true outcome variable τ_x itself, whereas the last two use the residual ε_x of τ_x with respect to the conditional expectation $E(\tau_x \mid 1_{X=x}, Z)$ [see Eq. (6.30)]. According to Equation (6.29), the conditional expectation $E^{X=x}(Y|Z)$ is unbiased if and only if τ_x is Z-conditionally mean-independent from the indicator variable $1_{X=x}$.

In the next corollary we present some conditions that are equivalent to unbiasedness of the conditional expectation $E(Y|X,Z)$, provided that Z is a covariate of X and all true outcome variables τ_x are P-unique. This corollary immediately follows from Theorem 6.20 and Definition 6.18 (i).

Corollary 6.21 [Conditions Equivalent to Unbiasedness of $E(Y|X,Z)$]
Let the Assumptions 6.1 (a) to (e) and (h) hold, and assume that for all $x \in X(\Omega)$, τ_x is P-unique. Then

$$E(Y|X,Z) \vdash D_X \quad \Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E^{X=x}(\tau_x|Z) \underset{P}{=} E(\tau_x|Z) \tag{6.33}$$

$$\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E(\tau_x|\,1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z). \tag{6.34}$$

If we additionally assume that ε_x is the residual defined by Equation (6.30), then

$$E(Y|X,Z) \vdash D_X \quad \Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E^{X=x}(\varepsilon_x|Z) \underset{P}{=} E(\varepsilon_x|Z) \tag{6.35}$$

$$\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ E(\varepsilon_x|\,1_{X=x}, Z) \underset{P}{=} E(\varepsilon_x|Z). \tag{6.36}$$

Hence, according to Equation (6.34), the conditional expectation $E(Y|X,Z)$ is unbiased if and only if, for all $x \in X(\Omega)$, the true outcome variable τ_x is Z-conditionally mean-independent from the indicator variable $1_{X=x}$. Furthermore, $E(Y|X,Z)$ is also unbiased if and only if, for all $x \in X(\Omega)$, the residual ε_x is Z-conditionally mean-independent from $1_{X=x}$ [see Eq. (6.36)].

In the next theorem we present two conditions each of which is equivalent to unbiasedness of $E^{X=x}(Y|Z=z)$. According to Equation (6.15), these conditions are also equivalent to unbiasedness of $E(Y|X=x, Z=z)$.

Theorem 6.22 [Condition Equivalent to Unbiasedness of $E^{X=x}(Y|Z=z)$]
Let the Assumptions 6.1 (a) to (f) hold. If τ_x is $P^{Z=z}$-unique, then

$$E^{X=x}(Y|Z=z) \vdash D_X \quad \Leftrightarrow \quad E^{X=x}(\tau_x|Z=z) = E(\tau_x|Z=z). \tag{6.37}$$

If we additionally assume that ε_x is defined by Equation (6.30), then

$$E^{X=x}(Y|Z=z) \vdash D_X \quad \Leftrightarrow \quad E^{X=x}(\varepsilon_x|Z=z) = E(\varepsilon_x|Z=z). \tag{6.38}$$

(Proof p. 190)

6.3 Unbiasedness of Prima Facie Effects

Now we introduce the concept of unbiasedness of a *prima facie effect*

$$PFE_{xx'} := E(Y|X=x) - E(Y|X=x'). \tag{6.39}$$

Furthermore, for a covariate Z of X, we also extend the concept of unbiasedness to a *Z-conditional prima facie effect function* $PFE_{Z;xx'}: \Omega'_Z \to \mathbb{R}$, which satisfies

$$PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z), \tag{6.40}$$

where $PFE_{Z;xx'}(Z)$ denotes the composition of Z and $PFE_{Z;xx'}$. While $PFE_{Z;xx'}$ assigns to each value $z \in \Omega'_Z$ a $(Z{=}z)$-conditional prima facie effect of x compared to x', the composition $PFE_{Z;xx'}(Z)$ is a random variable on $(\Omega, \mathscr{A}, P)$ assigning values to each $\omega \in \Omega$. We call the composition $PFE_{Z;xx'}(Z)$ a *Z-conditional prima facie effect variable*. Finally, presuming $P(X{=}x, Z{=}z),\ P(X{=}x', Z{=}z) > 0$, we define the *$(Z{=}z)$-conditional prima facie effect*

$$PFE_{Z;xx'}(z) := E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z). \tag{6.41}$$

Definition 6.23 [Unbiasedness of a Prima Facie Effect]
Let the Assumptions 6.1 (a) to (d) and (g) hold.

(i) *Then the prima facie effect $PFE_{xx'}$ is called unbiased, denoted $PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}$, if*
  - (a) τ_x *and* $\tau_{x'}$ *are* P*-unique*
  - (b) $PFE_{xx'} = E(\tau_x - \tau_{x'})$.

(ii) *If we additionally assume 6.1 (e), then $PFE_{Z;xx'}$ and $PFE_{Z;xx'}(Z)$ are called unbiased, denoted $PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$, if*
  - (a) τ_x *and* $\tau_{x'}$ *are* P*-unique*
  - (b) $PFE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'}|Z)$.

(iii) *If we additionally assume 6.1 (f), $P(X{=}x, Z{=}z) > 0$, and $P(X{=}x', Z{=}z) > 0$, then $PFE_{Z;xx'}(z)$ is called unbiased, denoted $PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}$, if*
  - (a) τ_x *and* $\tau_{x'}$ *are* $P^{Z=z}$*-unique*
  - (b) $PFE_{Z;xx'}(z) = E(\tau_x - \tau_{x'}|Z{=}z)$.

Note again that all three concepts of unbiasedness refer to total effects and that Z denotes a covariate of X.

Now we show that unbiasedness of the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x')$ implies unbiasedness of the prima facie effect $PFE_{xx'}$.

Theorem 6.24 [Unbiasedness of the Prima Facie Effect $PFE_{xx'}$]
Let the Assumptions 6.1 (a) to (d) and (g) hold. Then

$$\big(E(Y|X{=}x) \vdash D_X \wedge E(Y|X{=}x') \vdash D_X\big) \quad \Rightarrow \quad PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}. \tag{6.42}$$

(Proof p. 190)

Hence, under the Assumptions 6.1 (a) to (d) and (g), unbiasedness of $E(Y|X{=}x)$ and $E(Y|X{=}x')$ implies that the prima facie effect $PFE_{xx'}$ is unbiased.

Next, we show that unbiasedness of the Z-conditional expectations $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ implies unbiasedness of the Z-conditional prima facie effect function $PFE_{Z;xx'}$.

Theorem 6.25 [Unbiasedness of Prima Facie Effect Function $PFE_{Z;xx'}$]
Let the Assumptions 6.1 (a) to (e), and (g) hold. Then

$$\big(E^{X=x}(Y|Z) \vdash D_X \wedge E^{X=x'}(Y|Z) \vdash D_X\big) \quad \Rightarrow \quad PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}. \tag{6.43}$$

(Proof p. 190)

Hence, under the Assumptions 6.1 (a) to (e) and (g), unbiasedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ implies that the Z-conditional prima facie effect function $PFE_{Z;xx'}$ and the composition $PFE_{Z;xx'}(Z)$ of $PFE_{Z;xx'}$ and Z, the Z-conditional prima facie effect variable, are unbiased [see Def. 6.23 (ii)].

In the following theorem, we explicate the relationship between unbiasedness of the $(Z{=}z)$-conditional expectation values $E^{X=x}(Y|Z{=}z)$, $E^{X=x'}(Y|Z{=}z)$, and unbiasedness of the $(Z{=}z)$-conditional prima facie effect $PFE_{Z;xx'}(z)$.

Theorem 6.26 [Unbiasedness of a $(Z{=}z)$-Conditional Prima Facie Effect $PFE_{Z;xx'}(z)$]
Let the Assumptions 6.1 (a) to (f) and (g) hold. Then

$$\big(E^{X=x}(Y|Z{=}z) \vdash D_X \wedge E^{X=x'}(Y|Z{=}z) \vdash D_X\big) \quad \Rightarrow \quad PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}. \tag{6.44}$$

(Proof p. 191)

Hence, under the Assumptions 6.1 (a) to (f) and (g), unbiasedness of conditional expectation values $E^{X=x}(Y|Z{=}z)$ and $E^{X=x'}(Y|Z{=}z)$ implies that the $(Z{=}z)$-conditional prima facie effect $PFE_{Z;xx'}(z)$ is unbiased.

Box 6.1 summarizes the definitions of unbiasedness of various conditional expectations, their values, prima facie effect functions, and prima facie effects.

Remark 6.27 [Estimability of Conditional Expectations] While true outcome variables τ_x and their values are not directly estimable in a data sample unless rather restrictive assumptions are introduced, the conditional expectation values $E(Y|X{=}x)$, $E(Y|X{=}x,Z{=}z)$, and the conditional expectations $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$ *can* be estimated under realistic assumptions, and the same is true for the conditional and unconditional prima facie effects. As we shall see, this implies that causal average total effects and causal $(Z{=}z)$-conditional total effects can be estimated as well, provided that we assume that the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$, and $E(Y|X,Z)$ mentioned above are unbiased (see sect. 6.4 for details). ◁

Remark 6.28 [Unbiasedness and Randomization] In chapter 8 we show that unbiasedness of the conditional expectations, their values, and the prima facie effects can be created by randomized assignment of the observational unit to one of several treatment conditions. Unbiasedness can also be strived for by covariate selection. That is, we may try to select covariates $Z_1, \ldots, Z_m$ of X such that unbiasedness of the conditional expectations $E^{X=x}(Y|Z)$ holds for the m-variate covariate $Z := (Z_1, \ldots, Z_m)$ and all values x of X. ◁

Remark 6.29 [Unbiasedness and Covariate Selection] Unfortunately, unbiasedness cannot be used as a criterion for covariate selection. The reason is that it cannot be tested empirically because the definitions involve the true outcome variables τ_x. These variables

Box 6.1 Unbiasedness

Unbiasedness of various conditional expectations, their values, prima facie effect functions, and prima facie effects is symbolized and defined as follows:

$E(Y \mid X{=}x) \vdash D_X$	$E(Y \mid X{=}x)$ *is unbiased*. Under the Assumptions 6.1 (a) to (d) it is defined by P-uniqueness of τ_x and $E(Y \mid X{=}x) = E(\tau_x)$.
$E(Y \mid X) \vdash D_X$	$E(Y \mid X)$ *is unbiased*. Under the Assumptions 6.1 (a) to (d) and (h) it is defined by unbiasedness of $E(Y \mid X{=}x)$ for all values $x \in X(\Omega)$.
$PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}$	*The prima facie effect* $PFE_{xx'} = E(Y \mid X{=}x) - E(Y \mid X{=}x')$ *is unbiased*. Under the Assumptions 6.1 (a) to (d) and (g) it is defined by P-uniqueness of τ_x, $\tau_{x'}$ and $PFE_{xx'} = E(\tau_x - \tau_{x'})$.
$E^{X=x}(Y \mid Z) \vdash D_X$	$E^{X=x}(Y \mid Z)$ *is unbiased*. Under the Assumptions 6.1 (a) to (e) it is defined by P-uniqueness of τ_x and $E^{X=x}(Y \mid Z) \underset{P}{=} E(\tau_x \mid Z)$.
$E(Y \mid X, Z) \vdash D_X$	$E(Y \mid X, Z)$ *is unbiased*. Under the Assumptions 6.1 (a) to (e) and (h) it is defined by unbiasedness of $E^{X=x}(Y \mid Z)$ for all values $x \in X(\Omega)$.
$PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$	*The prima facie effect function* $PFE_{Z;xx'} \colon \Omega'_Z \to \mathbb{R}$ *and the prima facie effect variable* $PFE_{xx'}(Z) \underset{P}{=} E^{X=x}(Y \mid Z) - E^{X=x'}(Y \mid Z)$ *are unbiased*. Under the Assumptions 6.1 (a) to (e), and (g) it is defined by P-uniqueness of τ_x, $\tau_{x'}$, and $PFE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'} \mid Z)$.
$E^{X=x}(Y \mid Z{=}z) \vdash D_X$	$E^{X=x}(Y \mid Z{=}z)$ *is unbiased*. Under the Assumptions 6.1 (a) to (f), and $P(X{=}x, Z{=}z) > 0$ it is defined by $P^{Z=z}$-uniqueness of τ_x and $E^{X=x}(Y \mid Z{=}z) = E(\tau_x \mid Z{=}z)$.
$E^{Z=z}(Y \mid X) \vdash D_X$	$E^{Z=z}(Y \mid X)$ *is unbiased*. Under the Assumptions 6.1 (a) to (f), (h), and $P(X{=}x, Z{=}z) > 0$ for all $x \in X(\Omega)$, it is defined by unbiasedness of $E^{X=x}(Y \mid Z{=}z)$ for all values $x \in X(\Omega)$.
$PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}$	*The prima facie effect* $PFE_{Z;xx'}(z) = E^{X=x}(Y \mid Z{=}z) - E^{X=x'}(Y \mid Z{=}z)$ *is unbiased*. Under the Assumptions 6.1 (a) to (g) and $P(X{=}x, Z{=}z)$, $P(X{=}x', Z{=}z) > 0$, it is defined by $P^{Z=z}$-uniqueness of τ_x, $\tau_{x'}$ and $PFE_{Z;xx'}(z) = E(\tau_x - \tau_{x'} \mid Z{=}z)$.

even cannot be estimated unless very restrictive assumptions are introduced. This has been discussed in some detail by Holland (1986) and has been called the "fundamental problem of causal inference" (see also the preface). However, in chapters 7 to 10 we introduce other causality conditions that *can* be tested empirically and that imply unbiasedness. It is those causality conditions that can be used for covariate selection in empirical causal research. $\triangleleft$

6.4 Identification of Causal Total Effects

In chapter 5 we introduced causal average total effects and causal conditional total effect functions, which, in the first place, are of a purely theoretical nature. They just define what we are interested in, for example, in studies evaluating the causal effects of a treatment, an intervention, or an exposition. Now we study how causal total effects can be *identified*, that is, how they can be computed from parameters that can be estimated in samples, and how the causal conditional total effect functions can be identified by functions that can be estimated in samples.

6.4.1 Identification of the Causal Average Total Effect

In Definition 5.26 we introduced the causal average total effect

$$ATE_{xx'} \;=\; E(\tau_x - \tau_{x'}), \tag{6.45}$$

presuming that τ_x and $\tau_{x'}$ are P-unique. Taking its expectation, the true total effect variable $\delta_{xx'} = \tau_x - \tau_{x'}$ is coarsened to a single number. With such a coarsening, we often lose information. However, the resulting causal average total effect is still unbiased. In this context, instead of coarsening, we also use the term *aggregation* or *re-aggregation*. To emphasize, re-aggregation does not mean to ignore the potential confounders of a putative cause variable X. By definition, a potential confounder of X is a random variable on the probability space considered that is measurable with respect to the global potential confounder D_X [see Def. 4.11 (iii)]. Re-aggregation only means coarsening and loosing information about more fine-grained conditional effects. It does not reintroduce bias. Instead, re-aggregation maintains causal interpretability.

Remark 6.30 [Basic Idea of the True-Outcome Theory of Causal Effects] In the construction of the theory of causal effects, we first condition on a global potential confounder D_X in order to control for all potential confounders. Doing this, we obtain the most fine-grained causal total effect variable $CTE_{D_X;xx'}(D_X) = \tau_x - \tau_{x'}$ [see Def. 5.18 (i)]. Then we re-aggregate it and obtain a coarsened causal effect function or effect parameter that can be computed from an empirically estimable function or parameter. ◁

In Definition 6.23 we defined unbiasedness of the prima facie effect $PFE_{xx'}$ by P-uniqueness of τ_x, $\tau_{x'}$, and $PFE_{xx'} \;=\; E(\tau_x - \tau_{x'})$. Hence, the definitions of $ATE_{xx'}$ and unbiasedness of the prima facie effect $PFE_{xx'}$ immediately yield the following corollary.

Corollary 6.31 [Identifying the Causal Average Total Effect via $PFE_{xx'}$]
Let the Assumptions 6.1 (a) to (d) and (g) hold, and assume that $PFE_{xx'}$ is unbiased. Then

$$ATE_{xx'} \;=\; PFE_{xx'}. \tag{6.46}$$

Remark 6.32 [Ignoring Potential Confounders of X] In contrast to re-aggregation of the true total effect variable $\tau_x - \tau_{x'}$, considering the prima facie effect $E(Y|X{=}x) - E(Y|X{=}x')$ — and in this sense, *ignoring potential confounders of X* — may lead to a completely wrong conclusion about the causal average total effect (of x compared to x' on Y), unless the prima facie effect $E(Y|X{=}x) - E(Y|X{=}x')$ is unbiased. Even a reversal of effects is possible. That is, there may be a positive prima facie effect while the causal average total effect is negative and vice versa. This is exemplified in the following example. ◁

Example 6.33 [Joe and Ann With Self-Selection] Table 6.1 displays the crucial parameters of a random experiment, in which the effect of treatment 1 compared to treatment 0 is reversed if the person variable U is ignored. This table has already been shown in a similar form in Table 1.2. However, now it is written in the terms introduced in chapters 5 and 6. In this example, the person variable U is a global potential confounder of X. The causal average total effect is

$$\begin{aligned} ATE_{10} &= E\big(E^{X=1}(Y|U) - E^{X=0}(Y|U)\big) \;=\; E\big(CTE_{U;10}(U)\big) \;=\; E(\tau_1 - \tau_0) \\ &= E(\tau_1) \;-\; E(\tau_0) \\ &= \sum_u E^{X=1}(Y|U=u)\cdot P(U=u) \;-\; \sum_u E^{X=0}(Y|U=u)\cdot P(U=u) \\ &= \left(.8\cdot\frac{1}{2} + .4\cdot\frac{1}{2}\right) \;-\; \left(.7\cdot\frac{1}{2} + .2\cdot\frac{1}{2}\right) \;=\; .6 - .45 \;=\; .15 \end{aligned}$$

[see Box 6.2 (ii)]. In contrast, the corresponding prima facie effect is

$$\begin{aligned} PFE_{10} &= E(Y|X=1) \;-\; E(Y|X=0) \\ &= E(\tau_1|X=1) \;-\; E(\tau_0|X=0) \\ &= \sum_u E^{X=1}(Y|U=u)\cdot P(U=u|X=1) \;-\; \sum_u E^{X=0}(Y|U=u)\cdot P(U=u|X=0) \\ &= \left(.8\cdot\frac{1}{20} + .4\cdot\frac{19}{20}\right) \;-\; \left(.7\cdot\frac{4}{5} + .2\cdot\frac{1}{5}\right) \;=\; -.18 \end{aligned}$$

[see Box 6.2 (i)]. Considering Box 6.2 and comparing Equations (i) and (ii) to each other reveals why such a reversal of effects (.15 vs. −.18) can occur: Computing the conditional expectation value $E(Y|X=x)$ we weigh the conditional expectation values $E^{X=x}(Y|U=u)$ by the conditional probabilities $P(U=u|X=x)$ [see Eq. (i) in Box 6.2], whereas computing the expectation $E(\tau_x)$ of the true outcome variable we weigh them by the probabilities $P(U=u)$ [see Eq. (ii) in that box]. [For a proof of Equations (i) to (iv) of Box 6.2 see Exercise 6-11]. ◁

According to the following theorem we can also identify the causal average total effect $ATE_{xx'}$ if the prima facie effect $PFE_{xx'}$ is not unbiased. It suffices to assume that the Z-conditional prima facie effect variable $PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$ is unbiased. Reading this theorem, note that Z may be a multivariate covariate of X such that $Z = (Z_1, \ldots, Z_m)$ consists of m univariate covariates Z_i, $i = 1, \ldots, m$.

Theorem 6.34 [Identifying the Causal Average Total Effect via $PFE_{Z;xx'}$]
Let the Assumptions 6.1 (a) to (e) and (g) hold, and assume that $PFE_{Z;xx'}$ is unbiased. Then

$$ATE_{xx'} \;=\; E\big(PFE_{Z;xx'}(Z)\big). \tag{6.47}$$

(Proof p. 191)

In Equation (6.47) we re-aggregate the prima facie effect function $PFE_{Z;xx'}$ to obtain a single number. If $PFE_{Z;xx'}$ is unbiased, then this does *not* mean to ignore the potential

Table 6.1. Joe and Ann with self-selection to treatment

Person u	$P(U{=}u)$	$P(X{=}1 \mid U{=}u)$	$E^{X=0}(Y \mid U{=}u)$	$E^{X=1}(Y \mid U{=}u)$	$CTE_{U;10}(u)$	$P(U{=}u \mid X{=}0)$	$P(U{=}u \mid X{=}1)$
Joe	1/2	.04	.7	.8	.1	4/5	1/20
Ann	1/2	.76	.2	.4	.2	1/5	19/20

	$x=0$	$x=1$	
$E(\tau_x)$:	.45	.6	$ATE_{10} = \ \ .15$
$E(Y \mid X{=}x)$:	.6	.42	$PFE_{10} = -.18$

confounders of X. Instead, we just re-aggregate (coarsen) the causal Z-conditional total effect variable to obtain a single number, the causal average total effect $ATE_{xx'}$.

Remark 6.35 [Foundation for the Analysis of Causal Total Effects] Theorem 6.34 is the theoretical foundation for the analysis of causal average total effects beyond the simple randomized experiment. The crucial assumption is unbiasedness of $PFE_{Z;xx'}$, and this assumption may also hold in observational studies. Of course, finding a (possibly multivariate) covariate Z of X for which $PFE_{Z;xx'}$ is unbiasedness is often a challenge for empirical research. Also note that if $PFE_{Z;xx'}$ is unbiased, then it is identical to the more fine-grained causal total effect variable $CTE_{Z;xx'}(Z)$ (see Cor. 6.37), which is much more informative than the causal average total effect $ATE_{xx'}$. In the chapters to come we will learn more about sufficient conditions for unbiasedness of $PFE_{Z;xx'}$. ◁

Remark 6.36 [Z-Adjusted $(X{=}x)$-Conditional Expectation Value of Y] If we insert Equation (6.40) into the right-hand side of Equation (6.47), then we obtain

$$ATE_{xx'} \ = \ E\big(E^{X=x}(Y \mid Z)\big) \ - \ E\big(E^{X=x'}(Y \mid Z)\big). \tag{6.48}$$

The first term on the right-hand side of this equation is called the *Z-adjusted $(X{=}x)$-conditional expectation value of Y*, and the second term the *Z-adjusted $(X{=}x')$-conditional expectation value of Y*. Again, we presume that Z is a covariate of X. ◁

6.4.2 Identification of a Causal Conditional Total Effect Function

In Definition 5.38 we introduced the causal conditional total effect function $CTE_{Z;xx'}$ and the composition $CTE_{Z;xx'}(Z)$ by

$$CTE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'} \mid Z), \tag{6.49}$$

assuming that τ_x and $\tau_{x'}$ are P-unique and that Z is s random variable on $(\Omega, \mathscr{A}, P)$. In Definition 6.23 (ii) we defined *unbiasedness* of the prima facie effect function $PFE_{Z;xx'}$ by

Box 6.2 Conditional expectation values in the examples of this chapter

Consider the examples presented in Tables 6.1 to 6.4, let $x \in X(\Omega)$ denote a value of the treatment variable X, and let U denote the observational-unit variable. In these examples, $U = D_X$, $P(X{=}x, U{=}u) > 0$ for all values $u \in U(\Omega)$, and $\tau_x = E^{X=x}(Y|U) \in \mathscr{E}^{X=x}(Y|U)$. Then:

$$E(Y|X{=}x) = E(\tau_x|X{=}x) = \sum_u E(Y|X{=}x, U{=}u) \cdot P(U{=}u|X{=}x), \tag{i}$$

whereas

$$E(\tau_x) = \sum_u E(Y|X{=}x, U{=}u) \cdot P(U{=}u). \tag{ii}$$

Additionally, let Z be a covariate of X and let $z \in Z(\Omega)$. For the examples in Tables 6.1 to 6.4, in which U is finite, this implies $P(X{=}x, Z{=}z) > 0$ and

$$E(Y|X{=}x, Z{=}z) = E(\tau_x|X{=}x, Z{=}z) = \sum_u E(Y|X{=}x, U{=}u) \cdot P(U{=}u|X{=}x, Z{=}z), \tag{iii}$$

whereas

$$E(\tau_x|Z{=}z) = \sum_u E(Y|X{=}x, U{=}u) \cdot P(U{=}u|Z{=}z). \tag{iv}$$

$$PFE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'}|Z), \tag{6.50}$$

assuming that Z is a covariate of X as well as P-uniqueness of τ_x and $\tau_{x'}$. Because the covariate Z of X is a random variable on $(\Omega, \mathscr{A}, P)$, these two definitions immediately imply the following corollary, according to which a conditional prima facie effect function $PFE_{Z;xx'}$ is a causal conditional total effect function $CTE_{Z;xx'}$ if it is unbiased.

Corollary 6.37 [Identifying a Causal Z-Conditional Effect Function via $PFE_{Z;xx'}$ I] *Let the Assumptions 6.1 (a) to (e) and (g) hold, and assume that $PFE_{Z;xx'}$ is unbiased. Then*

$$CTE_{Z;xx'}(Z) \underset{P}{=} PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z). \tag{6.51}$$

Remark 6.38 [A Measurability Assumption for Z] Under the assumptions of Theorem 6.34, the expectation of an unbiased prima facie effect variable $PFE_{Z;xx'}(Z)$ is identical to the causal average total effect $ATE_{xx'}$. This implies a complete re-aggregation of a causal Z-conditional total effect function to a single number. In Theorem 6.39 we extend this result to a partial re-aggregation that yields a causal V-conditional total effect variable $CTE_{V;xx'}(V)$. In this theorem we do not only assume $PFE_{Z;xx'}(Z)$ to be unbiased (and with it that Z is a covariate of X) and P-uniqueness of the true outcome variables τ_x and $\tau_{x'}$, but also

$$\sigma(V) \subset \sigma(1_{X=x}, 1_{X=x'}, Z). \tag{6.52}$$

Note that this assumption implies $\sigma(V) \subset \sigma(1_{X=x}, Z)$ and $\sigma(V) \subset \sigma(1_{X=x'}, Z)$. It holds, for example, if one of the following conditions applies:

(a) $\sigma(V) \subset \sigma(Z)$
(b) $\sigma(V) \subset \sigma(1_{X=x})$
(c) $\sigma(V) \subset \sigma(1_{X=x'})$.

Conditions implying (a) are found in RS-Lemma 2.35 and RS-Corollary 2.36. ◁

Theorem 6.39 [Identifying a Causal V-Conditional Effect Function via $PFE_{Z;xx'}$ II] *Let the Assumptions 6.1 (a) to (e) and (g) hold, let V be a random variable on $(\Omega, \mathscr{A}, P)$, assume that Proposition (6.52) holds and $PFE_{Z;xx'}(Z)$ is unbiased. Then*

$$CTE_{V;xx'}(V) \underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big) \tag{6.53}$$

$$\underset{P}{=} E\big(E^{X=x}(Y|Z)\,\big|\,V\big) - E\big(E^{X=x'}(Y|Z)\,\big|\,V\big). \tag{6.54}$$

(Proof p. 191)

Remark 6.40 [Comparing $PFE_{V;xx'}(V)$ to $E(PFE_{Z;xx'}(Z)\,|\,V)$] Even if V is Z-measurable, then the right-hand side of Equation (6.53) is not necessarily almost surely identical to the V-conditional prima facie effect function

$$PFE_{V;xx'}(V) \underset{P}{=} E^{X=x}(Y\,|\,V) - E^{X=x'}(Y\,|\,V).$$

The two terms $PFE_{V;xx'}(V)$ and $E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big)$ are functions of V, which implies that they are V-measurable (see RS-Lem. 2.35). However, if $PFE_{Z;xx'}(Z)$ is unbiased and (6.52) holds for V, then $E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big)$ is almost surely identical to a causal conditional total effect variable $CTE_{V;xx'}(V)$ [see Eq. (6.53)]. In contrast, even if V is Z-measurable, then $PFE_{V;xx'}(V)$ does not have a causal meaning unless it is unbiased. Unbiasedness of $PFE_{V;xx'}(V)$ does not follow from the assumptions of Theorem 6.39. Hence, Theorem 6.39 offers a way to identify $CTE_{V;xx'}(V)$ even if $PFE_{V;xx'}(V)$ is biased. The two crucial assumptions implying Equation (6.53) are unbiasedness of $PFE_{Z;xx'}(Z)$ and the measurability assumption (6.52). ◁

Now we extend the result of Theorem 6.39 on the identification of a causal V-conditional total effect function $CTE_{V;xx'}$. In Theorem 6.39 we assumed $\sigma(V) \subset \sigma(1_{X=x}, 1_{X=x'}, Z)$. In contrast, in Theorem 6.41 we assume

$$\sigma(V) \subset \sigma(X, Z). \tag{6.55}$$

This assumption holds, for example, if one of the following conditions applies:

(a) $\sigma(V) \subset \sigma(Z)$
(b) $\sigma(V) \subset \sigma(X)$.

More general conditions implying Proposition (6.55) are found in RS-Lemma 2.35 and RS-Corollary 2.36.

If $\sigma(X) \not\subset \sigma(1_{X=x}, 1_{X=x'})$, then Theorem 6.41 extends the results of Theorem 6.39 to a larger set of random variables taking the role of V. However, the price is to assume P-uniqueness of τ_x and $\tau_{x'}$ as well as Z-conditional mean-independence of τ_x and $\tau_{x'}$ from X, that is,

$$\tau_x \vDash X\,|\,Z \qquad \text{and} \qquad \tau_{x'} \vDash X\,|\,Z. \tag{6.56}$$

Remember, under P-uniqueness of τ_x and $\tau_{x'}$, these assumptions are equivalent to

$$E(\tau_x|X,Z) \underset{P}{=} E(\tau_x|Z) \qquad \text{and} \qquad E(\tau_{x'}|X,Z) \underset{P}{=} E(\tau_{x'}|Z), \tag{6.57}$$

respectively. Some sufficient conditions for the equations in Proposition (6.57) will be treated in chapter 7 (see, e. g., Table 7.2), which is devoted to the Rosenbaum-Rubin causality conditions.

Under P-uniqueness of τ_x and $\tau_{x'}$, assuming (6.57) implies unbiasedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$, that is, it implies

$$E^{X=x}(Y|Z) \vdash D_X \qquad \text{and} \qquad E^{X=x'}(Y|Z) \vdash D_X \tag{6.58}$$

(see Exercise 6-12). Finally, note that unbiasedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ is equivalent to $\tau_x \vDash 1_{X=x}|Z$ and $\tau_x \vDash 1_{X=x'}|Z$, provided that τ_x and $\tau_{x'}$ are P-unique [see Th. 6.20 (i)].

Theorem 6.41 [Identifying a Causal V-Conditional Effect Function via $PFE_{Z;xx'}$ III] *Let the Assumptions 6.1 (a) to (e) and (g) hold, and let V be a random variable on $(\Omega, \mathscr{A}, P)$ satisfying $\sigma(V) \subset \sigma(X,Z)$. Furthermore, assume that τ_x and $\tau_{x'}$ are P-unique, and that Equations (6.57) hold. Then*

$$CTE_{V;xx'}(V) \underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big) \tag{6.59}$$

$$\underset{P}{=} E\big(E^{X=x}(Y|Z)\,\big|\,V\big) - E\big(E^{X=x'}(Y|Z)\,\big|\,V\big). \tag{6.60}$$

(Proof p. 192)

Remark 6.42 [Foundation for the Analysis of Causal Conditional Total Effects] Corollary 6.37, Theorem 6.39, and Theorem 6.41 are the theoretical foundations for the analysis of causal conditional total effect functions. The crucial assumptions in Theorem 6.39 are unbiasedness of $PFE_{Z;xx'}$ and Proposition (6.52). In contrast, the crucial assumptions in Theorem 6.41 are that τ_x and $\tau_{x'}$ are P-unique, $\sigma(V) \subset \sigma(X,Z)$, and that the Equations (6.57) hold. In the chapters to come we will study various sufficient conditions for these requirements. $\lhd$

6.4.3 *Identification of a Causal Conditional Total Effect*

The causal conditional total effect $CTE_{Z;xx'}(z)$ has been defined by

$$CTE_{Z;xx'}(z) = E(\tau_x - \tau_{x'}|Z{=}z), \tag{6.61}$$

assuming that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique [see Def. 5.38 (ii)]. Furthermore, additionally assuming $P(X{=}x, Z{=}z) > 0$ and $P(X{=}x', Z{=}z) > 0$, unbiasedness of the conditional prima facie effect $PFE_{Z;xx'}(z)$ has been defined by the conjunction of

$$PFE_{Z;xx'}(z) = E(\tau_x - \tau_{x'}|Z{=}z) \tag{6.62}$$

and $P^{Z=z}$-uniqueness of τ_x and $\tau_{x'}$ [see Def. 6.23 (iii)]. The assumption that $PFE_{Z;xx'}(z)$ is unbiased comprises the assumption that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique. This assumption has already been explained in more detail in Remarks 5.34 and 5.36.

Equations (6.61) and (6.62) immediately imply the following corollary.

Corollary 6.43 [Identifying a Causal $(Z{=}z)$-Conditional Total Effect by $PFE_{Z;xx'}(z)$]
Let the Assumptions 6.1 (a) to (g) hold. Furthermore, assume that $P(X{=}x, Z{=}z) > 0$, $P(X{=}x', Z{=}z) > 0$, *and that* $PFE_{Z;xx'}(z)$ *is unbiased. Then*

$$CTE_{Z;xx'}(z) \;=\; PFE_{Z;xx'}(z) \;=\; E^{X=x}(Y \mid Z{=}z) - E^{X=x'}(Y \mid Z{=}z)\,. \tag{6.63}$$

Hence, under the assumptions of Corollary (6.43), the $(Z{=}z)$-conditional causal total effect on Y comparing x to x', that is, $CTE_{Z;xx'}(z)$, is identical to the corresponding prima facie effect $PFE_{Z;xx'}(z)$.

Now we consider again re-aggregation of a causal Z-conditional effect variable, assuming $CTE_{Z;xx'}(Z) \underset{P}{=} PFE_{Z;xx'}(Z)$, that is, assuming unbiasedness of $PFE_{Z;xx'}(Z)$. Theorem 6.39 and Theorem 6.41 imply the following corollary about the identification of the causal $(V{=}v)$-conditional total effect $CTE_{V;xx'}(v)$, which is a uniquely defined number because we assume $P(V{=}v) > 0$ (see RS-Rem. 4.26).

Corollary 6.44 [Identifying a Causal $(V{=}v)$-Conditional Total Effect via $PFE_{Z;xx'}$]
Let the assumptions of Theorem 6.39 or Theorem 6.41 hold and assume $P(V{=}v) > 0$. *Then*

$$\begin{aligned} CTE_{V;xx'}(v) \;&=\; E\big(PFE_{Z;xx'}(Z) \,\big|\, V{=}v\big) & (6.64)\\ &=\; E\big(E^{X=x}(Y \mid Z) \,\big|\, V{=}v\big) - E\big(E^{X=x'}(Y \mid Z) \,\big|\, V{=}v\big). & (6.65) \end{aligned}$$

(Proof p. 192)

Hence, under the assumptions of Corollary (6.44), the $(V{=}v)$-conditional causal total effect on Y comparing x to x', that is, $CTE_{V;xx'}(v)$, is identical to the $(V{=}v)$-conditional expected value of the prima facie effect variable $PFE_{Z;xx'}(Z)$. Note that the terms on the right-hand sides of these equations are estimable in an appropriate sampling model. Also note, the true outcome variables τ_x and $\tau_{x'}$ are implicitly involved in the term $CTE_{V;xx'}(v)$ because $CTE_{V;xx'}(v) = E(\tau_x \mid V{=}v) - E(\tau_{x'} \mid V{=}v)$. In contrast, true outcome variables are not involved in the terms on the right-hand sides of Equations (6.64) and (6.65).

Remark 6.45 [Unbiasedness of $PFE_{Z;xx'}(z)$ vs. Unbiasedness of $PFE_{V;xx'}(v)$] Again, note that even if V is Z-measurable, then $CTE_{V;xx'}(v)$ is not necessarily identical to

$$PFE_{V;xx'}(v) \;=\; E^{X=x}(Y \mid V{=}v) - E^{X=x'}(Y \mid V{=}v)\,.$$

Comparing the right-hand side of this equation to the right-hand side of Equation (6.65) reveals the difference. However, $CTE_{V;xx'}(v) = PFE_{V;xx'}(v)$, if $PFE_{V;xx'}(v)$ is unbiased. Note that unbiasedness of $PFE_{V;xx'}(v)$ is not implied by the assumptions of Corollary 6.44. Hence, Corollary 6.44 offers a way to identify $CTE_{V;xx'}(v)$ even if $PFE_{V;xx'}(v)$ is biased. The crucial assumptions are mentioned in Theorems 6.39 and 6.41. ◁

Remark 6.46 [Some Special Cases of V] A special case of Corollary 6.44 is $V = X = 1_{X=1}$. Hence, if the assumptions of Theorem 6.39 hold, then

Box 6.3 Identification of causal total effects and causal conditional effect functions

Various causal total effects and causal total effect functions can be identified ...

... by the equations	... under the Assumptions
$ATE_{xx'} = PFE_{xx'}$	6.1 (a) to (d), (g), and $PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}$.
$ATE_{xx'} = E\big(PFE_{Z;xx'}(Z)\big)$	6.1 (a) to (e), (g), and $PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$.
$CTE_{Z;xx'}(Z) \underset{P}{=} PFE_{Z;xx'}(Z)$	6.1 (a) to (e), (g), and $PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$.
$CTE_{Z;xx'}(z) = PFE_{Z;xx'}(z)$	6.1 (a) to (g), $P(X{=}x, Z{=}z) > 0$, $P(X{=}x', Z{=}z) > 0$, and $PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}$.
$CTE_{V;xx'}(V) \underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big\vert\, V\big)$	6.1 (a) to (e), (g), $PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$, V is a random variable on $(\Omega, \mathscr{A}, P)$ satisfying $\sigma(V) \subset \sigma(1_{X=x}, 1_{X=x'}, Z)$. An alternative set of assumptions under which this identification equation holds is: 6.1 (a) to (e), (g), Equations (6.57), V is a random variable on $(\Omega, \mathscr{A}, P)$ satisfying $\sigma(V) \subset \sigma(X, Z)$.
$CTE_{V;xx'}(v) = E\big(PFE_{Z;xx'}(Z)\,\big\vert\, V{=}v\big)$	$CTE_{V;xx'}(V) \underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big\vert\, V\big)$ and $P(V{=}v) > 0$.

See Box 6.1 for the definitions of the unbiasedness assumptions such as $PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}$.

$$
\begin{aligned}
CTE_{X;10}(0) &= E\big(PFE_{Z;10}(Z)\,\big\vert\, X{=}0\big) && (6.66)\\
&= E\big(E^{X=1}(Y\,|\,Z)\,\big\vert\, X{=}0\big) - E\big(E^{X=0}(Y\,|\,Z)\,\big\vert\, X{=}0\big) && (6.67)
\end{aligned}
$$

and

$$
\begin{aligned}
CTE_{X;10}(1) &= E\big(PFE_{Z;10}(Z)\,\big\vert\, X{=}1\big) && (6.68)\\
&= E\big(E^{X=1}(Y\,|\,Z)\,\big\vert\, X{=}1\big) - E\big(E^{X=0}(Y\,|\,Z)\,\big\vert\, X{=}1\big). && (6.69)
\end{aligned}
$$

In empirical applications, in which X is an indicator variable representing treatment $(X{=}1)$ and control $(X{=}0)$, the term $CTE_{X;10}(0)$ is known as the *causal total effect on the untreated of treatment* 1 *compared to treatment* 0. Similarly, $CTE_{X;10}(1)$ is known as the *causal total effect* 0 *on the treated of treatment* 1 *compared to treatment*. Note, however, that this wording is not in line with the pre facto perspective of probability theory (see Remarks 5.47 and 5.48). Hence, it is better to replace 'on the untreated' and 'on the treated' by 'given no treatment' and 'given treatment', respectively. Also note that, in particular under self-selection to treatment, the $(X{=}1)$-conditional distributions of potential confounders of X may differ from their $(X{=}0)$-conditional distributions. In informal terms, the subjects who select treatment are likely to differ before treatment from those who don't in important attributes such as 'severity of symptoms', 'motivation for treatment', and so on, which often leads to bias of prima facie effects.

Another special case of Corollary 6.44 is $V = (1_{X=1}, W)$, where W is a Z-measurable random variable. Hence, if the assumptions of Theorem 6.39 hold, X is a putative cause variable, and $P(X{=}x, W{=}w) > 0$ for $x = 0, 1$, then

$$CTE_{XW;\,10}(0, w) = E\big(PFE_{Z;10}(Z)\,\big|\,X{=}0, W{=}w\big) \tag{6.70}$$
$$= E\big(E^{X=1}(Y\,|\,Z)\,\big|\,X{=}0, W{=}w\big) - E\big(E^{X=0}(Y\,|\,Z)\,\big|\,X{=}0, W{=}w\big) \tag{6.71}$$

is the causal conditional total effect (comparing treatment to control) given control and the value w of W. Correspondingly,

$$CTE_{XW;\,10}(1, w) = E\big(PFE_{Z;10}(Z)\,\big|\,X{=}1, W{=}w\big) \tag{6.72}$$
$$= E\big(E^{X=1}(Y\,|\,Z)\,\big|\,X{=}1, W{=}w\big) - E\big(E^{X=0}(Y\,|\,Z)\,\big|\,X{=}1, W{=}w\big). \tag{6.73}$$

is the causal conditional total effect given treatment and the value w of W. ◁

6.5 Three Examples

Tables 6.2 to 6.4 show parameters pertaining to fictive random experiments such as the single-unit trials described in chapter 2. Among these parameters are the individual expectation values $E(Y\,|\,X{=}x, U{=}u)$ given the treatment conditions, and the individual treatment probabilities $P(X{=}1\,|\,U{=}u)$. The parameters presented in the tables can be used to generate sample data that would result if the random experiments to which the tables refer were conducted n times.[1]

The Regular Probabilistic Causality Space in All Examples

Remark 6.47 [The Probability Space] For simplicity, we consider random experiments in which no *fallible* covariates are observed and in which there is neither a second treatment variable nor any other variable that is simultaneous to the treatment variable. In this case, the set

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3 = \Omega_U \times \Omega_X \times \mathbb{R} \tag{6.74}$$

suffices to describe the set of possible outcomes of the random experiment, where $\Omega_U = \{Tom, Tim, Joe, Jim, Ann, Sue\}$ and $\Omega_X = \{treatment, control\}$. Furthermore, we consider the product σ-algebra $\mathscr{A} = \mathscr{P}(\Omega_U) \otimes \mathscr{P}(\Omega_X) \otimes \mathscr{B}$, where $\mathscr{B}$ denotes the Borel σ-algebra on $\mathbb{R}$, the set of real numbers (see RS-Rem. 1.14). The probability measure P on $(\Omega, \mathscr{A})$ is only partly known. Looking at Table 6.2, for example, we only know the conditional expectation $E(Y\,|\,X, U)$. In contrast, we do not know the conditional distribution $P_{Y|X,U}$, which would be known only if additional information were added, such as 'Y is conditionally normally distributed given (X, U)', with a specified conditional variance of Y given (X, U). However, for our purpose, the conditional distribution of Y is not relevant because we only consider

[1] Although the focus of this book is on theory and not on data analysis, we also provide sample data for each table on the home page of this book: www.causal-effects.de. These and other examples of this type as well as a data sample generated by these examples can easily be created with the PC-program CausalEffectsXplorer that is also provided on www.causal-effects.de, together with an extensive help file providing the most important concepts and formulas.

Table 6.2. Self-selection to treatment

Fundamental parameters

Person u	Sex z	$P(U=u)$	$P(X=1 \mid U=u)$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$CTE_{U;10}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$
Tom	m	1/6	4/7	88	100	12	3/21	4/21
Tim	m	1/6	3/7	98	103	5	4/21	3/21
Joe	m	1/6	6/7	68	81	13	1/21	6/21
Jim	m	1/6	5/7	78	86	8	2/21	5/21
Ann	f	1/6	2/7	106	114	8	5/21	2/21
Sue	f	1/6	1/7	116	130	14	6/21	1/21

	$x=0$	$x=1$	
$E(\tau_x)$:	92.333	102.333	$ATE_{10} = 10$
$E(Y \mid X=x)$:	100.286	94.429	$PFE_{10} = -5.857$
$E(\tau_x \mid Z=m)$:	83	92.5	$CTE_{Z;10}(m) = 9.5$
$E(Y \mid X=x, Z=m)$:	88	90.278	$PFE_{Z;10}(m) = 2.278$
$E(\tau_x \mid Z=f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y \mid X=x, Z=f)$:	111.455	119.333	$PFE_{Z;10}(f) = 7.879$

causal total effects that are defined in terms of the conditional expectation $E(Y \mid X, U)$ and its values $E(Y \mid X=x, U=u)$ (see ch. 5). ◁

Remark 6.48 [Filtration and Global Potential Confounder] The filtration $(\mathscr{F}_t)_{t \in T}$, $t \in T = \{1,2,3\}$ consists of the σ-algebras

$$\mathscr{F}_1 := \sigma(\pi_1), \quad \mathscr{F}_2 := \sigma(\pi_1, \pi_2), \quad \mathscr{F}_3 := \sigma(\pi_1, \pi_2, \pi_3), \tag{6.75}$$

where π_1, π_2, and π_3 are the projections $\pi_t : \Omega \to \Omega_t$, $t \in T$ (see Def. 4.6). ◁

Remark 6.49 [Random Variables] In all examples of this section, we consider the person variable $U: \Omega \to \Omega_U$ with value space $\big(\Omega_U, \mathscr{P}(\Omega_U)\big)$, the treatment variable $X: \Omega \to \Omega'_X = \{0,1\}$ with value space $\big(\Omega'_X, \mathscr{P}(\Omega'_X)\big)$, where X takes on the value 1 for treatment and 0 for control, and the real-valued outcome variable $Y: \Omega \to \mathbb{R}$ with value space $(\mathbb{R}, \mathscr{B})$. ◁

Remark 6.50 [Cause and Confounder σ-Algebras] Because there are no other variables that are simultaneous to the treatment variable X, the index sets J and K used in Definition 4.6 are $J = K = \{1\}$. Hence, the cause σ-algebra is $\mathscr{C} = \sigma(\pi_{2j}, j \in K) = \sigma(\pi_2)$ and the confounder σ-algebra is $\mathscr{D}_{\mathscr{C}} = \sigma(\pi_1, \pi_{2j}, j \in J \setminus K) = \sigma(\pi_1)$.

Note that $U = \pi_1$, which is a global potential confounder of X [see Def. 4.11 (iii)]. Furthermore, U is prior in $(\mathscr{F}_t)_{t \in T}$ to X and Y, and X is prior to Y (see sect. 3.1). ◁

$$\begin{tikzcd}
& \Omega_U = \{\mathit{Tom}, \mathit{Tim}, \ldots, \mathit{Sue}\} \arrow[dr, "g_1"] & \\
\Omega = \Omega_U \times \Omega_X \times \mathbb{R} \arrow[ur, "U"] \arrow[rr, "\tau_1 = E^{X=1}(Y|U) = g_1(U)"'] & & \mathbb{R}
\end{tikzcd}$$

Figure 6.1. The person variable U, the function g_1, and their composition, the true outcome variable $\tau_1 = g_1(U)$.

Hence, for all examples of this section, we specified the regular probabilistic causality setup

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$$

[see Def. 4.11 (v)] and asserted that U is a global potential confounder of X.

True Outcome Variables and Individual Treatment Probabilities

Remark 6.51 [True Outcome Variables] In all examples of this chapter, $P(X{=}x, U{=}u) > 0$ for all pairs (x, u) of values of X and U. Therefore, and because $\sigma(U) = \mathscr{D}_{\mathscr{C}}$, the two true outcomes variables are

$$\tau_x \;:=\; E^{X=x}(Y|D_X) \;=\; E^{X=x}(Y|U), \quad x \in \{0, 1\}. \tag{6.76}$$

According to Equation (5.4), the true outcome variable τ_x can also be written as a function of the person variable U. More specifically, it can be written as the composition of U and the function $g_x\colon \Omega_U \to \mathbb{R}$ defined by

$$g_x(u) \;=\; E(Y|X{=}x, U{=}u), \quad \text{for all } u \in \Omega_U. \tag{6.77}$$

Note, that in this definition, x is a fixed value of X, and $\tau_x = g_x(U)$ is the composition of U and g_x, that is,

$$\forall\, \omega \in \Omega: \quad \tau_x(\omega) \;=\; g_x\big(U(\omega)\big) \;=\; g_x(u), \quad \text{if } \omega \in \{U{=}u\}. \tag{6.78}$$

This implies that the values of the conditional expectation $E^{X=x}(Y|U)$ are identical to the conditional expectation values $E(Y|X{=}x, U{=}u)$ [see Eq. (5.3)]. Figure 6.1 illustrates these equations for treatment $x = 1$. ◁

Remark 6.52 [Individual Treatment Probabilities] In all examples of this chapter,

$$P(X{=}1|D_X) \;=\; P(X{=}1|U). \tag{6.79}$$

Note, by definition, $P(X{=}1|D_X) = E(X|D_X)$ and $P(X{=}1|U) = E(X|U)$ (see RS-Rem. 4.12), provided that X is dichotomous with values 0 and 1. In all examples of this chapter, the D_X-conditional treatment probability is uniquely defined and it is identical to the U-conditional treatment probability $P(X{=}1|U)$, whose values are denoted by $P(X{=}1|U{=}u)$ and also called the *individual treatment probabilities*. ◁

Description of the Examples

In the *first example* (see Table 6.2), the individual treatment probabilities are *different for each and every person*, and they strongly depend on the individual expectation values of the outcomes under control, that is, they depend on the true outcome variable τ_0 and also on τ_1. The conditional expectation values $E(Y|X{=}1)$ and $E(Y|X{=}0)$ are biased. In fact, the prima facie effect PFE_{10} is negative, whereas the causal average total effect ATE_{10} is positive. Furthermore, the conditional expectation values $E(Y|X{=}1, Z{=}z)$ and $E(Y|X{=}0, Z{=}z)$ are biased as well. Although the causal $(Z{=}z)$-conditional total effects and the causal average total effect are defined (and can be computed from the parameters displayed in the upper left part of the table), they cannot be estimated from empirically estimable parameters such as the conditional expectation values $E(Y|X{=}1)$ and $E(Y|X{=}0)$ or $E(Y|X{=}1, Z{=}z)$ and $E(Y|X{=}0, Z{=}z)$.

In the *second example* (see Table 6.3), the treatment probabilities are *identical for all persons*, implying that X and U are independent, which has many implications that are studied in detail in chapter 8. Among these implications are that the conditional expectation $E(Y|X)$ and its values $E(Y|X{=}x)$ as well as the conditional expectation $E(Y|X,Z)$ and its values $E(Y|X{=}x, Z{=}z)$ are unbiased.

In the *third example* (see Table 6.4), the treatment probabilities *differ between males and females*. Furthermore, males (m) and females (f) also differ in their conditional expectation values of the true outcome variable τ_0, that is, $E(\tau_0|Z{=}m) \neq E(\tau_0|Z{=}f)$. However, given the value m or f of Z, the treatment probabilities do not differ from each other. This implies that X and U are Z-conditionally independent. In this example, in which $P(X{=}x, U{=}u) > 0$ for all pairs of values of X and U, implying that the true outcome variables τ_0 and τ_1 are P-unique (see Rem. 5.17), we can conclude that $E(Y|X,Z)$ is unbiased (see Th. 8.31). Hence, in this third example, the conditional expectation values $E(Y|X{=}x)$ are biased, whereas the conditional expectation values $E(Y|X{=}x, Z{=}z)$ are unbiased. Again, this case is studied extensively in chapter 8.

Tables 6.2, 6.3, and 6.4 display the true outcomes and the individual treatment probabilities. According to Equation (6.76), the values of τ_x, are also the individual conditional expectation values $E(Y|X{=}x, U{=}u)$, and, according to Equation (6.79), the values of the conditional probability $P(X{=}1|D_X) = P(X{=}1|U)$ are identical to the individual treatment probabilities $P(X{=}1|U{=}u)$. The tables also display the values of $Z := sex$, which is a the covariate of X because it is U-measurable.

Looking at the first four numerical columns, the three tables differ only in the treatment probabilities $P(X{=}1|U{=}u)$. All other entries in these first four columns, such as the true outcomes are the same. However, if we look at the other parameters, the three tables differ in important aspects.

$(X{=}x)$-Conditional Expectation Values

We start computing the $(X{=}x)$-conditional expectation values of Y and check whether or not they are unbiased. In the first example (see Table 6.2), $E(Y|X{=}0) = 100.286$, whereas $E(\tau_0) = 92.333$, that is, $E(Y|X{=}0)$ is much larger than $E(\tau_0)$. In contrast, $E(Y|X{=}1) = 94.429$, whereas $E(\tau_1) = 102.333$, that is, $E(Y|X{=}1)$ is much smaller than $E(\tau_0)$. Hence, according to Definition 6.3 (i), the conditional expectation values $E(Y|X{=}x)$ are biased, and according Definition 6.3 (ii) this is also true for the conditional expectation $E(Y|X)$.

Table 6.3. Randomized assignment of the person to a treatment

Fundamental parameters

Person u	Sex z	$P(U=u)$	$P(X=1 \mid U=u)$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$CTE_{U;10}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$
Tom	m	1/6	3/4	68	81	13	1/6	1/6
Tim	m	1/6	3/4	78	86	8	1/6	1/6
Joe	m	1/6	3/4	88	100	12	1/6	1/6
Jim	m	1/6	3/4	98	103	5	1/6	1/6
Ann	f	1/6	3/4	106	114	8	1/6	1/6
Sue	f	1/6	3/4	116	130	14	1/6	1/6

	$x=0$	$x=1$	
$E(\tau_x)$:	92.333	102.333	$ATE_{10} = 10$
$E(Y \mid X=x)$:	92.333	102.333	$PFE_{10} = 10$
$E(\tau_x \mid Z=m)$:	83	92.5	$CTE_{Z;10}(m) = 9.5$
$E(Y \mid X=x, Z=m)$:	83	92.5	$PFE_{Z;10}(m) = 9.5$
$E(\tau_x \mid Z=f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y \mid X=x, Z=f)$:	111	122	$PFE_{Z;10}(f) = 11$

These expectations and conditional expectations are easy to compute from the parameters displayed in Table 6.2. The expectation $E(\tau_x)$ of the true outcome variable τ_x is obtained using the *unconditional probabilities* $P(U=u)$ as weights [see Box 6.2 (ii)]. In contrast, the corresponding conditional expectations values $E(Y \mid X=x)$ are identical to the $(X=x)$-conditional expectation values $E(\tau_x \mid X=x)$ of the true outcome variables, using as weights the *conditional probabilities* $P(U=u \mid X=x)$ [see Box 6.2 (i)].

If used for the evaluation of the total treatment effect, the $(X=x)$-conditional expectation values would lead to completely wrong conclusions. First, the direction of the prima facie effect

$$PFE_{10} \;=\; E(Y \mid X=1) - E(Y \mid X=0) \;= -5.857$$

is reversed if compared to

$$ATE_{10} \;=\; E\big(CTE_{U;10}(U)\big) \;=\; E(\tau_1) - E(\tau_0) \;= 10.$$

And second, it is also reversed if compared to each and every individual total effect [see the column $CTE_{U;10}(u)$ in Table 6.2]. *All individual total effects are positive* in this example, ranging between 5 and 14. The bias in this example is due to strong inter-individual differences in the true outcomes and to the fact that the individual treatment probabilities $P(X=1 \mid U=u)$ heavily depend on the true outcome variables, and, therefore, on the person variable U. For instance, *Tom* has a true outcome under control of $E^{X=0}(Y \mid U=Tom) =$

Table 6.4. Conditionally randomized assignment of the person to a treatment

Fundamental parameters

Person u	Sex z	$P(U=u)$	$P(X=1 \mid U=u)$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$CTE_{U;10}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$
Tom	m	1/6	3/4	68	81	13	1/10	3/14
Tim	m	1/6	3/4	78	86	8	1/10	3/14
Joe	m	1/6	3/4	88	100	12	1/10	3/14
Jim	m	1/6	3/4	98	103	5	1/10	3/14
Ann	f	1/6	1/4	106	114	8	3/10	1/14
Sue	f	1/6	1/4	116	130	14	3/10	1/14

	$x=0$	$x=1$	
$E(\tau_x)$:	92.333	102.333	$ATE_{10} = 10$
$E(Y \mid X=x)$:	99.8	96.714	$PFE_{10} = -3.086$
$E(\tau_x \mid Z=m)$:	83	92.5	$CTE_{Z;10}(m) = 9.5$
$E(Y \mid X=x, Z=m)$:	83	92.5	$PFE_{Z;10}(m) = 9.5$
$E(\tau_x \mid Z=f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y \mid X=x, Z=f)$:	111	122	$PFE_{Z;10}(f) = 11$

68 and a treatment probability of 6/7, while *Sue* has a true outcome under control of $E^{X=0}(Y \mid U=Sue) = 116$ and a treatment probability of 1/7. Such a constellation is to be expected under self-selection of subjects to treatments, if the subjects base their decisions to take treatment on the *severity of their dysfunction before treatment* and if *severity of their dysfunction after treatment* is assessed as the outcome variable.

This example once again drastically demonstrates the necessity for the distinction between a difference $E(Y \mid X=1) - E(Y \mid X=0)$ of conditional expectation values and the causal average total effect $E(\tau_1) - E(\tau_0)$. Obviously, only the latter is of interest if we want to evaluate the treatment.

For the second example presented in Table 6.3, the situation is completely different. Although the true outcome variables are the same as in Table 6.2, here, the conditional expectation values $E(Y \mid X=x)$ and the expectations $E(\tau_x)$ of the true outcome variables $\tau_x = E^{X=x}(Y \mid U)$ are identical to each other, and this applies to both values $x=0$ and $x=1$ of X. Hence, in this example, the conditional expectation values $E(Y \mid X=x)$ are unbiased and can be used for the evaluation of the treatment effect. This is due to the fact that the individual treatment probabilities *do not depend on the persons*. This constellation occurs in a perfect randomized experiment, in which the experimenter decides that each person is in treatment 1 with probability $P(X=1)$ and in treatment 0 with probability $1 - P(X=1)$, provided, of course, that the persons comply with the experimenters decisions. In our second example, $P(X=1) = 3/4$. Note, however, that $P(X=1)$ could be *any* number between 0 and 1, exclusively. The only important point is that the individual treatment probabilities *do not differ between persons*, that is, $P(X{=}1 \,|\, U{=}u) = P(X{=}1)$ for all persons $u \in \Omega_U$. Such a randomized assignment may be performed by drawing a ball from an urn with three black balls and one white ball, adopting the rule that the subject is treated if a black ball is drawn.

In the third example (see Table 6.4), the conditional expectation values $E(Y \,|\, X{=}x)$ are biased again. Here, $E(Y \,|\, X{=}0) = 99.8$, whereas $E(\tau_0) = 92.333$. Again, $E(Y \,|\, X{=}0)$ is much larger than $E(\tau_0)$. In contrast, $E(Y \,|\, X{=}1) = 96.714$, whereas $E(\tau_1) = 102.333$, that is, again $E(Y \,|\, X{=}1)$ is much smaller than $E(\tau_1)$. Hence, in this example, the conditional expectation values $E(Y \,|\, X{=}x)$ are strongly biased as well. However, in contrast to the first example, the conditional expectation values $E(Y \,|\, X{=}x, Z{=}z)$ are unbiased. In this example, the treatment probability is $3/4$ *for all male units*, while it is $1/4$ *for all female units*. The crucial point is that these probabilities are identical given a value m or a value f of the covariate Z, that is, $P(X{=}1 \,|\, Z{=}z, U{=}u) = P(X{=}1 \,|\, Z{=}z)$ for each person u and both values z of the covariate Z. This constellation holds in a perfect *conditionally randomized experiment* in which we assign the sampled person to treatment with probability $P(X{=}1 \,|\, Z{=}m)$ if he is male and with probability $P(X{=}1 \,|\, Z{=}f)$ if the sampled person is female.

$(X{=}x, Z{=}z)$-Conditional Expectation Values

The conditional expectation values $E(Y \,|\, X{=}x, Z{=}z)$, can be computed from the parameters displayed in Table 6.4, applying Equation (iii) of Box 6.2. For this purpose we also need the formula

$$P(U{=}u \,|\, X{=}x, Z{=}z) \;=\; \frac{P(X{=}x \,|\, U{=}u) \cdot P(U{=}u, Z{=}z)}{P(X{=}x \,|\, Z{=}z) \cdot P(Z{=}z)}, \tag{6.80}$$

where

$$P(X{=}x \,|\, Z{=}z) \;=\; \frac{\sum_u P(X{=}x \,|\, U{=}u) \cdot P(U{=}u, Z{=}z)}{P(Z{=}z)} \tag{6.81}$$

(see Exercise 6-13). Note that in the three examples $P(X{=}x \,|\, U{=}u, Z{=}z) = P(X{=}x \,|\, U{=}u)$ because in theses examples Z is U-measurable. Intuitively speaking, this means that Z (*sex*) does not contain any information that is not already contained in U (the person variable). All terms on the right-hand side of Equation (6.80) are displayed in Table 6.4 or can be computed from the parameters displayed in this table.[2]

Remark 6.53 [How Realistic Are These Examples?] In empirical applications, assuming $D_X = U$ is correct if (a) there is neither a second treatment variable nor another variable that is simultaneous to X and if (b) no fallible covariate is observed. In this case, u signifies the observational unit *at the onset of treatment*. If, however, a fallible covariate of X is observed and u represents the observational unit *at the time at which the covariate is assessed*, then there may very well be covariates that are not measurable with respect to U, which affect the outcome variable Y and/or the treatment probability (see section 2.2). Hence, in this case, $D_X = U$ would not hold. In this case a global potential confounder would be the multivariate random variable (U, Z), where $Z = (Z_1, \dots, Z_m)$ consists of the fallible covariates Z_i, $i = 1, \dots, m$, to be assessed before treatment. That is, in this case $D_X = (U, Z)$ is a global potential confounder of X [see Def. 4.11 (iii)]. $\lhd$

[2] An alternative is using the Causal Effects Xplorer provided at www.causal-effects.de, the home page of this book.

Conditional Total Effects

Comparing the conditional prima facie effects to the causal conditional total effects reveals that the conditional prima facie effects are still biased with respect to total effects in the random experiment presented in Table 6.2, but not in the examples displayed in Tables 6.3 and 6.4. Hence, in the first example, the conditional prima facie effect and the causal conditional total effect for males are *not* identical, while they *are* identical in the second and third example, and the same applies to the corresponding conditional prima facie effects for females.

The bias of the conditional prima facie effects in the example presented in Table 6.2 is no surprise because there are still individual differences within the two sets of males and females with respect to (a) the true outcomes under treatment and under control, as well as (b) in the individual treatment probabilities $P(X{=}1|U{=}u)$. In contrast, in the second and third example, the individual treatment probabilities are all the same *within each of the two sets of males and females*.

Average Total Effect

In all three examples, the average over the individual total effects is equal to the causal average total effect. [Remember, the causal average total effect is defined as the expectation of true total effect variable $CTE_{D_X;10}(D_X) = \delta_{10}$ (see Def. 5.26) and in this example $U{=}D_X$.] However, only in the second and third example, the expectation of the Z-conditional *prima facie effects* is equal to the causal average total effect. Because this is no coincidence, this fact can be used for causal inference even in those cases in which the *unconditional* prima facie effects are biased, provided that the *conditional* prima facie effects are unbiased, that is, provided that $PFE_{Z;10}(z) = E(\delta_{10}|Z{=}z)$ for each value z of the covariate Z (see Th. 6.34).

Whether or not a causal average total effect is meaningful if there are different causal conditional total effects — some of which may even be negative, while some are positive — needs judgement with regards to content in the specific applications considered. In some applications it might be meaningful, in others it might not. Clearly, causal *conditional* total effects give more specific information than the causal *average* total effect. However, there are also advantages of causal average total effects. First, they give a summary evaluation of a treatment in *a single number* and different treatments may be compared to each other with respect to this number. Second, in samples of limited size, causal average effects can be estimated with more accuracy than the plenitude of causal conditional effects. And third, one should keep in mind that even causal conditional total effects are only causal average total effects (see, e. g., Table 6.4). Hence, it is always a matter of case-specific judgement how fine-grained the analysis should be.

First Conclusions

The three examples show that conditioning on a covariate of X does not *necessarily* yield unbiasedness given the values of the covariate. While there is no bias at all in the second example, the third example shows that conditioning *may* remove bias. Comparing Tables 6.3 and 6.4 to each other shows that unbiasedness of the conditional expectation values $E(Y|X{=}x, Z{=}z)$ relies on specific conditions. In these two tables, $P(X{=}1|U) = P(X{=}1|Z)$, that is, in these two tables there are equal individual treatment probabilities for all units with an identical value z of Z. Such conditions implying unbiasedness of $E(Y|X)$ or of $E(Y|X,Z)$ and their values are called *causality conditions*. Note, however, that there are several of such causality conditions that do *not* involve the U-conditional treatment probabilities (see ch. 9).

6.6 An Example With Accidental Unbiasedness

Now we treat an example demonstrating that there can be *unbiasedness of the conditional expectation* $E(Y|X)$ and at the same time *bias of the conditional expectation* $E(Y|X,Z)$, where Z is a covariate of X. This example shows that unbiasedness can be accidental, that is, there are cases in which unbiasedness is not a logical consequence of experimental design but an 'accident of numbers'. In chapter 8 we show that the experimental design technique of randomized assignment *always* induces unbiasedness of the conditional expectations $E(Y|X)$ and $E(Y|X,Z)$, whenever Z is a covariate of X, that is, whenever Z is measurable with respect to D_X.[3]

Table 6.5 displays the relevant parameters. We assume that it is a simple experiment having the same structure as the examples treated in section 6.5. That is,

$$\big((\Omega,\mathscr{A},P),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big),$$

as specified in section 6.5, is the regular probabilistic causality setup. The only difference is that $\Omega_1 = \Omega_U = \{Joe, Jim, Ann, Sue\}$ now consists of four (instead of six) persons. Again, $D_X = U$ is a global potential confounder of X.

Conditional Expectation Values $E(Y|X{=}x)$

In this specific example, the causal individual total effects are the same for all persons, namely 5, implying that the causal average total effect is also 5. The prima facie effect can be computed from the difference between the two conditional expectation values $E(Y|X{=}0)$ and $E(Y|X{=}1)$. In this example, Box 6.2 (i) yields

$$\begin{aligned}E(Y|X{=}0) &= \sum_u E(Y|X{=}0,U{=}u)\cdot P(U{=}u|X{=}0)\\ &= 95\cdot\frac{3}{16}+65\cdot\frac{1}{16}+80\cdot\frac{7}{16}+50\cdot\frac{5}{16} \;=\; 72.5\end{aligned}$$

and

$$\begin{aligned}E(Y|X{=}1) &= \sum_u E(Y|X{=}1,U{=}u)\cdot P(U{=}u|X{=}1)\\ &= 100\cdot\frac{5}{16}+70\cdot\frac{7}{16}+85\cdot\frac{1}{16}+55\cdot\frac{3}{16} \;=\; 77.5.\end{aligned}$$

Using Equation (ii) of Box 6.2, the corresponding expectations of the true outcome variables are

[3] Note again that unbiasedness does not refer to a sample and that there is no (successful) randomized assignment if there is systematic attrition, that is, if persons do not comply with the assignments of the experimenter.

Table 6.5. Accidental unbiasedness

Fundamental parameters

Person u	Sex z	$P(U=u)$	$P(X=1 \mid U=u)$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$CTE_{U;10}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$
Joe	m	1/4	5/8	95	100	5	3/16	5/16
Jim	m	1/4	7/8	65	70	5	1/16	7/16
Ann	f	1/4	1/8	80	85	5	7/16	1/16
Sue	f	1/4	3/8	50	55	5	5/16	3/16

	$x=0$	$x=1$	
$E(\tau_x)$:	72.5	77.5	$ATE_{10} = 5$
$E(Y \mid X=x)$:	72.5	77.5	$PFE_{10} = 5$
$E(\tau_x \mid Z=m)$:	80	85	$CTE_{Z;10}(m) = 5$
$E(Y \mid X=x, Z=m)$:	87.5	82.5	$PFE_{Z;10}(m) = -5$
$E(\tau_x \mid Z=f)$:	65	70	$CTE_{Z;10}(f) = 5$
$E(Y \mid X=x, Z=f)$:	67.5	62.5	$PFE_{Z;10}(f) = -5$

$$\begin{aligned} E(\tau_0) &= \sum_u E(Y \mid X=0, U=u) \cdot P(U=u) \\ &= 95 \cdot \frac{1}{4} + 65 \cdot \frac{1}{4} + 80 \cdot \frac{1}{4} + 50 \cdot \frac{1}{4} \;=\; 72.5 \end{aligned}$$

and

$$\begin{aligned} E(\tau_1) &= \sum_u E(Y \mid X=1, U=u) \cdot P(U=u) \\ &= 100 \cdot \frac{1}{4} + 70 \cdot \frac{1}{4} + 85 \cdot \frac{1}{4} + 55 \cdot \frac{1}{4} \;=\; 77.5. \end{aligned}$$

Hence, the conditional expectation values $E(Y \mid X=0)$ and $E(Y \mid X=1)$ are unbiased because they are identical to the corresponding expectations $E(\tau_0)$ and $E(\tau_1)$ of the true outcome variables and because the two true outcome variables are P-unique.

Conditional Expectation Values $E(Y \mid X=x, Z=z)$

The conditional expectation values $E(Y \mid X=1, Z=z)$ and $E(Y \mid X=0, Z=z)$ can be computed from the parameters displayed in Table 6.5 using Equation (iii) of Box 6.2. This equation holds because, in this example, the random variable Z is measurable with respect to U (see RS-Cor. 2.36). While the individual expected outcomes $E(Y \mid X=x, U=u)$ are displayed in Table 6.5, the conditional probabilities $P(U=u \mid X=x, Z=z)$ have to be computed via Equation (6.80) (see Exercise 6-13).

For $Z=m$ (males), Equation (iii) of Box 6.2 yields

$$\begin{aligned} E(Y|X=0,Z=m) &= \sum_u E(Y\,|X=0,U=u)\cdot P(U=u\,|X=0,Z=m) \\ &= 95\cdot\frac{9}{12}+65\cdot\frac{3}{12} \;+\; 80\cdot 0+50\cdot 0 \;=\; 87.5 \end{aligned}$$

and

$$\begin{aligned} E(Y|X=1,Z=m) &= \sum_u E(Y\,|X=1,U=u)\cdot P(U=u\,|X=1,Z=m) \\ &= 100\cdot\frac{5}{12}+70\cdot\frac{7}{12} \;+\; 85\cdot 0 \;+\; 55\cdot 0 \;=\; 82.5. \end{aligned}$$

In contrast, using Equation (iv) of Box 6.2, the $(Z=m)$-conditional expectation values of the true outcome variables are

$$\begin{aligned} E(\tau_0|Z=m) &= \sum_u E(Y\,|X=0,U=u)\cdot P(U=u\,|Z=m) \\ &= 95\cdot\frac{1}{2}+65\cdot\frac{1}{2} \;+\; 80\cdot 0+50\cdot 0 \;=\; 80 \end{aligned}$$

and

$$\begin{aligned} E(\tau_1|Z=m) &= \sum_u E(Y\,|X=1,U=u)\cdot P(U=u\,|Z=m) \\ &= 100\cdot\frac{1}{2}+70\cdot\frac{1}{2} \;+\; 85\cdot 0 \;+\; 55\cdot 0 \;=\; 85. \end{aligned}$$

For $Z=f$ (females), Equation (iii) of Box 6.2 yields

$$\begin{aligned} E(Y|X=0,Z=f) &= \sum_u E(Y|X=0,U=u)\cdot P(U=u\,|X=0,Z=f) \\ &= 95\cdot 0+65\cdot 0 \;+\; 80\cdot 7/12+50\cdot 5/12 \;=\; 67.5 \end{aligned}$$

and

$$\begin{aligned} E(Y|X=1,Z=f) &= \sum_u E(Y|X=1,U=u)\cdot P(U=u\,|X=1,Z=m) \\ &= 100\cdot 0+70\cdot 0 \;+\; 85\cdot 3/12+55\cdot 9/12 \;=\; 62.5. \end{aligned}$$

In contrast, using Equation (iv) of Box 6.2, the $(Z=f)$-conditional expectation values of the true outcome variables are

$$\begin{aligned} E(\tau_0|Z=f) &= \sum_u E(Y\,|X=0,U=u)\cdot P(U=u\,|Z=f) \\ &= 95\cdot 0+65\cdot 0 \;+\; 80\cdot\frac{1}{2}+50\cdot\frac{1}{2} \;=\; 65 \end{aligned}$$

and

$$\begin{aligned} E(\tau_1|Z=f) &= \sum_u E(Y\,|X=1,U=u)\cdot P(U=u\,|Z=f) \\ &= 100\cdot 0+70\cdot 0 \;+\; 85\cdot\frac{1}{2} \;+\; 55\cdot\frac{1}{2} \;=\; 70. \end{aligned}$$

Obviously, the conditional expectation values $E(Y|X=x,Z=z)$ of the outcome variable Y are not identical to the corresponding conditional expectation values $E(\tau_x|Z=z)$ of the true outcome variables. Hence, the $E(Y|X=x,Z=z)$ are biased although the conditional expectation values $E(Y|X=x)$ are unbiased.

Remark 6.54 [Methodological Implications] This example shows that unbiasedness of the conditional expectation values $E(Y|X{=}x)$ does not imply unbiasedness of the conditional expectation values $E(Y|X{=}x, Z{=}z)$, even if Z is a covariate (or potential confounder) of X. Hence, this example shows that *unbiasedness can be accidental*, that is, it may be a fortunate coincidence, an 'accident of numbers', not a logical consequence of experimental design. In chapter 8, however, we will show that experimental design techniques such as randomized assignment of the unit to one of the treatment conditions *always* leads to unbiasedness of the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x, Z{=}z)$, if Z is a covariate (or potential confounder) of X. Note, however, that this beneficial implication of randomized assignment *only applies to unbiasedness with respect to total effects*. Unfortunately, it does not apply to unbiasedness with respect to direct effects (see, e. g., Mayer et al., 2014). ◁

6.7 Summary and Conclusions

In this chapter we introduced the concepts of unbiasedness of conditional expectations such as $E(Y|X)$, $E(Y|X,Z)$, and $E^{X=x}(Y|Z)$. We also treated unbiasedness of the prima facie effects $E(Y|X{=}x) - E(Y|X{=}x')$ and $E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z)$, and of the conditional prima facie effect variables $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$. In these expressions, X is the putative cause variable, Y the outcome variable, and Z a covariate of X.

Unbiasedness

Unbiasedness is a first kind of causality conditions, which, together with the additional structural components listed in a regular probabilistic causality space, distinguishes a conditional expectation that has a causal meaning from an ordinary conditional expectation. Several kinds of conditional expectations and their values as well as their differences can be unbiased (see Box 6.1). The general insight of this chapter is that comparing conditional expectation values (true means) does not allow to draw any conclusions on the effects of a treatment or intervention *unless they are unbiased*. In terms of the metaphor discussed in the preface, conditional expectation values and their differences are like the shadow of the invisible man. The length of this shadow is identical to the height of the invisible man only under very specific conditions, in particular the angle between the sun and the surface of the earth at the point on which the man stands.

Identification

The unbiasedness conditions are the weakest assumptions under which we can *identify* causal average total effects and causal conditional total effects and effect functions. Box 6.3 summarizes the identification equations. Note that the right-hand sides of these equations are empirically estimable parameters or empirically estimable functions that can be computed from the conditional expectations $E(Y|X)$ or $E(Y|X,Z)$. That is, only the putative cause variable X, the outcome variable Y, and the random variables Z and V are involved. All other causality conditions that will be treated in the chapters to come imply unbiasedness, provided that true outcome theory applies. Some of these other causality conditions are empirically testable, at least in the sense of falsifiability. Unfortunately, this does not apply to unbiasedness itself.

Limitations

Hence, a first limitation of unbiasedness is that it cannot be tested empirically. Another drawback of unbiasedness has been exemplified by the numerical example displayed in Table 6.5. This example shows that, even if the conditional expectation values $E(Y|X{=}x)$ are unbiased and Z is a covariate of X, the conditional expectation values $E(Y|X{=}x, Z{=}z)$ and the $(Z{=}z)$-conditional prima facie effects can be biased (see also Greenland & Robins, 1986). In contrast, the sufficient conditions of unbiasedness treated in chapters 8 and 9 are less volatile, that is, they generalize to conditioning on a covariate Z of X. *Generalizability* and *falsifiability* are two important virtues of these alternative causality conditions.

6.8 Proofs

Proof of Theorem 6.9

In the proofs of all four equations, we assume that τ_x is P-unique.

Equation (6.6).

$$\begin{aligned}
& E(Y|X{=}x) \vdash D_X \\
\Leftrightarrow\quad & E(Y|X{=}x) = E(\tau_x) && \text{[Def. 6.3 (i)]} \\
\Leftrightarrow\quad & E^{X=x}(Y) = E(\tau_x) && \text{[(6.3)]} \\
\Leftrightarrow\quad & E^{X=x}\big(E^{X=x}(Y\,|D_X)\big) = E(\tau_x) && \text{[RS-Box 4.1 (iv)]} \\
\Leftrightarrow\quad & E^{X=x}(\tau_x) = E(\tau_x). && \text{[(5.1)]}
\end{aligned}$$

Equation (6.7).

$$\begin{aligned}
& E(Y|X{=}x) \vdash D_X \\
\Leftrightarrow\quad & E^{X=x}(\tau_x) = E(\tau_x) && \text{[(6.6)]} \\
\Leftrightarrow\quad & E(\tau_x|X{=}x) = E(\tau_x) && \text{[(6.3)]} \\
\Leftrightarrow\quad & \tau_x \vDash 1_{X=x} && \text{[RS-Th. 4.38 (ii)]} \\
\Leftrightarrow\quad & E(\tau_x|\,1_{X=x}) \underset{P}{=} E(\tau_x). && \text{[RS-(4.35)]}
\end{aligned}$$

Equation (6.9).

$$\begin{aligned}
& E(Y|X{=}x) \vdash D_X \\
\Leftrightarrow\quad & E^{X=x}(\tau_x) = E(\tau_x) && \text{[(6.6)]} \\
\Leftrightarrow\quad & E^{X=x}\big(\varepsilon_x + E(\tau_x|\,1_{X=x})\big) = E\big(\varepsilon_x + E(\tau_x|\,1_{X=x})\big) && \text{[(6.8)]} \\
\Leftrightarrow\quad & E^{X=x}(\varepsilon_x) + E^{X=x}\big(E(\tau_x|\,1_{X=x})\big) = E(\varepsilon_x) + E\big(E(\tau_x|\,1_{X=x})\big) && \text{[RS-Box 3.1 (vii)]} \\
\Leftrightarrow\quad & E^{X=x}(\varepsilon_x) + E^{X=x}\big(E(\tau_x)\big) = E(\varepsilon_x) + E(\tau_x) && \text{[(6.7), RS-Box 4.1 (iv)]} \\
\Leftrightarrow\quad & E^{X=x}(\varepsilon_x) + E(\tau_x) = E(\varepsilon_x) + E(\tau_x) && \text{[RS-Box 3.1 (i)]}
\end{aligned}$$

$$\Leftrightarrow \quad E^{X=x}(\varepsilon_x) = E(\varepsilon_x).$$

Equation (6.10).

$$\begin{aligned}
& E(Y \mid X{=}x) \vdash D_X & \\
\Leftrightarrow \quad & E^{X=x}(\varepsilon_x) = E(\varepsilon_x) & [(6.9)]\\
\Leftrightarrow \quad & E(\varepsilon_x \mid X{=}x) = E(\varepsilon_x) & [(6.3)]\\
\Leftrightarrow \quad & \varepsilon_x \vDash 1_{X=x} & [\text{RS-Th. 4.38 (ii)}]\\
\Leftrightarrow \quad & E(\varepsilon_x \mid 1_{X=x}) \underset{P}{=} E(\varepsilon_x). & [\text{RS-(4.35)}]
\end{aligned}$$

Proof of Theorem 6.16

Let g_x denote the factorization of $E^{X=x}(Y \mid Z)$ [see RS-Eq. (5.23)] and g denote the factorization of $E(\tau_x \mid Z)$ [see RS-Eq. (4.14)]. Furthermore, note that $P(X{=}x, Z{=}z) > 0$ implies $P(Z{=}z) > 0$.

$$\begin{aligned}
E^{X=x}(Y \mid Z) \vdash D_X \quad \Leftrightarrow \quad & E^{X=x}(Y \mid Z) \underset{P}{=} E(\tau_x \mid Z) & [\text{Def. 6.13 (i)}]\\
\Leftrightarrow \quad & g_x(Z) \underset{P}{=} g(Z) & [\text{RS-(4.14), RS-(5.23)}]\\
\Rightarrow \quad & g_x(z) = g(z) & [P(Z{=}z) > 0, \text{RS-(2.68)}]\\
\Leftrightarrow \quad & E^{X=x}(Y \mid Z{=}z) = E(\tau_x \mid Z{=}z) & [\text{RS-(5.24), RS-(4.17)}]\\
\Leftrightarrow \quad & E^{X=x}(Y \mid Z{=}z) \vdash D_X & [\text{Def. 6.13 (ii)}]\\
\Leftrightarrow \quad & E(Y \mid X{=}x, Z{=}z) \vdash D_X. & [(6.20)]
\end{aligned}$$

Proof of Theorem 6.17

Again, let g_x denote the factorization of $E^{X=x}(Y \mid Z)$ [see RS-Eq. (5.23)] and g denote the factorization of $E(\tau_x \mid Z)$ [see RS-Eq. (4.14)].

$$\begin{aligned}
& \forall z \in Z(\Omega)\colon\ E^{X=x}(Y \mid Z{=}z) \vdash D_X & \\
\Leftrightarrow \quad & \forall z \in Z(\Omega)\colon E^{X=x}(Y \mid Z{=}z) = E(\tau_x \mid Z{=}z) & [\text{Def. 6.13 (ii)}]\\
\Leftrightarrow \quad & \forall z \in Z(\Omega)\colon\ g_x(z) = g(z) & [\text{RS-(5.24), RS-(4.17)}]\\
\Rightarrow \quad & g_x(Z) \underset{P}{=} g(Z) & [\text{RS-(2.69)}]\\
\Rightarrow \quad & E^{X=x}(Y \mid Z) \underset{P}{=} E(\tau_x \mid Z) & [\text{RS-(5.23), RS-(4.14)}]\\
\Leftrightarrow \quad & E^{X=x}(Y \mid Z) \vdash D_X. & [\text{Def. 6.13 (i)}]
\end{aligned}$$

Proof of Theorem 6.19

$$
\begin{aligned}
& \forall x \in X(\Omega)\colon E^{Z=z}(Y \mid X=x) = E^{Z=z}(\tau_x) \\
\Leftrightarrow \quad & \forall x \in X(\Omega)\colon E^{X=x}(Y \mid Z=z) = E(\tau_x|Z=z) && [(6.15), \text{RS-}(3.24)] \\
\Leftrightarrow \quad & \forall x \in X(\Omega)\colon E^{X=x}(Y \mid Z=z) \vdash D_X && [\text{Def. } 6.13 \text{ (ii)}] \\
\Leftrightarrow \quad & \forall x \in X(\Omega)\colon E^{Z=z}(Y \mid X=x) \vdash D_X && [(6.25)] \\
\Leftrightarrow \quad & E^{Z=z}(Y \mid X) \vdash D_X. && [\text{Def. } 6.18 \text{ (ii)}]
\end{aligned}
$$

Proof of Theorem 6.20

Proposition (6.28). If Z is a covariate of X, then $\sigma(Z) \subset \sigma(D_X)$ [see Def. 4.11 (iv) and Rem. 4.16]. Hence,

$$
\begin{aligned}
& E^{X=x}(Y|Z) \vdash D_X \\
\Leftrightarrow \quad & E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z) && [\text{Def. } 6.13 \text{ (i)}] \\
\Leftrightarrow \quad & E^{X=x}\big(E^{X=x}(Y \mid D_X)\big|Z\big) \underset{P}{=} E(\tau_x|Z) && [\text{RS-Box } 4.1 \text{ (xiii)}] \\
\Leftrightarrow \quad & E^{X=x}(\tau_x|Z) \underset{P}{=} E(\tau_x|Z). && [(5.1)]
\end{aligned}
$$

Proposition (6.29).

$$
\begin{aligned}
& E^{X=x}(Y|Z) \vdash D_X \\
\Leftrightarrow \quad & E^{X=x}(\tau_x|Z) \underset{P}{=} E(\tau_x|Z) && [(6.28)] \\
\Leftrightarrow \quad & \tau_x \vDash 1_{X=x}|Z && [\text{RS-Th. } 5.50] \\
\Leftrightarrow \quad & E(\tau_x|\,1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z). && [\text{RS-}(4.48)]
\end{aligned}
$$

Proposition (6.31).

$$
\begin{aligned}
& E^{X=x}(Y|Z) \vdash D_X \\
\Leftrightarrow \quad & E^{X=x}(\tau_x|Z) \underset{P}{=} E(\tau_x|Z) && [(6.28)] \\
\Leftrightarrow \quad & E^{X=x}\big(\varepsilon_x + E(\tau_x|\,1_{X=x}, Z)\big|Z\big) \underset{P}{=} E\big(\varepsilon_x + E(\tau_x|\,1_{X=x}, Z)\big|Z\big) \\
& && [\text{RS-Box } 4.1 \text{ (xiv)}, (6.30)] \\
\Leftrightarrow \quad & E^{X=x}(\varepsilon_x|Z) + E^{X=x}\big(E(\tau_x|\,1_{X=x}, Z)\big|Z\big) \underset{P}{=} E(\varepsilon_x|Z) + E\big(E(\tau_x|\,1_{X=x}, Z)\big|Z\big) \\
& && [\text{RS-Box } 4.1 \text{ (xvii)}] \\
\Leftrightarrow \quad & E^{X=x}(\varepsilon_x|Z) + E^{X=x}\big(E(\tau_x|Z)\big|Z\big) \underset{P}{=} E(\varepsilon_x|Z) + E\big(E(\tau_x|Z)\big|Z\big) && [(6.29)] \\
\Leftrightarrow \quad & E^{X=x}(\varepsilon_x|Z) + E(\tau_x|Z) \underset{P}{=} E(\varepsilon_x|Z) + E(\tau_x|Z) && [\text{RS-Box } 4.1 \text{ (xi)}] \\
\Leftrightarrow \quad & E^{X=x}(\varepsilon_x|Z) \underset{P}{=} E(\varepsilon_x|Z).
\end{aligned}
$$

Proposition (6.32).

$$
\begin{aligned}
& E^{X=x}(Y|Z) \vdash D_X \\
\Leftrightarrow \quad & E^{X=x}(\varepsilon_x|Z) \underset{P}{=} E(\varepsilon_x|Z) && [(6.31)] \\
\Leftrightarrow \quad & \varepsilon_x \vDash 1_{X=x}|Z && [\text{RS-Th. } 5.50] \\
\Leftrightarrow \quad & E(\varepsilon_x|\,1_{X=x}, Z) \underset{P}{=} E(\varepsilon_x|Z). && [\text{RS-}(4.48)]
\end{aligned}
$$

Proof of Theorem 6.22

Equation (6.37). If Z is a covariate of X, then $\sigma(Z) \subset \sigma(D_X)$ [see Def. 4.11 (iv) and Rem. 4.16], which implies $\{Z{=}z\} = \{\omega \in \Omega\colon Z(\omega) = z\} \in \sigma(D_X)$. Also note that $E^{X=x}(Y|D_X) \underset{P}{=} E^{X=x}(Y|\mathscr{D}_{\mathscr{C}})$ [see Def. 4.11 (iii) and RS-Def. 4.4]. Hence,

$$
\begin{aligned}
& E^{X=x}(Y|Z{=}z) \vdash D_X \\
\Leftrightarrow\quad & E^{X=x}(Y|Z{=}z) = E(\tau_x|Z{=}z) && \text{[Def. 6.13 (ii)]} \\
\Leftrightarrow\quad & E^{X=x}(Y|\{Z{=}z\}) = E(\tau_x|Z{=}z) && \text{[RS-(3.23)]} \\
\Leftrightarrow\quad & E^{X=x}\big(E^{X=x}(Y|D_X)\,\big|\,\{Z{=}z\}\big) = E(\tau_x|Z{=}z) && [\{Z{=}z\} \in \sigma(D_X),\ \text{RS-Box 4.1 (xii)}] \\
\Leftrightarrow\quad & E^{X=x}(\tau_x|Z{=}z) = E(\tau_x|Z{=}z). && \text{[(5.1), RS-(3.23)]}
\end{aligned}
$$

Equation (6.38).

$$
\begin{aligned}
& E^{X=x}(Y|Z{=}z) \vdash D_X \\
\Leftrightarrow\quad & E^{X=x}(\tau_x|Z{=}z) = E(\tau_x|Z{=}z) && \text{[(6.37)]} \\
\Leftrightarrow\quad & E^{X=x}\big(\varepsilon_x + E(\tau_x|\,1_{X=x}, Z)\,\big|\,Z{=}z\big) = E\big(\varepsilon_x + E(\tau_x|\,1_{X=x}, Z)\,\big|\,Z{=}z\big) && \text{[(6.30)]} \\
\Leftrightarrow\quad & E^{X=x}(\varepsilon_x|Z{=}z) + E^{X=x}\big(E(\tau_x|\,1_{X=x}, Z)\,\big|\,Z{=}z\big) = E(\varepsilon_x|Z{=}z) + E\big(E(\tau_x|\,1_{X=x}, Z)\,\big|\,Z{=}z\big) \\
& && \text{[RS-Box 3.1 (vii)]} \\
\Leftrightarrow\quad & E^{X=x}(\varepsilon_x|Z{=}z) + E^{X=x}\big(E(\tau_x|\,1_{X=x}, Z)\,\big|\,\{Z{=}z\}\big) = E(\varepsilon_x|Z{=}z) + E\big(E(\tau_x|\,1_{X=x}, Z)\,\big|\,\{Z{=}z\}\big) \\
& && \text{[RS-(3.23)]} \\
\Leftrightarrow\quad & E^{X=x}(\varepsilon_x|Z{=}z) + E^{X=x}(\tau_x|Z{=}z) = E(\varepsilon_x|Z{=}z) + E(\tau_x|Z{=}z) \\
& && [\{Z{=}z\} \in \sigma(1_{X=x}, Z),\ \text{RS-Box 4.1 (xii), RS-(3.23)}] \\
\Leftrightarrow\quad & E^{X=x}(\varepsilon_x|Z{=}z) = E(\varepsilon_x|Z{=}z). && \text{[(6.37)]}
\end{aligned}
$$

Proof of Theorem 6.24

$$
\begin{aligned}
& E(Y|X{=}x) \vdash D_X \ \wedge\ E(Y|X{=}x') \vdash D_X \\
\Leftrightarrow\quad & E(Y|X{=}x) = E(\tau_x) \ \wedge\ E(Y|X{=}x') = E(\tau_{x'}) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P\text{-unique} && \text{[Def. 6.3 (i)]} \\
\Rightarrow\quad & E(Y|X{=}x) - E(Y|X{=}x') = E(\tau_x) - E(\tau_{x'}) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P\text{-unique} \\
\Leftrightarrow\quad & PFE_{xx'} = E(\tau_x) - E(\tau_{x'}) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P\text{-unique} && \text{[(6.39)]} \\
\Leftrightarrow\quad & PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}. && \text{[Def. 6.23 (i)]}
\end{aligned}
$$

Proof of Theorem 6.25

$$
E^{X=x}(Y|Z) \vdash D_X \ \wedge\ E^{X=x'}(Y|Z) \vdash D_X
$$

$$\Leftrightarrow \quad E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z) \ \wedge\ E^{X=x'}(Y|Z) \underset{P}{=} E(\tau_{x'}|Z) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P\text{-unique} \qquad \text{[Def. 6.13 (i)]}$$
$$\Rightarrow \quad E^{X=x}(Y|Z) - E^{X=x'}(Y|Z) \underset{P}{=} E(\tau_x|Z) - E(\tau_{x'}|Z) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P\text{-unique} \qquad \text{[SN-Rem. 2.76 (ii)]}$$
$$\Leftrightarrow \quad PFE_{xx'}(Z) = E(\tau_x|Z) - E(\tau_{x'}|Z) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P\text{-unique} \qquad \text{[(6.40)]}$$
$$\Leftrightarrow \quad PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}. \qquad \text{[Def. 6.23 (ii)]}$$

Proof of Theorem 6.26

$$E^{X=x}(Y|Z{=}z) \vdash D_X \ \wedge\ E^{X=x'}(Y|Z{=}z) \vdash D_X$$
$$\Leftrightarrow \quad E^{X=x}(Y|Z{=}z) = E(\tau_x|Z{=}z) \ \wedge\ E^{X=x'}(Y|Z{=}z) = E(\tau_{x'}|Z{=}z)$$
$$\wedge\ \tau_x, \tau_{x'} \text{ are } P^{Z=z}\text{-unique} \qquad \text{[Def. 6.13 (ii)]}$$
$$\Rightarrow \quad E^{X=x}(Y|Z{=}z) - E^{X=x'}(Y|Z{=}z) = E(\tau_x|Z{=}z) - E(\tau_{x'}|Z{=}z) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P^{Z=z}\text{-unique}$$
$$\Leftrightarrow \quad PFE_{Z;xx'}(z) = E(\tau_x - \tau_{x'}|Z{=}z) \ \wedge\ \tau_x, \tau_{x'} \text{ are } P^{Z=z}\text{-unique} \qquad \text{[RS-(3.35), (6.41)]}$$
$$\Leftrightarrow \quad PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}. \qquad \text{[Def. 6.23 (iii)]}$$

Proof of Theorem 6.34

$$\begin{aligned} E\big(PFE_{Z;xx'}(Z)\big) &= E\big(E(\tau_x - \tau_{x'}|Z)\big) && \text{[Def. 6.23 (ii)]} \\ &= E(\tau_x - \tau_{x'}) && \text{[RS-Box 4.1 (iv)]} \\ &= ATE_{xx'}. && \text{[(5.20), (5.21)]} \end{aligned}$$

Proof of Theorem 6.39

Equation (6.53).

$$\begin{aligned} & E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big) \\ \underset{P}{=}\ & E\big(E(\tau_x - \tau_{x'}|Z)\,\big|\,V\big) && \text{[Def. 6.23 (ii)]} \\ \underset{P}{=}\ & E\big(E(\tau_x|Z) - E(\tau_{x'}|Z)\,\big|\,V\big) && \text{[RS-Box 4.1 (xviii)]} \\ \underset{P}{=}\ & E\big(E(\tau_x|\,1_{X=x}, Z) - E(\tau_{x'}|\,1_{X=x'}, Z)\,\big|\,V\big) && \text{[(6.29)]} \\ \underset{P}{=}\ & E\big(E(\tau_x|\,1_{X=x}, Z)\,\big|\,V\big) - E\big(E(\tau_{x'}|\,1_{X=x'}, Z)\,\big|\,V\big) && \text{[RS-Box 4.1 (xviii)]} \\ \underset{P}{=}\ & E(\tau_x|V) - E(\tau_{x'}|V) && [\sigma(V) \subset \sigma(1_{X=x}, Z), \sigma(V) \subset \sigma(1_{X=x'}, Z), \text{RS-Box 4.1 (xiii)}] \\ \underset{P}{=}\ & E(\tau_x - \tau_{x'}|V) && \text{[RS-Box 4.1 (xviii)]} \\ \underset{P}{=}\ & CTE_{V;xx'}(V). && \text{[(5.34)]} \end{aligned}$$

Equation (6.54).

$$
\begin{aligned}
CTE_{V;xx'}(V) &\underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big) && [(6.53)] \\
&\underset{P}{=} E\big(E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)\,\big|\,V\big) && [(6.40)] \\
&\underset{P}{=} E\big(E^{X=x}(Y|Z)\,\big|\,V\big) - E\big(E^{X=x'}(Y|Z)\,\big|\,V\big). && [\text{RS-Box 4.1 (xviii)}]
\end{aligned}
$$

Proof of Theorem 6.41

Equation (6.59).

$$
\begin{aligned}
&E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big) \\
&\underset{P}{=} E\big(E(\tau_x - \tau_{x'}|Z)\,\big|\,V\big) && [(6.58), \text{Def. 6.23 (ii)}] \\
&\underset{P}{=} E\big(E(\tau_x|Z) - E(\tau_{x'}|Z)\,\big|\,V\big) && [\text{RS-Box 4.1 (xviii)}] \\
&\underset{P}{=} E\big(E(\tau_x|X,Z) - E(\tau_{x'}|X,Z)\,\big|\,V\big) && [(6.57)] \\
&\underset{P}{=} E\big(E(\tau_x|X,Z)\,\big|\,V\big) - E\big(E(\tau_{x'}|X,Z)\,\big|\,V\big) && [\text{RS-Box 4.1 (xviii)}] \\
&\underset{P}{=} E(\tau_x|V) - E(\tau_{x'}|V) && [\sigma(V) \subset \sigma(X,Z), \text{RS-Box 4.1 (xiii)}] \\
&\underset{P}{=} E(\tau_x - \tau_{x'}|V) && [\text{RS-Box 4.1 (xvii)}] \\
&\underset{P}{=} CTE_{V;xx'}(V). && [(5.30)]
\end{aligned}
$$

Equation (6.60).

$$
\begin{aligned}
&CTE_{V;xx'}(V) \\
&\underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big) && [(6.59)] \\
&\underset{P}{=} E\big(E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)\,\big|\,V\big) && [(6.40)] \\
&\underset{P}{=} E\big(E^{X=x}(Y|Z)\,\big|\,V\big) - E\big(E^{X=x'}(Y|Z)\,\big|\,V\big). && [\text{RS-Box 4.1 (xvii)}]
\end{aligned}
$$

Proof of Corollary 6.44

Both sets of assumptions, those of Theorem 6.39 and those of Theorem 6.41, yield

$$CTE_{V;xx'}(V) \underset{P}{=} E\big(PFE_{Z;xx'}(Z)\,\big|\,V\big).$$

Because both sides are compositions of V and some numerical functions, RS-Remark 2.55 implies Equation (6.64). The same kind of argument proves Equation (6.65).

6.9 Exercises

▷ **Exercise 6-1** What is the difference between the two terms $E(\tau_x)$ and $E(Y|X{=}x)$?

▷ **Exercise 6-2** Compute the probabilities $P(Z{=}z)$ occurring in Equation (6.81) for both values of Z in the example displayed in Table 6.2.

▷ **Exercise 6-3** Which are the probabilities $P(U{=}Tom, Z{=}m)$ and $P(U{=}Ann, Z{=}m)$ occurring in Equation (6.81) for the example displayed in Table 6.2.

▷ **Exercise 6-4** Compute the two conditional probabilities $P(U{=}u \,|\, Z{=}m)$ for $u = Tom$ and $u = Ann$ displayed in Table 6.2.

▷ **Exercise 6-5** Use RS-Theorem 1.38 to compute the probability $P(X{=}1)$ for the example displayed in Table 6.2.

▷ **Exercise 6-6** Compute the probabilities $P(U{=}u \,|\, X{=}0)$ and $P(U{=}u \,|\, X{=}1)$ displayed in Table 6.2 for all six persons.

▷ **Exercise 6-7** Compute the conditional probabilities $P(U{=}u \,|\, X{=}1, Z{=}m)$ occurring in Equation (iii) of Box 6.2 for the example of Table 6.2.

▷ **Exercise 6-8** Compute the conditional expectation values $E(Y \,|\, X{=}0)$ and $E(Y \,|\, X{=}1)$ for the example in Table 6.4.

▷ **Exercise 6-9** Compute the conditional expectation values $E(\tau_1 \,|\, Z{=}f)$ and $E(\tau_0 \,|\, Z{=}f)$ displayed in Table 6.2.

▷ **Exercise 6-10** Download *Kbook Table 8.4.sav* from *www.causal-effects.de*. This data set has been generated from Table 6.4 for a sample of size $N = 10{,}000$. Estimate the conditional expectations $E^{X=0}(Y \,|\, Z)$ and $E^{X=1}(Y \,|\, Z)$ and use them to compute the four conditional expectation values $E(Y \,|\, X{=}x, Z{=}z)$ displayed in Table 6.4.

▷ **Exercise 6-11** Prove the four equations presented in Box 6.2.

▷ **Exercise 6-12** Show that $\tau_x \models X \,|\, Z$ implies unbiasedness of $E^{X=x}(Y \,|\, Z)$ if τ_x is P-unique.

▷ **Exercise 6-13** Show that Equation (6.80) holds.

Solutions

▷ **Solution 6-1** The term $E(\tau_x)$ denotes the expectation of a true outcome variable τ_x, where x is a value of a putative cause variable. It is these true outcome variables that are of interest in the empirical sciences because the difference between τ_x and $\tau_{x'}$ is the conditional effect function of x compared to x' controlling for all potential confounders of X. This implies that the true outcome variables cannot be biased, and this also applies to their expectations $E(\tau_x)$. In causal research, we often aim at estimating the differences $E(\tau_x) - E(\tau_{x'})$. In contrast, the differences between the conditional expectation values $E(Y \,|\, X{=}x)$ and $E(Y \,|\, X{=}x')$ of the outcome variable Y are not of interest in causal research because they do not have a causal interpretation unless $E(Y \,|\, X{=}x) = E(\tau_x)$ and $E(Y \,|\, X{=}x') = E(\tau_{x'})$, that is, unless $E(Y \,|\, X{=}x)$ and $E(Y \,|\, X{=}x')$ are unbiased.

▷ **Solution 6-2** The events that U takes on the value u_i and that U takes on the value u_j, $i \neq j$, are disjoint. Therefore, we can use the theorem of total probability (see RS-Th. 1.38):

$$\begin{aligned}
P(Z{=}m) &= P(Z{=}m, U{=}Tom) + \ldots + P(Z{=}m, U{=}Sue) \\
&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + 0 + 0 \;=\; \frac{4}{6}. \\
P(Z{=}f) &= P(Z{=}f, U{=}Tom) + \ldots + P(Z{=}f, U{=}Sue) \\
&= 0 + 0 + 0 + 0 + \frac{1}{6} + \frac{1}{6} \;=\; \frac{2}{6}.
\end{aligned}$$

▷ **Solution 6-3** $P(U{=}Tom, Z{=}m) = 1/6$ and $P(U{=}Ann, Z{=}m) = 0$.

▷ **Solution 6-4**

$$\begin{aligned} P(U{=}Tom \,|\, Z{=}m) &= \frac{P(U{=}Tom, Z{=}m)}{P(Z{=}m)} = \frac{1/6}{4/6} = \frac{1}{4}. \\ P(U{=}Ann | Z{=}m) &= \frac{P(U{=}Ann, Z{=}m)}{P(Z{=}m)} = \frac{0}{4/6} = 0. \end{aligned}$$

▷ **Solution 6-5** The events $\{U{=}Tom\}, \ldots, \{U{=}Sue\}$ are disjoint and all these events have positive probabilities. Hence we can apply the theorem of total probability (see RS-Th. 1.38):

$$\begin{aligned} P(X{=}1) &= P(X{=}1 | U{=}Tom) \cdot P(U{=}Tom) + \ldots + P(X{=}1 | U{=}Sue) \cdot P(U{=}Sue) \\ &= \frac{6}{7} \cdot \frac{1}{6} + \frac{5}{7} \cdot \frac{1}{6} + \frac{4}{7} \cdot \frac{1}{6} + \frac{3}{7} \cdot \frac{1}{6} + \frac{2}{7} \cdot \frac{1}{6} + \frac{1}{7} \cdot \frac{1}{6} \\ &= \frac{21}{42} = \frac{1}{2}. \end{aligned}$$

▷ **Solution 6-6** We have to use the equation

$$P(U{=}u \,|\, X{=}x) = \frac{P(X{=}x \,|\, U{=}u) \cdot P(U{=}u)}{P(X{=}x)}.$$

For $u{=}Tom$ and $x{=}1$ this equation yields:

$$\begin{aligned} P(U{=}Tom \,|\, X{=}1) &= \frac{P(X{=}1 | U{=}Tom) \cdot P(U{=}Tom)}{P(X{=}1)} \\ &= \frac{6/7 \cdot 1/6}{1/2} = \frac{6}{21}. && \text{[Exercise 6-5]} \end{aligned}$$

Using the same procedure, we obtain $5/21, 4/21, \ldots, 1/21$, the corresponding probabilities for the other five persons $Tim, \ldots, Sue$, respectively. For $u{=}Tom$ and $x{=}0$, we obtain

$$\begin{aligned} P(U{=}Tom \,|\, X{=}0) &= \frac{P(X{=}0 \,|\, U{=}Tom) \cdot P(U{=}Tom)}{P(X{=}0)} \\ &= \frac{1/7 \cdot 1/6}{1/2} = \frac{1}{21}. && \text{[Exercise 6-5]} \end{aligned}$$

Using the same procedure, we obtain $2/21,\ 3/21, \ldots, 6/21$ for the other five persons $Tim, \ldots, Sue$, respectively.

▷ **Solution 6-7** According to Equation (6.80) we need the conditional probabilities $P(X{=}1 \,|\, U{=}u)$ displayed in Table 6.2. The other probabilities occurring in this equation can be computed from the probabilities displayed in the table.

One of these other probabilities that needs some computation is $P(X{=}1 \,|\, Z{=}m)$. For $x{=}1$ and $z{=}m$ in the example of Table 6.2, Equation (6.81) results in:

$$\begin{aligned} P(X{=}1 \,|\, Z{=}m) &= \frac{\sum_u P(X{=}1 \,|\, U{=}u) \cdot P(U{=}u, Z{=}m)}{P(Z{=}m)} \\ &= \frac{(6/7) \cdot (1/6) + \ldots + (3/7) \cdot (1/6) + (2/7) \cdot 0 + (1/7) \cdot 0}{4/6} = \frac{27}{42}. \end{aligned}$$

Using this result, Equation (6.80) yields:

$$\begin{aligned} P(U{=}Tom \,|\, X{=}1, Z{=}m) &= \frac{P(X{=}1 \,|\, U{=}Tom) \cdot P(U{=}Tom, Z{=}m)}{P(X{=}1 \,|\, Z{=}m) \cdot P(Z{=}m)} \\ &= \frac{(6/7) \cdot (1/6)}{(27/42) \cdot (4/6)} = \frac{6/7}{(27/42) \cdot 4} = \frac{6}{18}, \end{aligned}$$

as well as 5/18, 4/18, and 3/18 for the corresponding conditional probabilities for $u{=}Tim$, $u{=}Joe$, and $u{=}Jim$. The conditional probabilities $P(U{=}Ann \,|\, X{=}1, Z{=}m)$ and $P(U{=}Sue \,|\, X{=}1, Z{=}m)$ are zero.

▷ **Solution 6-8** According to Equation (i) of Box 6.2,

$$\begin{aligned} E(Y \,|\, X{=}0) &= \sum_u E(Y \,|\, X{=}0, U{=}u) \cdot P(U{=}u \,|\, X{=}0) \\ &= (68+78+88+98) \cdot \frac{1}{10} + (106+116) \cdot \frac{3}{10} \\ &= 33.2 + 66.6 \;=\; 99.8, \end{aligned}$$

and

$$\begin{aligned} E(Y \,|\, X{=}1) &= \sum_u E(Y \,|\, X{=}1, U{=}u) \cdot P(U{=}u \,|\, X{=}1) \\ &= (81+86+100+103) \cdot \frac{3}{14} + (114+130) \cdot \frac{1}{14} \\ &\approx 79.286 + 17.429 \;\approx\; 96.715. \end{aligned}$$

▷ **Solution 6-9** Remember again, in this example, $D_X = U$. Therefore, the values of the true outcome variable τ_x are the conditional expectation values $E^{X=x}(Y \,|\, U{=}u) = E(Y \,|\, X{=}x, U{=}u)$. Hence,

$$\begin{aligned} E(\tau_0 \,|\, Z{=}f) &= \sum_u E(Y \,|\, X{=}0, U{=}u) \cdot P(U{=}u \,|\, Z{=}f) \\ &= 68 \cdot 0 + \ldots + 98 \cdot 0 + 106 \cdot \frac{1}{2} + 116 \cdot \frac{1}{2} \;=\; 111. \\ E(\tau_1 \,|\, Z{=}f) &= \sum_u E(Y \,|\, X{=}1, U{=}u) \cdot P(U{=}u \,|\, Z{=}f) \\ &= 81 \cdot 0 + \ldots + 103 \cdot 0 + 114 \cdot \frac{1}{2} + 130 \cdot \frac{1}{2} \;=\; 122. \end{aligned}$$

▷ **Solution 6-10** Because Z is dichotomous, one way to estimate the conditional expectations $E^{X=0}(Y \,|\, Z)$ and $E^{X=1}(Y \,|\, Z)$ is to estimate the linear regressions of Y on the indicator $1_{Z=m}$ in treatments 0 and 1, that is, within the data subsamples with $x{=}0$ and $x{=}1$, respectively.

▷ **Solution 6-11** (i). We only have to prove $E(Y \,|\, X{=}x) = E(\tau_x \,|\, X{=}x)$ because the second equation is Equation (ii) of RS-Box 3.2. Hence, for $\tau_x \in \mathscr{E}^{X=x}(Y \,|\, U)$,

$$\begin{aligned} E(Y \,|\, X{=}x) &= E^{X=x}(Y) && \text{[RS-(3.24)]} \\ & E^{X=x}\big(E^{X=x}(Y \,|\, U)\big) && \text{[RS-Box 4.1 (iv)]} \\ & E^{X=x}(\tau_x) && [\tau_x = E^{X=x}(Y \,|\, U)] \\ & E(\tau_x \,|\, X{=}x). && \text{[RS-(3.24)]} \end{aligned}$$

(ii). Note that there is a measurable mapping g_x such that $E^{X=x}(Y \,|\, U) = g_x(U)$ [see Rem. 5.11]. Furthermore, under the assumptions of Box 6.2, in particular $P(X{=}x, U{=}u) > 0$, for all $u \in U(\Omega)$, the true outcome variable τ_x is P-unique. Hence, according to RS-Box 3.1 (v), for $\tau_x \in \mathscr{E}^{X=x}(Y \,|\, U)$,

$$\begin{aligned} E(\tau_x) &= E\big(E^{X=x}(Y \,|\, U)\big) && [\tau_x \underset{P}{=} E^{X=x}(Y \,|\, U)] \\ &= E\big(g_x(U)\big) && [E^{X=x}(Y \,|\, U) \underset{P}{=} g_x(U)] \\ &= \sum_u g_x(u) \cdot P(U{=}u) && \text{[RS-(3.13)]} \\ &= \sum_u E^{X=x}(Y \,|\, U{=}u) \cdot P(U{=}u) && \text{[RS-(5.24)]} \\ &= \sum_u E(Y \,|\, X{=}x, U{=}u) \cdot P(U{=}u). && \text{[RS-(5.26)]} \end{aligned}$$

(iii). If Z is measurable with respect to U and Ω'_Z is finite, then, according to RS-Corollary 2.36, there is a measurable mapping $f\colon (\Omega_U, \mathscr{P}(\Omega_U)) \to (\Omega'_Z, \mathscr{A}'_Z)$ such that $Z = f(U)$. Hence,

$$
\begin{aligned}
& P(X{=}x, U{=}u) > 0 \\
\Leftrightarrow\quad & P^{X=x}(U{=}u) > 0 && [P^{X=x}(U{=}u) = P(X{=}x, U{=}u)/P(X{=}x)] \\
\Rightarrow\quad & P^{X=x}(Z{=}z) > 0 && \Big[P^{X=x}(Z{=}z) = \sum_{u:\, f(u)=z} P^{X=x}(U{=}u)\Big] \\
\Leftrightarrow\quad & P(X{=}x, Z{=}z) > 0. && [P^{X=x}(Z{=}z) = P(X{=}x, Z{=}z)/P(X{=}x)]
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
& E^{X=x}(Y|Z) \underset{P^{X=x}}{=} E^{X=x}\big(E^{X=x}(Y|U)\,\big|\,Z\big) && \text{[RS-Box 4.1 (xiii)]} \\
\Rightarrow\quad & E^{X=x}(Y|Z{=}z) = E^{X=x}\big(E^{X=x}(Y|U)\,\big|\,Z{=}z\big) && [P^{X=x}(Z{=}z) > 0, \text{RS-Rem. 2.55}] \\
\Leftrightarrow\quad & E^{X=x}(Y|Z{=}z) = E^{X=x}\big(g_x(U)\,\big|\,Z{=}z\big) && \text{[RS-(5.23)]} \\
\Leftrightarrow\quad & E^{X=x}(Y|Z{=}z) = \sum_u g_x(u) \cdot P^{X=x}(U{=}u|Z{=}z) && \text{[RS-(3.28)]} \\
\Leftrightarrow\quad & E^{X=x}(Y|Z{=}z) = \sum_u E^{X=x}(Y|U{=}u) \cdot P^{X=x}(U{=}u|Z{=}z) && \text{[RS-(5.24)]} \\
\Leftrightarrow\quad & E(Y|X{=}x, Z{=}z) = \sum_u E(Y|X{=}x, U{=}u) \cdot P(U{=}u|X{=}x, Z{=}z). && \text{[RS-(5.26)]}
\end{aligned}
$$

Note that, in Box 6.2, we assume $\tau_x \underset{P}{=} E^{X=x}(Y|U)$. Therefore,

$$
E^{X=x}\big(E^{X=x}(Y|U)\,\big|\,Z{=}z\big) = E^{X=x}(\tau_x|Z{=}z) = E(\tau_x|X{=}x, Z{=}z),
$$

which, together with the equivalence propositions above, proves the first equation of (iii).

(iv).

$$
\begin{aligned}
E(\tau_x|Z{=}z) &= E\big(E^{X=x}(Y|U)\,\big|Z{=}z\big) && [\tau_x \underset{P}{=} E^{X=x}(Y|U)] \\
&= E\big(g_x(U)\big|Z{=}z\big) && \text{[RS-(5.23)]} \\
&= \sum_u g_x(u) \cdot P(U{=}u|Z{=}z) && \text{[RS-(3.28)]} \\
&= \sum_u E^{X=x}(Y|U{=}u) \cdot P(U{=}u|Z{=}z) && \text{[RS-(5.24)]} \\
&= \sum_u E(Y|X{=}x, U{=}u) \cdot P(U{=}u|Z{=}z). && \text{[RS-(5.26)]}
\end{aligned}
$$

▷ **Solution 6-12**

$$
\begin{aligned}
& \tau_x \vDash X\,|\,Z \\
\Leftrightarrow\quad & E(\tau_x|X,Z) \underset{P}{=} E(\tau_x|Z) && \text{[RS-Def. 4.41]} \\
\Rightarrow\quad & E\big(E(\tau_x|X,Z)\,\big|\,1_{X=x}, Z\big) \underset{P}{=} E\big(E(\tau_x|Z)\,\big|\,1_{X=x}, Z\big) && \text{[RS-Box 4.1 (xiv)]} \\
\Leftrightarrow\quad & E(\tau_x|\,1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z) && \text{[RS-Box 4.1 (xiii), (xi)]} \\
\Rightarrow\quad & E^{X=x}(Y|Z) \vdash D_X. && [\tau_x \text{ is } P\text{-unique, Th. 6.20 (i)}]
\end{aligned}
$$

▷ **Solution 6-13** In the examples presented in Tables 6.2 to 6.4, Z (sex) is U-measurable. According to RS-Corollary 2.36, there is a mapping $g\colon \Omega_U \to \{m, f\}$ such that Z is the composite function of U and g, that is, $Z = g(U)$. Therefore,

$$\{U{=}u, Z{=}z\} \;=\; \begin{cases} \{U{=}u\}, & \text{if } g(u) = z, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{6.82}$$

According to Equation (6.82), the event to sample person u and that the sampled person is male is identical to the event to sample person u, if that person is male [i. e., if $g(u) = m$]. Correspondingly, the event to sample person u and that the sampled person is female [i. e., $g(u) = f$] is identical to the event to sample person u, if that person is female. In contrast, the event to sample a male person u and to observe $Z(\omega) = g\big(U(\omega)\big) = f$ is the empty set. The same applies to the event to sample a female person u and to observe $Z(\omega) = g\big(U(\omega)\big) = m$. Hence, Equation (6.82) implies

$$P(U{=}u, Z{=}z) \;=\; \begin{cases} P(U{=}u), & \text{if } g(u) = z, \\ 0, & \text{otherwise,} \end{cases} \tag{6.83}$$

Therefore, we can conclude, that in these examples,

$$\begin{aligned} P(X{=}x \,|\, U{=}u, Z{=}z) \;&=\; \frac{P(X{=}x, U{=}u, Z{=}z)}{P(U{=}u, Z{=}z)} \\ &=\; \frac{P(X{=}x, U{=}u)}{P(U{=}u)} \\ &=\; P(X{=}x \,|\, U{=}u), \quad \text{if } P(U{=}u, Z{=}z) > 0. \end{aligned} \tag{6.84}$$

Furthermore, in our examples, $P(X{=}x, Z{=}z) > 0$. Therefore, if $P(U{=}u, Z{=}z) > 0$, then

$$\begin{aligned} P(U{=}u \,|\, X{=}x, Z{=}z) \;&=\; \frac{P(X{=}x, U{=}u, Z{=}z)}{P(X{=}x, Z{=}z)} \\ &=\; \frac{P(X{=}x \,|\, U{=}u, Z{=}z) \cdot P(U{=}u, Z{=}z)}{P(X{=}x \,|\, Z{=}z) \cdot P(Z{=}z)} \\ &=\; \frac{P(X{=}x \,|\, U{=}u) \cdot P(U{=}u, Z{=}z)}{P(X{=}x \,|\, Z{=}z) \cdot P(Z{=}z)}, \qquad \text{[Eq. (6.84)]} \end{aligned}$$

which is Equation (6.80). If $P(U{=}u, Z{=}z) = 0$, then $P(X{=}x, U{=}u, Z{=}z) = 0$. Hence, if $P(U{=}u, Z{=}z) = 0$, then

$$P(U{=}u \,|\, X{=}x, Z{=}z) \;=\; \frac{P(X{=}x, U{=}u, Z{=}z)}{P(X{=}x, Z{=}z)} \;=\; 0.$$

The same result is also obtained applying Equation (6.80).

Chapter 7
Rosenbaum-Rubin Conditions

In chapter 6, we introduced *unbiasedness* of various conditional expectations, conditional expectation values, prima facie effects, and prima facie effect functions. We also studied how various causal total effects and total effect functions can be identified by empirically estimable parameters and functions if we can assume that certain terms are unbiased. Those *unbiasedness conditions* are a first of several kinds of causality conditions, which, together with the structural components listed in a regular probabilistic causality setup, distinguish *causal total effects* from differences between conditional expectation values that have no causal meaning.

In this chapter we introduce some other causality conditions, all of them dealing with the relationship between the true outcome variables and the focused putative cause variable (the treatment, intervention, exposition variable). For simplicity, these causality conditions will summarily be referred to as the Rosenbaum-Rubin conditions. They include *strong ignorability* that has been introduced by Rosenbaum and Rubin (1983b), which is the most restrictive of all causality conditions treated in this chapter. Note, however, that we adapted the original Rosenbaum-Rubin condition by replacing their (deterministic) potential outcome variables by the (probabilistic) true outcome variables, which have been introduced in chapter 5.

Requirements

Reading this chapter we assume that the reader is familiar with the concepts treated in all chapters of Steyer (2024). Chapters 4 to 6 are now crucial, dealing with the concepts of a conditional expectation, a conditional expectation with respect to a conditional probability measure, and conditional independence. Furthermore, we assume familiarity with chapters 4 to 6 of the present book.

In the present chapter we will often refer to the following notation and assumptions.

Notation and Assumptions 7.1

(a) *Let* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *be a regular probabilistic causality setup and let* D_X *denote a global potential confounder of* X.

(b) *Let* $(\Omega'_X, \mathscr{A}'_X)$ *denote the value space of* X*, assume that the image* $X(\Omega)$ *of* Ω *under* X *is finite or countable. Let* $x \in \Omega'_X$, $\{x\} \in \mathscr{A}'_X$*, and let* $1_{X=x}$ *denote the indicator variable of the event* $\{X{=}x\} = \{\omega \in \Omega\colon X(\omega) = x\}$.

(c) *Let* Y *be real-valued with positive variance. Assume that* $P(X{=}x) > 0$*, define the probability measure* $P^{X=x}\colon \mathscr{A} \to [0,1]$ *by* $P^{X=x}(A) = P(A \mid X{=}x)$*, for all* $A \in \mathscr{A}$*, and let* $\tau_x = E^{X=x}(Y \mid D_X)$ *denote a true outcome variable of* Y *given* x.

(d) *Assume that τ_x is P-unique.*
(e) *Let $X(\Omega) = \{0,1,\ldots,J\}$ denote the image of Ω under X, for all $x \in X(\Omega)$, let $\{x\} \in \mathscr{A}'_X$ and assume $0 < P(X{=}x) < 1$. Finally, let $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$ denote the $(J+1)$-variate random variable consisting of the true outcome variables τ_x, $x \in X(\Omega)$.*
(f) *Assume that all τ_x, $x \in X(\Omega)$, are P-unique.*
(g) *Let Z be a random variable on $(\Omega, \mathscr{A}, P)$ and let $(\Omega'_Z, \mathscr{A}'_Z)$ denote its value space.*
(h) *Let $z \in \Omega'_Z$ be a value of Z, let $\{z\} \in \mathscr{A}'_Z$, assume $P(Z{=}z) > 0$, and define the probability measure $P^{Z=z}: \mathscr{A} \to [0,1]$ by $P^{Z=z}(A) = P(A \,|\, Z{=}z)$, for all $A \in \mathscr{A}$.*

7.1 RR-Conditions for $E(Y|X{=}x)$ and $E(Y|X)$

In chapter 6 we introduced *unbiasedness of* $E(Y|X{=}x)$, denoted $E(Y|X{=}x) \vdash D_X$ and defined by the conjunction of P-uniqueness of τ_x and $E(Y|X{=}x) = E(\tau_x)$. Presuming P-uniqueness of τ_x, we also showed that $E(Y|X{=}x) \vdash D_X$ is equivalent to *mean-independence of* τ_x *from* $1_{X=x}$, denoted $\tau_x \vDash 1_{X=x}$ and defined by $E(\tau_x|\, 1_{X=x}) = E(\tau_x)$ [see Th. 6.9 (i)]. In this section, we introduce more causality conditions for $E(Y|X{=}x)$ and $E(Y|X)$ involving the true outcome variables τ_x. These causality conditions are presented in Box 7.1 and the implication relations between them are summarized in Table 7.1.

7.1.1 Mean-Independence Conditions

The general concept of mean-independence has already been introduced in RS-Definition 4.36. However, in order to apply it to a true outcome variable τ_x we have to assume that $\tau_x = E^{X=x}(Y|D_X)$ is P-unique (see RS-Rem. 5.14, and RS-Th. 5.27). Hence, under the Assumptions 7.1 (a) to (d) we define

$$\tau_x \vDash 1_{X=x} \quad :\Leftrightarrow \quad E(\tau_x|\, 1_{X=x}) \underset{P}{=} E(\tau_x) \tag{7.1}$$

and call it *mean-independence of τ_x from $1_{X=x}$*. Correspondingly, *mean-independence of τ_x from X* is defined by

$$\tau_x \vDash X \quad :\Leftrightarrow \quad E(\tau_x|X) \underset{P}{=} E(\tau_x). \tag{7.2}$$

These are the first two causality conditions listed in Box 7.1.

Furthermore, under the Assumptions 7.1 (a) to (f) we define

$$\forall x\colon \tau_x \vDash 1_{X=x} \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; E(\tau_x|\, 1_{X=x}) \underset{P}{=} E(\tau_x) \tag{7.3}$$

and call it *mean-independence of τ_x from $1_{X=x}$ for all x*. Under the same assumptions, *mean-independence of all τ_x from X* is defined by

$$\forall x\colon \tau_x \vDash X \quad :\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; E(\tau_x|X) \underset{P}{=} E(\tau_x). \tag{7.4}$$

These are conditions (v) and (vi), respectively, of Box 7.1.

Remark 7.2 [Consequences of P-Uniqueness of the True Outcome Variables] If τ_x, τ_x^* are two versions of a true outcome variable, then P-uniqueness of τ_x does not only imply that the two versions are P-equivalent, that is, $\tau_x \underset{P}{=} \tau_x^*$ (see RS-Def. 2.46), but also that their expectations are identical and their X-conditional expectations are P-equivalent [see RS-Box 4.1 (xiv)]. Without P-uniqueness of τ_x, neither $\tau_x \underset{P}{=} \tau_x^*$, nor $E(\tau_x) = E(\tau_x^*)$, or $E(\tau_x|X) \underset{P}{=} E(\tau_x^*|X)$ are guaranteed. Hence, if a true outcome variable τ_x is not P-unique, then assuming $E(\tau_x|X) \underset{P}{=} E(\tau_x)$ does not make sense because $E(\tau_x)$ would not be uniquely defined. ◁

Remark 7.3 [τ_x Is the Composition of D_X and a Function g_x] At first sight, mean-independence of τ_x from X seems paradoxical, however, it is not. The reason is that $\tau_x = E^{X=x}(Y|D_X)$ is a function of the global potential confounder D_X. More precisely, τ_x is a composition $g_x(D_X)$ of D_X and a function $g_x\colon \Omega'_{D_X} \to \mathbb{R}$ [see RS-Eq. (5.23)], where $(\Omega'_{D_X}, \mathscr{A}'_{D_X})$ denotes the value space of D_X. Although τ_x refers to a specific value x of X, a true outcome variable τ_x is *not a function of* X. Therefore, postulating $E(\tau_x|X) \underset{P}{=} E(\tau_x)$ *does* make sense, provided that τ_x is P-unique. ◁

Remark 7.4 [Mean-Independence of τ_x From X Implies Unbiasedness of $E(Y|X{=}x)$] If the Assumptions 7.1 (a) to (d) hold, then

$$\tau_x \vDash X \quad\Rightarrow\quad \tau_x \vDash 1_{X=x} \tag{7.5}$$

$$\Leftrightarrow\quad E(Y|X{=}x) \vdash D_X \tag{7.6}$$

(see Exercise 7-1). Hence, mean-independence of τ_x from X implies mean-independence of τ_x from $1_{X=x}$, which itself is equivalent to unbiasedness of $E(Y|X{=}x)$ [see Th. 6.9 (i)]. However, unless X is dichotomous, $\tau_x \vDash X$ is not equivalent to $\tau_x \vDash 1_{X=x}$. More precisely, $\tau_x \vDash 1_{X=x}$ does not imply $\tau_x \vDash X$ unless $E(\tau_x|X) \underset{P}{=} E(\tau_x|\,1_{X=x})$. ◁

Remark 7.5 [Mean-Independence of All τ_x From X Implies Unbiasedness of $E(Y|X)$] If the Assumptions 7.1 (a) to (f) hold, then

$$\forall x\colon \tau_x \vDash X \quad\Rightarrow\quad \forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} \tag{7.7}$$

$$\Leftrightarrow\quad \forall\, x \in X(\Omega)\colon E(Y|X{=}x) \vdash D_X \tag{7.8}$$

$$\Leftrightarrow\quad E(Y|X) \vdash D_X \tag{7.9}$$

(see Exercise 7-2). Hence, mean-independence of all τ_x from X implies unbiasedness of all conditional expectation values $E(Y|X{=}x)$, and under the Assumptions 7.1 (a) to (f), this is equivalent to unbiasedness of the conditional expectation $E(Y|X)$. ◁

7.1.2 Independence Conditions

In this section, we treat some independence conditions involving the true outcome variables τ_x. These conditions also imply unbiasedness of $E(Y|X{=}x)$ and $E(Y|X)$. Remember, $X \perp\!\!\!\perp Y$ denotes *independence of two random variables X and Y* (see RS-Def. 2.59). Furthermore, $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$, that is, τ is a $J+1$-variate random variable consisting of the true outcome variables $\tau_x = E^{X=x}(Y|D_X)$, $x \in X(\Omega) = \{0, 1, \ldots, J\}$.

Box 7.1 Rosenbaum-Rubin conditions for $E(Y|X{=}x)$ and $E(Y|X)$

RR-Conditions implying unbiasedness of $E(Y|X{=}x)$

$\tau_x \vDash 1_{X=x}$ — *Mean-independence of τ_x from $1_{X=x}$*. If the Assumptions 7.1 (a) to (d) hold, then it is defined by

$$E(\tau_x|\, 1_{X=x}) \underset{P}{=} E(\tau_x). \tag{i}$$

$\tau_x \vDash X$ — *Mean-independence of τ_x from X*. If the Assumptions 7.1 (a) to (d) hold, then it is defined by

$$E(\tau_x|X) \underset{P}{=} E(\tau_x). \tag{ii}$$

$\tau_x \perp\!\!\!\perp 1_{X=x}$ — *Independence of τ_x and $1_{X=x}$*. If the Assumptions 7.1 (a) to (c) hold, then it is equivalent to

$$P(X{=}x|\tau_x) \underset{P}{=} P(X{=}x). \tag{iii}$$

$\tau_x \perp\!\!\!\perp X$ — *Independence of τ_x and X*. If the Assumptions 7.1 (a) to (c) hold, then it is equivalent to

$$\forall\, x' \in X(\Omega)\colon\; P(X{=}x'|\tau_x) \underset{P}{=} P(X{=}x'). \tag{iv}$$

The last two conditions are well-defined under the Assumptions 7.1 (a) to (c). However, only in conjunction with Assumption 7.1 (d), *each* of conditions (i) to (iv) implies $E(Y|X{=}x) \vdash D_X$.

RR-Conditions implying unbiasedness of $E(Y|X)$ and all $E(Y|X{=}x)$

The

$\forall x\colon \tau_x \vDash 1_{X=x}$ — *Mean-independence of τ_x from $1_{X=x}$ for all x*. If the Assumptions 7.1 (a) to (f) hold, then it is defined by

$$\forall\, x \in X(\Omega)\colon\; E(\tau_x|\, 1_{X=x}) \underset{P}{=} E(\tau_x). \tag{v}$$

$\forall x\colon \tau_x \vDash X$ — *Mean-independence of all τ_x from X*. If the Assumptions 7.1 (a) to (f) hold, then it is defined by

$$\forall\, x \in X(\Omega)\colon\; E(\tau_x|X) \underset{P}{=} E(\tau_x). \tag{vi}$$

$\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x}$ — *Independence of τ_x and $1_{X=x}$ for all x*. If the Assumptions 7.1 (a) to (c) and (e) hold, then it is equivalent to

$$\forall\, x \in X(\Omega)\colon\; P(X{=}x|\tau_x) \underset{P}{=} P(X{=}x). \tag{vii}$$

$\forall x\colon \tau_x \perp\!\!\!\perp X$ — *Independence of τ_x and X for all x*. If the Assumptions 7.1 (a) to (c) and (e) hold, then it is equivalent to

$$\forall\, x, x' \in X(\Omega)\colon\; P(X{=}x'|\tau_x) \underset{P}{=} P(X{=}x'). \tag{viii}$$

$\tau \perp\!\!\!\perp X$ — *Independence of τ and X*. If the Assumptions 7.1 (a) to (c) and (e) hold, then it is equivalent to

$$\forall\, x \in X(\Omega)\colon\; P(X{=}x|\tau) \underset{P}{=} P(X{=}x). \tag{ix}$$

last three conditions are well-defined under the Assumptions 7.1 (a) to (c) and (e). However, only in conjunction with Assumption 7.1 (f), *each* of conditions (v) to (ix) implies unbiasedness of $E(Y|X)$ and all $E(Y|X{=}x)$, $x \in X(\Omega)$.

Remark 7.6 [Equivalent Formulations of the Independence Conditions] These independence conditions and their symbols are listed in Box 7.1 [see conditions (iii), (iv), and (vii) to (ix)]. The assumptions specified in this box include that the image $X(\Omega)$ of Ω under X is finite or countable, and in this case the conditions specified in the table and on the right-hand sides of Propositions (7.10) to (7.14) are equivalent to the corresponding independence condition (see RS-Cors. 6.17 and 6.18). For convenience, we repeat these independence conditions and the conditions that are equivalent to them, provided that the image $X(\Omega)$ of Ω under X is finite or countable:

$$\tau_x \perp\!\!\!\perp 1_{X=x} \quad \Leftrightarrow \quad P(X=x|\tau_x) \underset{P}{=} P(X=x) \tag{7.10}$$

$$\tau_x \perp\!\!\!\perp X \quad \Leftrightarrow \quad \forall\, x' \in X(\Omega):\; P(X=x'|\tau_x) \underset{P}{=} P(X=x') \tag{7.11}$$

$$\forall x\colon \tau_x \vDash 1_{X=x} \quad \Leftrightarrow \quad \forall\, x \in X(\Omega):\; P(X=x|\tau_x) \underset{P}{=} P(X=x) \tag{7.12}$$

$$\forall x\colon \tau_x \perp\!\!\!\perp X \quad \Leftrightarrow \quad \forall\, x, x' \in X(\Omega):\; P(X=x'|\tau_x) \underset{P}{=} P(X=x') \tag{7.13}$$

$$\tau \perp\!\!\!\perp X \quad \Leftrightarrow \quad \forall\, x \in X(\Omega):\; P(X=x|\tau) \underset{P}{=} P(X=x). \tag{7.14}$$

Note again that independence of two random variables X and Y is also defined if neither X nor Y are finite or countable (see RS-Def. 2.59). However, because we assume that X is finite or countable, the propositions on the right-hand sides above are more convenient and more intuitive than the general definition. ◁

7.1.3 Implications Among RR-Conditions for $E(Y|X=x)$ *and* $E(Y|X)$

Table 7.1 displays the implications among the causality conditions listed in Box 7.1. Note that the propositions summarized in this table are special cases of the corresponding propositions presented in Table 7.2. (For proofs see Exercises 7-3 and 7-4.)

In the sequel we study some consequences of some of the independence conditions, starting with the consequences of $\tau_x \perp\!\!\!\perp 1_{X=x}$, that is, of independence of the true outcome variable τ_x and the indicator variable $1_{X=x}$. According to Theorem 7.7, under the Assumptions 7.1 (a) to (d), $\tau_x \perp\!\!\!\perp 1_{X=x}$ implies *mean-independence of* τ_x *from* $1_{X=x}$, which itself is equivalent to unbiasedness of $E(Y|X=x)$.

Theorem 7.7 [$\tau_x \perp\!\!\!\perp 1_{X=x}$ **Implies Unbiasedness of** $E(Y|X=x)$]
Let the Assumptions 7.1 (a) to (d) hold. Then

$$\tau_x \perp\!\!\!\perp 1_{X=x} \quad \Rightarrow \quad \tau_x \vDash 1_{X=x} \tag{7.15}$$

$$\Leftrightarrow \quad E(Y|X=x) \vdash D_X. \tag{7.16}$$

(Proof p. 220)

Hence, under the assumptions of this theorem, independence of τ_x and $1_{X=x}$ implies mean-independence of τ_x from $1_{X=x}$, which is equivalent to unbiasedness of the conditional expectation value $E(Y|X=x)$.

According to the next theorem, independence of the true outcome variable τ_x and the putative cause variable X implies $\tau_x \perp\!\!\!\perp 1_{X=x}$. If we additionally assume P-uniqueness of τ_x, then it also implies mean-independence of τ_x from X and from the indicator variable $1_{X=x}$, which is equivalent to unbiasedness of $E(Y|X=x)$ [see Prop. (7.16)].

Table 7.1. Implications among RR-conditions for $E(Y|X{=}x)$ and $E(Y|X)$

	$\tau_x \vDash \mathit{1}_{X=x}$	$\tau_x \vDash X$	$\tau_x \perp\!\!\!\perp \mathit{1}_{X=x}$	$\tau_x \perp\!\!\!\perp X$	$\forall x\colon \tau_x \vDash \mathit{1}_{X=x}$	$\forall x\colon \tau_x \vDash X$	$\forall x\colon \tau_x \perp\!\!\!\perp \mathit{1}_{X=x}$	$\forall x\colon \tau_x \perp\!\!\!\perp X$
$\tau_x \vDash X$	(a)-(d)							
$\tau_x \perp\!\!\!\perp \mathit{1}_{X=x}$	(a)-(d)							
$\tau_x \perp\!\!\!\perp X$	(a)-(d)	(a)-(d)	(a)-(c)					
$\forall x\colon \tau_x \vDash \mathit{1}_{X=x}$	(a)-(f)							
$\forall x\colon \tau_x \vDash X$	(a)-(f)	(a)-(f)						
$\forall x\colon \tau_x \perp\!\!\!\perp \mathit{1}_{X=x}$	(a)-(e)	(a)-(e)	(a)-(c),(e)		(a)-(f)			
$\forall x\colon \tau_x \perp\!\!\!\perp X$	(a)-(e)	(a)-(e)	(a)-(c),(e)	(a)-(c),(e)	(a)-(f)	(a)-(f)	(a)-(c),(e)	
$\tau \perp\!\!\!\perp X$	(a)-(e)	(a)-(e)	(a)-(c),(e)	(a)-(c),(e)	(a)-(f)	(a)-(f)	(a)-(c),(e)	(a)-(c),(e)

Note: An entry such as (a)-(d) means that the condition in the row implies the condition in the column, provided that the Assumptions 7.1 (a) to (d) hold. The symbols involving $\vDash$ or $\perp\!\!\!\perp$ are explained in Box 7.1. Trivial equivalences such as $\tau_x \perp\!\!\!\perp X \Leftrightarrow \tau_x \perp\!\!\!\perp X$ are omitted. The first three conditions imply unbiasedness of $E(Y|X{=}x)$, provided that the Assumptions 7.1 (a) to (d) hold and, under the Assumptions 7.1 (a) to (f), the last five imply unbiasedness of $E(Y|X)$ and all $E(Y|X{=}x)$, $x \in X(\Omega)$.

Theorem 7.8 [Consequences of $\tau_x \perp\!\!\!\perp X$]
Let the Assumptions 7.1 (a) to (c) hold. Then:

$$\tau_x \perp\!\!\!\perp X \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp \mathit{1}_{X=x}\,. \tag{7.17}$$

If we additionally assume 7.1 (d), then

$$\begin{aligned} \tau_x \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau_x \vDash X && (7.18)\\ &\Rightarrow \quad \tau_x \vDash \mathit{1}_{X=x} && (7.19)\\ &\Leftrightarrow \quad E(Y|X{=}x) \vdash D_X\,. && (7.20)\end{aligned}$$

(Proof p. 221)

Remark 7.9 [$\tau \perp\!\!\!\perp X$ implies $\tau_x \perp\!\!\!\perp \mathit{1}_{X=x}$] Remember, according to RS-Corollary 6.18,

$$P(X{=}x|\tau_x) \underset{P}{=} P(X{=}x) \quad \Leftrightarrow \quad \tau_x \perp\!\!\!\perp \mathit{1}_{X=x}\,. \tag{7.21}$$

As mentioned before, even if we assume $\tau_x \perp\!\!\!\perp X$ for all $x \in X(\Omega)$, then this is less restrictive than $\tau \perp\!\!\!\perp X$. More precisely, if the Assumptions 7.1 (a) to (c) and (e) hold, then

$$\tau \perp\!\!\!\perp X \quad \Rightarrow \quad \big(\forall\, x \in X(\Omega): \; \tau_x \perp\!\!\!\perp X\big), \tag{7.22}$$

which follows from $\sigma(\tau_x) \subset \sigma(\tau)$ and RS-Box 2.1 (iv). Note that the term on the right-hand side of Proposition (7.22) does not imply $\tau \perp\!\!\!\perp X$. ◁

The following theorem summarizes some important consequences of $\tau \perp\!\!\!\perp X$. These consequences include unbiasedness of $E(Y|X)$ and all its values $E(Y|X=x)$. Note that this theorem is a special case of Theorem 7.21 for Z being a constant map.

Theorem 7.10 [Consequences of $\tau \perp\!\!\!\perp X$]
If the Assumptions 7.1 (a) to (c) and (e) hold, then

$$\begin{aligned} \tau \perp\!\!\!\perp X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} &(7.23)\\ &\Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X. &(7.24)\end{aligned}$$

If we additionally assume 7.1 (f), then

$$\begin{aligned} \forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \quad &\Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vDash X &(7.25)\\ &\Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} &(7.26)\\ &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E(Y|X=x) \vdash D_X &(7.27)\\ &\Leftrightarrow \quad E(Y|X) \vdash D_X. &(7.28)\end{aligned}$$

Remark 7.11 [Methodological Consequences] In empirical studies, there is neither a *direct* way to create one of the causality conditions treated in this section, nor to check if these conditions hold, unless the values of the true outcome variables are known. Hence, from a practical perspective, these conditions are useless. However, in chapter 8, we will treat another causality condition, $D_X \perp\!\!\!\perp X$, that implies $\tau \perp\!\!\!\perp X$ and P-uniqueness of all true outcome variables τ_x, $x \in X(\Omega)$. The condition $D_X \perp\!\!\!\perp X$ can be created via randomized assignment of the observational unit (e.g., the person) to a treatment condition x. Hence, there is an *indirect* way to create $\tau \perp\!\!\!\perp X$, and with it, unbiasedness of the conditional expectation $E(Y|X)$ and all its values $E(Y|X=x)$, $x \in X(\Omega)$. ◁

7.2 RR-Conditions for $E^{Z=z}(Y|X=x)$ and $E^{Z=z}(Y|X)$

Remember, $\tau \perp\!\!\!\perp X$ means independence of τ and X *with respect to the probability measure* P on the measurable space $(\Omega, \mathscr{A})$. A more explicit notation is $\tau \underset{P}{\perp\!\!\!\perp} X$. Hence, all propositions of Theorem 7.10 hold with respect to P. For example, $\tau_x \vDash X$ denotes mean-independence of τ_x from X with respect to P, which may also be denoted by $\tau_x \underset{P}{\vDash} X$. Such an explicit notation is necessary if we want to use Theorem 7.10 for another probability measure on $(\Omega, \mathscr{A})$ such as $P^{Z=z}\colon \mathscr{A} \to [0,1]$ [see Ass. 7.1 (h)].

Remark 7.12 [$(Z=z)$-Conditional Independence of X and Y] Hence, if X, Y, Z are random variables on $(\Omega, \mathscr{A}, P)$ and $P(Z=z) > 0$, then

$$\begin{aligned} X \perp\!\!\!\perp Y \,|\, (Z=z) \quad &:\Leftrightarrow \quad X \underset{P^{Z=z}}{\perp\!\!\!\perp} Y \\ &\Leftrightarrow \quad P^{Z=z}(A \cap B) = P^{Z=z}(A) \cdot P^{Z=z}(B), \quad \forall\, A, B \in \sigma(X) \times \sigma(Y) \end{aligned} \tag{7.29}$$

defines *$(Z=z)$-conditional independence of X and Y*. By definition, it is equivalent to *independence of X and Y with respect to the probability measure $P^{Z=z}$ on* $(\Omega, \mathscr{A})$. ◁

Remark 7.13 [$(Z{=}z)$**-Conditional Mean-Independence of** Y **From** X] Furthermore, if Y is numerical or with finite expectation $E^{Z=z}(Y)$ with respect to the measure $P^{Z=z}$, then

$$Y \vDash X \,|\, Z{=}z \quad :\Leftrightarrow \quad \underset{P^{Z=z}}{Y \vDash X} \quad \Leftrightarrow \quad E^{Z=z}(Y \,|\, X) \underset{P^{Z=z}}{=} E^{Z=z}(Y) \tag{7.30}$$

defines $(Z{=}z)$*-conditional mean-independence of Y from X*, which, by definition, is equivalent to *mean-independence of Y from X with respect to the probability measure* $P^{Z=z}$. ◁

Using this notation, Theorem 7.10 implies the following corollary about the consequences of $(Z{=}z)$-conditional independence of τ and X. Reading this corollary, note that P-uniqueness of τ_x implies $P^{Z=z}$-uniqueness of τ_x (see RS-Box 5.1) and that unbiasedness of the terms $E^{Z=z}(Y \,|\, X)$ and $E^{Z=z}(Y \,|\, X{=}x)$, $x \in X(\Omega)$, is defined only if Z is a covariate of X.

Corollary 7.14 [Consequences of $(Z{=}z)$**-Conditional Independence of** τ **and** X**]**
If the Assumptions 7.1 (a) to (c), (e), (g), and (h) hold, then

$$\begin{aligned} \tau \perp\!\!\!\perp X \,|\, (Z{=}z) \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} | (Z{=}z) && (7.31)\\ &\Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \,|\, (Z{=}z)\,. && (7.32)\end{aligned}$$

If we additionally assume that all τ_x, $x \in X(\Omega)$, *are* $P^{Z=z}$*-unique, then*

$$\begin{aligned} \forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \,|\, (Z{=}z) \quad &\Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vDash X \,|\, (Z{=}z) && (7.33)\\ &\Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} | (Z{=}z). && (7.34)\end{aligned}$$

If we additionally assume that Z *is a covariate of* X*, then*

$$\begin{aligned} \forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} | (Z{=}z) \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{Z=z}(Y \,|\, X{=}x) \vdash D_X && (7.35)\\ &\Leftrightarrow \quad E^{Z=z}(Y \,|\, X) \vdash D_X\,. && (7.36)\end{aligned}$$

Hence, according to this theorem $(Z{=}z)$-conditional independence of τ and X implies, among other things, unbiasedness of the X-conditional expectation of Y with respect to the probability measure $P^{Z=z}$ [see Prop. (7.36)] and unbiasedness of all its values $E^{Z=z}(Y \,|\, X{=}x) = E(Y \,|\, X{=}x, Z{=}z)$, $x \in X(\Omega)$ [see Prop. (7.35)], provided that Z is a covariate of X.

7.3 RR-Conditions for $E^{X=x}(Y \,|\, Z)$ and $E(Y \,|\, X, Z)$

Now we generalize the causality conditions treated in section 7.1, conditioning on a random variable Z on $(\Omega, \mathscr{A}, P)$. Under the appropriate assumptions, which include that Z is a covariate of X, these conditions imply unbiasedness of the conditional expectation $E(Y \,|\, X, Z)$ and all Z-conditional expectations $E^{X=x}(Y \,|\, Z)$, $x \in X(\Omega)$. These causality conditions can indirectly be created by conditionally randomized assignment of the observational unit to a treatment condition, but also via covariate selection (see chs. 8 to 10).

Box 7.2 lists all causality conditions treated in this section, including their symbols and definitions, and Table 7.2 summarizes the implications among them. The most restrictive of these causality conditions is Z-conditional independence of a $(J+1)$-variate true outcome variable τ and X, which is the translation (into true outcome theory) of *strong ignorability* (Rosenbaum & Rubin, 1983b).

7.3.1 Conditional Mean-Independence Conditions

The first condition in Box 7.2 is *Z-conditional mean-independence of τ_x from $1_{X=x}$*, denoted $\tau_x \vDash 1_{X=x}|Z$. Under the Assumptions 7.1 (a) to (d) and (g), it is defined by

$$\tau_x \vDash 1_{X=x}|Z \quad :\Leftrightarrow \quad E(\tau_x|\, 1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z). \tag{7.37}$$

According to Theorem 6.20 (i), this condition is equivalent to unbiasedness of $E^{X=x}(Y|Z)$, provided that τ_x is P-unique and Z is a covariate of X.

The second condition in Box 7.2 is *Z-conditional mean-independence of τ_x from X*, denoted $\tau_x \vDash X|Z$. Under the Assumptions 7.1 (a) to (d) and (g), it is defined by

$$\tau_x \vDash X|Z \quad :\Leftrightarrow \quad E(\tau_x|X,Z) \underset{P}{=} E(\tau_x|Z). \tag{7.38}$$

Remark 7.15 [Dichotomous X] If the Assumptions 7.1 (a) to (d) and (g) hold, then X being dichotomous implies

$$E(\tau_x|X,Z) \underset{P}{=} E(\tau_x|\, 1_{X=x}, Z) \underset{P}{=} E(\tau_x|\, 1_{X\neq x}, Z) \tag{7.39}$$

because $\sigma(X,Z) = \sigma(1_{X=x}, Z) = \sigma(1_{X\neq x}, Z)$ (see RS-Def. 4.4). Hence, under these assumptions, X being dichotomous implies

$$\tau_x \vDash X|Z \quad \Leftrightarrow \quad \tau_x \vDash 1_{X=x}|Z \quad \Leftrightarrow \quad \tau_x \vDash 1_{X\neq x}|Z. \tag{7.40}$$

If we additionally assume that Z is a covariate of X, then X being dichotomous also implies

$$\tau_x \vDash X|Z \quad \Leftrightarrow \quad E(Y|X,Z) \vdash D_X. \tag{7.41}$$

$\lhd$

7.3.2 Conditional Independence Conditions

Conditions three and four in Box 7.2 are *Z-conditional independence of τ_x and $1_{X=x}$*, denoted $\tau \perp\!\!\!\perp 1_{X=x}|Z$, and *$Z$-conditional independence of τ_x and X*, denoted $\tau \perp\!\!\!\perp X|Z$. Remember that the general concept of conditional independence of two random variables X and Y given a random variable Z, denoted $X \perp\!\!\!\perp Y|Z$, has been introduced in RS-Definition 6.2. However, under the Assumptions 7.1 (a) to (c) and (g),

$$\tau_x \perp\!\!\!\perp 1_{X=x}|Z \quad \Leftrightarrow \quad P(X{=}x|Z,\tau_x) \underset{P}{=} P(X{=}x|Z), \tag{7.42}$$

and

$$\tau_x \perp\!\!\!\perp X|Z \quad \Leftrightarrow \quad \forall\, x' \in X(\Omega):\; P(X{=}x'|\tau_x, Z) \underset{P}{=} P(X{=}x'|Z). \tag{7.43}$$

Box 7.2 Conditional Rosenbaum-Rubin conditions

RR-Conditions implying unbiasedness of $E^{X=x}(Y|Z)$

$\tau_x \vDash 1_{X=x}|Z$ — *Z-conditional mean-independence of τ_x from $1_{X=x}$*. Under the Assumptions 7.1 (a) to (d) and (g), it is defined by

$$E(\tau_x|\,1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z)\,. \tag{i}$$

$\tau_x \vDash X|\,Z$ — *Z-conditional mean-independence of τ_x from X*. Under the Assumptions 7.1 (a) to (d) and (g), it is defined by

$$E(\tau_x|X, Z) \underset{P}{=} E(\tau_x|Z)\,. \tag{ii}$$

$\tau_x \perp\!\!\!\perp 1_{X=x}|Z$ — *Z-conditional independence of τ_x and $1_{X=x}$*. Under the Assumptions 7.1 (a) to (c) and (g), it is equivalent to

$$P(X{=}x\,|\,\tau_x, Z) \underset{P}{=} P(X{=}x\,|\,Z). \tag{iii}$$

$\tau_x \perp\!\!\!\perp X|Z$ — *Z-conditional independence of τ_x and X*. Under the Assumptions 7.1 (a) to (c) and (g), it is equivalent to

$$x' \in X(\Omega):\; P(X{=}x'\,|\,\tau_x, Z) \underset{P}{=} P(X{=}x'\,|\,Z). \tag{iv}$$

If we additionally assume 7.1 (d) and Z is a covariate of X, then each of conditions (i) to (iv) implies $E^{X=x}(Y|Z) \vdash D_X$.

RR-Conditions implying unbiasedness of $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$

$\forall x\colon \tau_x \vDash 1_{X=x}|Z$ — *Z-conditional mean-independence of τ_x from $1_{X=x}$ for all x*. Under the Assumptions 7.1 (a) to (g), it is defined by

$$\forall\, x \in X(\Omega):\; E(\tau_x|\,1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z)\,. \tag{v}$$

$\forall x\colon \tau_x \vDash X|Z$ — *Z-conditional mean-independence of all τ_x from X*. Under the Assumptions 7.1 (a) to (g), it is defined by

$$\forall\, x \in X(\Omega):\; E(\tau_x|X, Z) \underset{P}{=} E(\tau_x|Z)\,. \tag{vi}$$

$\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x}|Z$ — *Z-conditional independence of τ_x and $1_{X=x}$ for all x*. Under the Assumptions 7.1 (a) to (c), (e) and (g), it is equivalent to

$$\forall\, x \in X(\Omega):\; P(X{=}x\,|\,\tau_x, Z) \underset{P}{=} P(X{=}x\,|\,Z). \tag{vii}$$

$\forall x\colon \tau_x \perp\!\!\!\perp X|Z$ — *Z-conditional independence of τ_x and X for all x*. Under the Assumptions 7.1 (a) to (c), (e) and (g), it is equivalent to

$$\forall\, x, x' \in X(\Omega):\; P(X{=}x'\,|\,\tau_x, Z) \underset{P}{=} P(X{=}x'\,|\,Z). \tag{viii}$$

$\tau \perp\!\!\!\perp X|Z$ — *Z-conditional independence of τ and X (strong ignorability)*. Under the Assumptions 7.1 (a) to (c), (e) and (g), it is equivalent to

$$\forall\, x \in X(\Omega):\; P(X{=}x\,|\,\tau, Z) \underset{P}{=} P(X{=}x\,|\,Z). \tag{ix}$$

If we additionally assume 7.1 (f) and Z is a covariate of X, then *each* of conditions (v) to (ix) implies $E(Y|X,Z) \vdash D_X$ and $E^{X=x}(Y|Z) \vdash D_X$, for all $x \in X(\Omega)$.

Also note that $\tau_x \perp\!\!\!\perp X|Z$ implies $\tau_x \perp\!\!\!\perp 1_{X=x}|Z$ but not vice versa, unless X is dichotomous.

Conditions (v) to (viii) of Box 7.2 postulate that conditions (i) to (iv) hold for *all* values x of X. Under the appropriate assumptions including that Z is a covariate of X, these conditions imply unbiasedness of $E(Y|X,Z)$ and $E^{X=x}(Y|Z)$, for all values $x \in X(\Omega)$. Note that condition (viii) has been proposed by Porta (2014, p. 142).

Finally, the last condition in Box 7.2 is *Z-conditional independence of τ and X*, denoted $\tau \perp\!\!\!\perp X|Z$, where $\tau = (\tau_0, \tau_1, \dots, \tau_J)$ is a $J+1$-dimensional random variable consisting of the true outcome variables τ_x, $x \in X(\Omega)$. Under the Assumptions 7.1 (a) to (c), (e) and (g), it is equivalent to

$$\tau \perp\!\!\!\perp X|Z \quad \Leftrightarrow \quad \forall\, x \in X(\Omega): \ P(X{=}x|Z,\tau) \underset{P}{=} P(X{=}x|Z) \tag{7.44}$$

(see RS-Theorem 6.5). Hence, assuming Z-conditional independence of τ and X is equivalent to assuming Z-conditional independence of the events $\{X{=}x\}$ from τ, for all $x \in X(\Omega)$.

Remark 7.16 [Strong Ignorability] Note that $\tau \perp\!\!\!\perp X|Z$ is the translation of Rosenbaum and Rubin's *strong ignorability* into true outcome theory. This condition has been presented by Rosenbaum and Rubin (1983b) and plays a crucial role in Rubin's potential outcome approach to causal effects. Also note that the additional assumption $P(X{=}x|Z) \underset{P}{>} 0$, for all $x \in X(\Omega)$, of Rosenbaum and Rubin (1983b) follows from P-uniqueness of τ_x, for all $x \in X(\Omega)$, which is Assumption 7.1 (f) (see Exercise 7-6). As mentioned before, P-uniqueness of a true outcome variable τ_x is equivalent to

$$P(X{=}x|D_X) \underset{P}{>} 0 \tag{7.45}$$

[see RS-Th. 5.27 (ii)]. If we assume that Z is a covariate of X, then $P(X{=}x|D_X) \underset{P}{>} 0$ implies $P(X{=}x|Z) \underset{P}{>} 0$. However, $P(X{=}x|Z) \underset{P}{>} 0$ does not imply $P(X{=}x|D_X) \underset{P}{>} 0$, also not in conjunction with $\tau_x \perp\!\!\!\perp X|Z$ or $\tau \perp\!\!\!\perp X|Z$, unless $P(X{=}x|Z) \underset{P}{=} P(X{=}x|D_X)$. ◁

7.3.3 Implications Among RR-Conditions for $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$

Table 7.2 displays the implications among the causality conditions listed in Box 7.2. In this section, we present some theorems in which most of these implications are proved. The solution to Exercise 7-4 provide a guide to the proofs of all these implications.

In the following theorem, we present some propositions about the consequences of $\tau_x \vDash X|Z$ concerning unbiasedness.

Theorem 7.17 [Consequences of Conditional Mean-Independence of τ_x From X]
If the Assumptions 7.1 (a) to (d) and (g) hold, then

$$\tau_x \vDash X|Z \quad \Rightarrow \quad \tau_x \vDash 1_{X=x}|Z\,. \tag{7.46}$$

If we additionally assume that Z is a covariate of X, then

$$\tau_x \vDash 1_{X=x}|Z \quad \Leftrightarrow \quad E^{X=x}(Y|Z) \vdash D_X\,. \tag{7.47}$$

(Proof p. 221)

Hence, under the assumptions of Theorem 7.17, if τ_x is Z-conditionally mean-independent from X, then it is also Z-conditionally mean-independent from an indicator $1_{X=x}$ for one of the values x of the putative cause variable X, which itself is equivalent to unbiasedness of $E^{X=x}(Y|Z)$, provided that we also assume that Z is a covariate of X.

In the following corollary we consider the consequences of assuming that, *for all* $x \in X(\Omega)$, the true outcome variable τ_x is Z-conditionally mean-independent from X.

Corollary 7.18 [Consequences of Conditional Mean-Independence of τ_x From $1_{X=x}$]
If the Assumptions 7.1 (a) to (d) and (g) hold, then

$$\forall\, x \in X(\Omega)\colon \tau_x \vDash X \,|\, Z \quad \Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} | Z\,. \tag{7.48}$$

If we additionally assume that Z is a covariate of X, then

$$\forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} | Z \quad \Leftrightarrow \quad E(Y|X,Z) \vdash D_X\,. \tag{7.49}$$

(Proof p. 221)

Hence, if we assume Z-conditional mean-independence of τ_x from X for all values x of X, and Z is a covariate of X, then $E(Y|X,Z)$ is unbiased, provided that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique.

Now we consider conditional (stochastic) independence of a true outcome variable τ_x from an indicator variable $1_{X=x}$ for a value x of the putative cause variable X. According to the following theorem, if the Assumptions 7.1 (a) to (d) and (g) hold, then $\tau_x \perp\!\!\!\perp 1_{X=x}|Z$ implies Z-conditional mean-independence of τ_x from $1_{X=x}$, and if Z is a covariate of X, then it also implies unbiasedness of $E^{X=x}(Y|Z)$.

Theorem 7.19 [Consequences of Conditional Independence of τ_x From $1_{X=x}$]
If the Assumptions 7.1 (a) to (d) and (g) hold, then

$$\tau_x \perp\!\!\!\perp 1_{X=x}|Z \quad \Rightarrow \quad \tau_x \vDash 1_{X=x}|Z\,. \tag{7.50}$$

If we additionally assume that Z is a covariate of X, then

$$\tau_x \vDash 1_{X=x}|Z \quad \Leftrightarrow \quad E^{X=x}(Y|Z) \vdash D_X\,. \tag{7.51}$$

(Proof p. 222)

Again, note that Proposition (7.50) holds for any value x of X for which τ_x is P-unique. Hence, if τ_x is P-unique for all $x \in X(\Omega)$, then Proposition (7.50) holds for all $x \in X(\Omega)$. Correspondingly, if τ_x is P-unique for all $x \in X(\Omega)$ and Z is a covariate of X, then Proposition (7.51) holds for all $x \in X(\Omega)$.

In Theorem 7.19 we considered Z-conditional independence of τ_x and an indicator variable $1_{X=x}$. In the following theorem we turn to Z-conditional independence of τ_x and X, which may take on more than just two different values.

Table 7.2. Implications among RR-conditions for $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$

	$\tau_x \models 1_{X=x} \mid Z$	$\tau_x \models X \mid Z$	$\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	$\tau_x \perp\!\!\!\perp X \mid Z$	$\forall x\colon \tau_x \models 1_{X=x} \mid Z$	$\forall x\colon \tau_x \models X \mid Z$	$\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	$\forall x\colon \tau_x \perp\!\!\!\perp X \mid Z$
$\tau_x \models X \mid Z$	(a)-(d), (g)							
$\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	(a)-(d), (g)							
$\tau_x \perp\!\!\!\perp X \mid Z$	(a)-(d), (g)	(a)-(d), (g)	(a)-(c), (g)					
$\forall x\colon \tau_x \models 1_{X=x} \mid Z$	(a)-(g)							
$\forall x\colon \tau_x \models X \mid Z$	(a)-(g)	(a)-(g)						
$\forall x\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	(a)-(e), (g)		(a)-(c), (e), (g)		(a)-(g)			
$\forall x\colon \tau_x \perp\!\!\!\perp X \mid Z$	(a)-(e), (g)	(a)-(e), (g)	(a)-(c), (e), (g)	(a)-(c), (e), (g)	(a)-(g)	(a)-(g)	(a)-(c), (e), (g)	
$\tau \perp\!\!\!\perp X \mid Z$	(a)-(e), (g)	(a)-(e), (g)	(a)-(c), (e), (g)	(a)-(c), (e), (g)	(a)-(g)	(a)-(g)	(a)-(c), (e), (g)	(a)-(c), (e), (g)

Note: An entry such as (a)-(g) means that the condition in the row implies the condition in the column, provided that we assume 7.1 (a) to (g). The symbols involving $\models$ or $\perp\!\!\!\perp$ are explained in Box 7.2. Trivial equivalences such as $\tau_x \perp\!\!\!\perp X \mid Z \Leftrightarrow \tau_x \perp\!\!\!\perp X \mid Z$ are omitted. If we additionally assume that Z is a covariate of X, then the first three conditions listed in the first column of the table imply unbiasedness of $E^{X=x}(Y|Z)$, provided that the Assumptions 7.1 (a) to (d) and (g) hold. The last five conditions imply unbiasedness of $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, if Z is a covariate of X and we assume 7.1 (a) to (g).

Theorem 7.20 [Consequences of $\tau_x \perp\!\!\!\perp X \mid Z$]
If the Assumptions 7.1 (a) to (d) and (g) hold, then

$$\tau_x \perp\!\!\!\perp X \mid Z \quad \Rightarrow \quad \tau_x \vDash X \mid Z\,. \tag{7.52}$$

Furthermore,

$$\tau_x \perp\!\!\!\perp X \mid Z \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z \tag{7.53}$$
$$\Rightarrow \quad \tau_x \vDash 1_{X=x} \mid Z\,. \tag{7.54}$$

(Proof p. 222)

The following theorem summarizes some important consequences of $\tau \perp\!\!\!\perp X \mid Z$. If Z is a covariate of X, then these consequences include unbiasedness of the conditional expectation $E(Y \mid X, Z)$, which is equivalent to unbiasedness of all conditional expectations $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$.

Theorem 7.21 [Consequences of Conditional Independence of τ and X and More]
If the Assumptions 7.1 (a) to (c), (e), and (g) hold, then

$$\tau \perp\!\!\!\perp X \mid Z \quad \Leftrightarrow \quad \forall\, x \in X(\Omega):\ \tau \perp\!\!\!\perp 1_{X=x} \mid Z \tag{7.55}$$
$$\Rightarrow \quad \forall\, x, x' \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x'} \mid Z \tag{7.56}$$
$$\Leftrightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp X \mid Z \tag{7.57}$$
$$\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z\,. \tag{7.58}$$

Furthermore, if the Assumptions 7.1 (a) to (g) hold, then

$$\forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp X \mid Z \quad \Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \vDash X \mid Z \tag{7.59}$$
$$\Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \vDash 1_{X=x} \mid Z \tag{7.60}$$

and

$$\forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z \quad \Rightarrow \quad \forall\, x \in X(\Omega):\ \tau_x \vDash 1_{X=x} \mid Z\,. \tag{7.61}$$

If we additionally assume that Z is a covariate of X, then

$$\forall\, x \in X(\Omega):\ \tau_x \vDash 1_{X=x} \mid Z \quad \Leftrightarrow \quad \forall\, x \in X(\Omega):\ E^{X=x}(Y \mid Z) \vdash D_X \tag{7.62}$$
$$\Leftrightarrow \quad E(Y \mid X, Z) \vdash D_X\,. \tag{7.63}$$

(Proof p. 222)

Remark 7.22 [Methodological Consequences] What has been said in Remark 7.11 about the causality conditions for $E(Y \mid X)$ and its values $E(Y \mid X{=}x)$ also applies to the causality conditions for $E(Y \mid X, Z)$ and for the conditional expectations $E^{X=x}(Y \mid Z)$. There is no *direct* way to create the conditions such as $\tau \perp\!\!\!\perp X \mid Z$ or $\forall x\colon \tau_x \vDash X \mid Z$. There is also no direct way to select the (possibly multivariate) random variable Z such that these conditions hold. However, in chapters 8 to 10 we will treat other causality conditions that imply $\tau \perp\!\!\!\perp X \mid Z$ and $\forall x\colon \tau_x \vDash X \mid Z$. These conditions can be created via conditionally randomized assignment of the observational unit (person) to a treatment condition, but also via an appropriate selection of Z. Hence, there are *indirect* ways to create $\tau \perp\!\!\!\perp X|Z$ and $\forall x\colon \tau_x \vDash X|Z$, and with them unbiasedness of the conditional expectations $E(Y|X,Z)$ and the conditional expectations $E^{X=x}(Y|Z)$, $x \in X(\Omega)$. Remember, it is these conditional expectations that can be estimated in data samples. ◁

7.4 Examples

In this section, we study two examples. In the first one, there is Z-conditional mean-independence of τ_x from $1_{X=x}$ for all values x of X. In the second one, Z-conditional independence of τ from X holds.

7.4.1 Z-Conditional Mean-Independence of τ_x From $1_{X=x}$ for All x

Table 7.3 displays the parameters of a random experiment in which Z-conditional mean-independence of τ_x from $1_{X=x}$ holds for each value x of X. However, neither Z-conditional mean-independence of all τ_x from X holds nor Z-conditional independence of τ from X.

We assume that this random experiment has the same structure as the random experiments treated in section 6.5. That is,

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big),$$

as specified in section 6.5, is the regular probabilistic causality setup. The only difference is that $\Omega_2 = \Omega_X = \{treatment\ 0, treatment\ 1, treatment\ 3\}$ now consists of three (instead of two) treatment conditions. Again, U is a global potential confounder of X.

The upper left part of Table 7.3 displays the true outcomes under treatments 0, 1, and 2, the probabilities $P(U=u)$ for each person u to be sampled, as well as the conditional probabilities $P(X=1|U=u)$ and $P(X=2|U=u)$ to be assigned to treatment 1 and treatment 2, respectively. All other parameters, such as the associated individual causal total effects $CTE_{U;10}(u)$ and $CTE_{U;20}(u)$ or the conditional probabilities $P(U=u|X=x)$, can be computed from those 'fundamental parameters'. The table also displays the values of the covariate (potential confounder) $Z = sex$. To emphasize, the table does not contain sample data; it displays the parameters describing the laws of a random experiment, the single-unit trial, which consists of (a) sampling a person from the set of (the six) persons, (b) assigning or registering the (self-) assignment of the person to one of the three treatment conditions, and (c) observing the value of the outcome variable. (For more details on such singl-unit trials see ch. 2).

Z-Conditional Mean-Independence of τ_x From $1_{X=x}$

In this random experiment, the true treatment probabilities and true outcomes are such that $\tau_x \vDash 1_{X=x}|Z$ holds for each of the three values x of X. By definition, $\tau_x \vDash 1_{X=x}|Z$ is equivalent to

$$E(\tau_x|\, 1_{X=x}, Z) \underset{P}{=} E(\tau_x|Z), \tag{7.64}$$

[see Eq. (7.37)]. In this example, $U = D_X$, $\tau_x = E^{X=x}(Y|U)$, and τ_x is P-unique. Because $Z(\Omega) = \{m, f\}$ and $P(X=x, Z=z) > 0$ for all values of X and Z, according to RS-Corollary

Table 7.3. Z-conditional mean-independence of τ_x from $1_{X=x}$ for all values x of X

Fundamental parameters

Person u	Sex z	$P(U=u)$	$P(X=1 \mid U=u)$	$P(X=2 \mid U=u)$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$E^{X=2}(Y \mid U=u)$	$CTE_{U;10}(u)$	$CTE_{U;20}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$	$P(U=u \mid X=2)$
Tom	m	1/6	3/5	1/3	75	87	97	12	22	4/63	12/59	1/6
Tim	m	1/6	1/2	1/3	70	80	92	10	22	10/63	10/59	1/6
Joe	m	1/6	3/5	1/3	65	73	107	8	42	4/63	12/59	1/6
Ann	f	1/6	1/2	1/3	106	114	90	8	−16	10/63	10/59	1/6
Sue	f	1/6	1/4	1/3	116	130	120	14	4	25/63	5/59	1/6
Eva	f	1/6	1/2	1/3	126	146	126	20	0	10/63	10/59	1/6

	$x=0$	$x=1$	$x=2$		$x=1$	$x=2$
$E(\tau_x)$:	93	105	105.333	ATE_{x0}:	12	12.333
$E(Y \mid X=x)$:	102.857	101.186	105.333	PFE_{x0}:	−1.671	2.476
$E(\tau_x \mid Z=m)$:	70	80	98.667	$CTE_{Z;x0}(m)$:	10	28.667
$E(Y \mid X=x, Z=m)$:	70	80	98.667	$PFE_{Z;x0}(m)$:	10	28.667
$E(\tau_x \mid Z=f)$:	116	130	112	$CTE_{Z;x0}(f)$:	14	−4
$E(Y \mid X=x, Z=f)$:	116	130	112	$PFE_{Z;x0}(f)$:	14	−4

5.51, RS-Equations (5.49) and (5.50), Equation (7.64) holds if and only if

$$E(\tau_x|X=x,Z=z) \;=\; E(\tau_x|Z=z), \quad \forall\, z \in \{m, f\}. \tag{7.65}$$

The values of the true outcome variable $\tau_x = E^{X=x}(Y|U) = g_x(U)$ [see RS-Eq. (5.23)] are the conditional expectation values $E^{X=x}(Y|U=u) = E(Y|X=x, U=u)$. Now,

$$\begin{aligned} E(\tau_x|X=x,Z=z) &= E\big(g_x(U)\,\big|\,X=x,Z=z\big) \\ &= \sum_u g_x(u)\cdot P(U=u\,|X=x,Z=z) \\ &= \sum_u E^{X=x}(Y|U=u)\cdot P(U=u\,|X=x,Z=z), \quad \forall\, z \in \{m, f\}. \end{aligned} \tag{7.66}$$

In contrast,

$$\begin{aligned} E(\tau_x\,|Z=z) &= E\big(g_x(U)\,|Z=z\big) \\ &= \sum_u g_x(u)\cdot P(U=u\,|Z=z) \\ &= \sum_u E^{X=x}(Y|U=u)\cdot P(U=u\,|Z=z), \quad \forall\, z \in \{m, f\}. \end{aligned} \tag{7.67}$$

According to Equation (7.66), the conditional expectation value $E(\tau_1|X=1, Z=z)$ can be computed from Tables 7.3 and 7.4 by

Table 7.4. Conditional probabilities $P(U{=}u\,|\,X{=}x, Z{=}z)$ supplementing Table 7.3

Person u	$P(U{=}u\,|\,X{=}0, Z{=}m)$	$P(U{=}u\,|\,X{=}1, Z{=}m)$	$P(U{=}u\,|\,X{=}2, Z{=}m)$	$P(U{=}u\,|\,X{=}0, Z{=}f)$	$P(U{=}u\,|\,X{=}1, Z{=}f)$	$P(U{=}u\,|\,X{=}2, Z{=}f)$
Tom	2/9	6/17	1/3	0	0	0
Tim	5/9	5/17	1/3	0	0	0
Joe	2/9	6/17	1/3	0	0	0
Ann	0	0	0	2/9	2/5	1/3
Sue	0	0	0	5/9	1/5	1/3
Eva	0	0	0	2/9	2/5	1/3

$$\begin{aligned} E(\tau_1|X{=}1, Z{=}m) &= \sum_u E^{X=1}(Y|U{=}u) \cdot P(U{=}u\,|\,X{=}1, Z{=}m) \\ &= 87 \cdot \frac{6}{17} + 80 \cdot \frac{5}{17} + 73 \cdot \frac{6}{17} = 80. \end{aligned}$$

This is exactly the same result as obtained for

$$\begin{aligned} E(\tau_1|Z{=}m) &= \sum_u E^{X=1}(Y|U{=}u) \cdot P(U{=}u\,|\,Z{=}m) \\ &= (87 + 80 + 73) \cdot \frac{1}{3} = 80. \end{aligned}$$

Hence,

$$E(\tau_1|X{=}1, Z{=}m) = E(\tau_1|Z{=}m) = 80,$$

and the corresponding equations hold for the second value of Z, namely

$$E(\tau_1|X{=}1, Z{=}f) = E(\tau_1|Z{=}f) = 130.$$

Therefore, according to RS-Corollary 5.51, we proved $\tau_1 \vDash 1_{X=1}|Z$. In the same way it can be shown that $\tau_0 \vDash 1_{X=0}|Z$ and $\tau_2 \vDash 1_{X=2}|Z$ hold as well.

Z-Conditional Mean-Independence of All τ_x From X

Now we show that $\tau_x \vDash X|Z$ does not hold in the example of Table 7.3 for at least one value x of X. Of course, this implies that $\forall x\colon \tau_x \vDash X|Z$ does not hold either. We start gathering the relevant equations. By definition, $\forall x\colon \tau_x \vDash X|Z$ is equivalent to

$$E(\tau_x|X,Z) \underset{P}{=} E(\tau_x|Z), \quad \forall\, x \in X(\Omega) \tag{7.68}$$

[see Eq. (7.38)]. As mentioned before, in this example U is a global potential confounder of X, the image of Ω under X is $X(\Omega) = \{0,1,2\}$, and $\tau_x = E^{X=x}(Y|U)$ is P-unique for all

$x \in X(\Omega)$. Hence because $P(X{=}x, Z{=}z) > 0$ for all pairs (x,z) of values of (X,Z), Equation (7.68) is equivalent to

$$E(\tau_x|X{=}x', Z{=}z) \;=\; E(\tau_x|Z{=}z), \quad \forall\,(x,x',z) \in \{0,1,2\}^2 \times \{m,f\} \tag{7.69}$$

[see RS-Th. 4.42 (ii)].

The values of the true outcome variable $\tau_x = E^{X=x}(Y|U) = g_x(U)$ are the conditional expectation values $E^{X=x}(Y|U{=}u) = E(Y|X{=}x, U{=}u)$. Hence, applying RS-Equation (3.28) to the left-hand side of Equation (7.69) yields

$$\begin{aligned} E(\tau_x|X{=}x', Z{=}z) \;&=\; E\big(g_x(U)\,\big|\,X{=}x', Z{=}z\big)\\ =& \sum_u g_x(u)\cdot P(U{=}u\,|X{=}x', Z{=}z)\\ =& \sum_u E^{X=x}(Y|U{=}u)\cdot P(U{=}u\,|X{=}x', Z{=}z), \quad \forall\,(x,x',z) \in \{0,1,2\}^2 \times \{m,f\}. \end{aligned} \tag{7.70}$$

In contrast, applying RS-Equation (3.28) to the right-hand side of Equation (7.69) yields

$$\begin{aligned} E(\tau_x|Z{=}z) \;&=\; E\big(g_x(U)\,|\,Z{=}z\big)\\ &= \sum_u g_x(u)\cdot P(U{=}u\,|Z{=}z)\\ &= \sum_u E^{X=x}(Y|U{=}u)\cdot P(U{=}u\,|Z{=}z), \quad \forall\,(x,z) \in \{0,1,2\} \times \{m,f\}. \end{aligned} \tag{7.71}$$

In order to show that $\forall x\colon \tau_x \vDash X|Z$ does not hold in this example, it suffices to show that there is a triple $(x,x',z) \in \{0,1,2\}^2 \times \{m,f\}$ for which $E(\tau_x|X{=}x', Z{=}z) \neq E(\tau_x|Z{=}z)$ [see Eq. (7.69)].

According to Equation (7.70), the conditional expectation value $E(\tau_2|X{=}1, Z{=}f)$ can be computed from Tables 7.3 and 7.4 by

$$\begin{aligned} E(\tau_2|X{=}1, Z{=}f) \;&=\; \sum_u E^{X=2}(Y|U{=}u)\cdot P(U{=}u\,|X{=}1, Z{=}f)\\ &= 90\cdot\frac{2}{5} + 120\cdot\frac{1}{5} + 126\cdot\frac{2}{5} \;=\; 110.4. \end{aligned}$$

In contrast,

$$\begin{aligned} E(\tau_2|Z{=}f) \;&=\; \sum_u E^{X=2}(Y|U{=}u)\cdot P(U{=}u\,|Z{=}f)\\ &= 90\cdot\frac{1}{3} + 120\cdot\frac{1}{3} + 126\cdot\frac{1}{3} \;=\; 112. \end{aligned}$$

Hence, $E(\tau_2|X{=}1, Z{=}f) \neq E(\tau_2|Z{=}f)$, and this proves that $\forall x\colon \tau_x \vDash X|Z$ does not hold in this example.

To summarize: Whereas, in this example, $\tau_x \vDash 1_{X=x}|Z$ holds for all values x of X — and with it, unbiasedness of the conditional expectations $E(Y|X,Z)$, $E^{X=x}(Y|Z)$, and the conditional expectation values $E(Y|X{=}x, Z{=}z)$ — the more restrictive condition $\forall x\colon \tau_x \vDash X|Z$ *does not* hold. Hence, $\tau \perp\!\!\!\perp X|Z$ (strong ignorability), which implies $\forall x\colon \tau_x \vDash X|Z$, does not hold as well in this example (see Exercise 7-5).

7.4.2 *Z*-Conditional Independence of τ and *X*

Table 7.5 displays the parameters of a random experiment in which there is Z-conditional independence of τ and X. Again we assume that this random experiment has the same structure as the random experiments treated in section 6.5. That is,

Table 7.5. Z-conditional independence of X and τ

Person u	Sex z_1	College z_2	$P(U=u)$	$P(X=1 \mid U=u)$	$P(X=1 \mid Z=(z_1, z_2))$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$CTE_{U;10}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$
Tom	*m*	*no*	1/6	7/8	6/8	72	83	11	1/22	7/26
Tim	*m*	*no*	1/6	5/8	6/8	72	83	11	3/22	5/26
Joe	*m*	*yes*	1/6	5/8	5/8	95	100	5	3/22	5/26
Jim	*m*	*yes*	1/6	5/8	5/8	100	105	5	3/22	5/26
Ann	*f*	*yes*	1/6	2/8	2/8	106	114	8	6/22	2/26
Sue	*f*	*yes*	1/6	2/8	2/8	116	130	14	6/22	2/26

	$x = 0$	$x = 1$	
$E(\tau_x)$:	93.5	102.5	$ATE_{10} = 9$
$E(Y \mid X=x)$:	100.227	96.5	$PFE_{10} = -3.727$
$E\big(\tau_x \,\big|\, Z=(m, no)\big)$:	72	83	$CTE_{Z;10}(m, no) = 11$
$E\big(Y \big| X=x, Z=(m, no)\big)$:	72	83	$PFE_{Z;10}(m, no) = 11$
$E\big(\tau_x \,\big|\, Z=(m, yes)\big)$:	97.5	102.5	$CTE_{Z;10}(m, yes) = 5$
$E\big(Y \big| X=x, Z=(m, yes)\big)$:	97.5	102.5	$PFE_{Z;10}(m, yes) = 5$
$E\big(\tau_x \,\big|\, Z=(f, yes)\big)$:	111	122	$CTE_{Z;10}(f, yes) = 11$
$E\big(Y \big| X=x, Z=(f, yes)\big)$:	111	122	$PFE_{Z;10}(f, yes) = 11$

Note: In this table $Z = (Z_1, Z_2)$ is a two-dimensional random variable.

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big),$$

as specified in section 6.5, is the regular probabilistic causality setup. This also includes the set $\Omega_2 = \Omega_X = \{control, treatment\}$, represented by the values 0 and 1 of the treatment variable X. And again, U is a global potential confounder of X.

The upper left part of Table 7.5 displays the true outcomes under treatment and under control as well as the probabilities for each person to be assigned to treatment condition 1. In the random experiment presented in this table, the true treatment probabilities and true outcomes are such that Z-conditional independence of τ and X (i. e., Z-conditional strong ignorability) holds.

We check if $\tau \perp\!\!\!\perp X \mid Z$ actually holds via checking

$$P(X=1 \mid Z, \tau) \underset{P}{=} P(X=1 \mid Z) \tag{7.72}$$

[see Prop. (7.44)]. According to RS-Theorem 6.6, Equation (7.72) is equivalent to

$$P(X=0 \mid Z, \tau) \underset{P}{=} P(X=0 \mid Z) \tag{7.73}$$

because X is binary. According to the same theorem, this equation is also equivalent to $\tau \perp\!\!\!\perp 1_{X=1} | Z$ and because, in this example, $1_{X=1} = X$, it is equivalent to $\tau \perp\!\!\!\perp X | Z$ as well. Furthermore, in this example, $P(Z{=}z, \tau{=}t) > 0$ for all pairs $(z, t) \in Z(\Omega) \times \tau(\Omega)$. Therefore, Equation (7.72) is also equivalent to

$$P(X{=}1 | Z{=}z, \tau{=}t) = P(X{=}1 | Z{=}z), \quad \forall (z, t) \in Z(\Omega) \times \tau(\Omega) \tag{7.74}$$

[see RS-Th. 4.42 and RS-Eq. (3.26)]. Note that, in this example, $Z = (Z_1, Z_2)$ and $\tau = (\tau_0, \tau_1)$ are two-dimensional random variables.

In the sequel, we use

$$P(X{=}x \,|\, V{=}v) = \sum_u P(X{=}x \,|\, V{=}v, U{=}u) \cdot P(U{=}u \,|\, V{=}v), \tag{7.75}$$

which is always true if $P(V{=}v, U{=}u) > 0$ for all values of U [see RS-Box 3.2 (ii) for $Y = 1_{X=x}$]. Using Equation (7.75) for the parameters displayed in Table 7.5 with $Z = (Z_1, Z_2)$ taking the role of V and considering $x{=}1$ and $z{=}(m, no)$ yields

$$\begin{aligned} &P\big(X{=}1 \big| Z{=}(m, no)\big) \\ &= \sum_u P\big(X{=}1 \big| Z{=}(m, no), U{=}u\big) \cdot P\big(U{=}u \big| Z{=}(m, no)\big) \\ &= 7/8 \cdot 1/2 + 5/8 \cdot 1/2 = 6/8. \end{aligned}$$

In this case, we only have to sum over the first two persons displayed in Table 7.5 because, for the other four persons, the probabilities $P\big(U{=}u \big| Z{=}(m, no)\big)$ are zero. Applying Equation (7.75) to $V = (Z, \tau)$ and the combination of values $z = (m, no)$ and $t = (72, 83)$ yields exactly the same probability

$$\begin{aligned} &P\big(X{=}1 \big| Z{=}(m, no), \tau{=}(72, 83)\big) \\ &= \sum_u P\big(X{=}1 \big| Z{=}(m, no), \tau{=}(72, 83), U{=}u\big) \cdot P\big(U{=}u \big| Z{=}(m, no), \tau{=}(72, 83)\big) \\ &= 7/8 \cdot 1/2 + 5/8 \cdot 1/2 = 6/8 \end{aligned}$$

(see the first two rows of Table 7.5). Hence, we have shown

$$P\big(X{=}1 \big| Z{=}(m, no)\big) = P\big(X{=}1 \big| Z{=}(m, no), \tau{=}(72, 83)\big) = 6/8.$$

Again using Equation (7.75) with Z taking the role of V and considering the case $x{=}1$ and $z{=}(m, yes)$ yields

$$\begin{aligned} &P\big(X{=}1 \big| Z{=}(m, yes)\big) \\ &= \sum_u P\big(X{=}1 \big| Z{=}(m, yes), U{=}u\big) \cdot P\big(U{=}u \big| Z{=}(m, yes)\big) \\ &= 5/8 \cdot 1/2 + 5/8 \cdot 1/2 = 5/8. \end{aligned}$$

In this case, we only have to sum over persons three and four displayed in Table 7.5; for the other four persons, the conditional probabilities $P\big(U{=}u \big| Z{=}(m, yes)\big)$ are zero.

Applying Equation (7.75) to $V = (Z, \tau)$ and the combination of values $Z = (m, yes)$ and $\tau = (95, 100)$ yields exactly the same conditional probability

$$
\begin{aligned}
&P\big(X{=}1\,\big|\,Z{=}(m, yes), \tau{=}(95,100)\big)\\
&= \sum_u P\big(X{=}1\,\big|\,Z{=}(m, yes), \tau{=}(95,100), U{=}u\big) \cdot P\big(U{=}u\,\big|\,Z{=}(m, yes), \tau{=}(95,100)\big)\\
&= 5/8 \cdot 1 \;=\; 5/8
\end{aligned}
$$

(see the third row of Table 7.5). The same result is obtained if we apply Equation (7.75) to $V = (Z, \tau)$ and the combination of values $z = (m, yes)$ and $t = (100, 105)$ (see the fourth row of Table 7.5). Hence we have shown

$$
\begin{aligned}
P\big(X{=}1\big|Z{=}(m, yes)\big) &= P\big(X{=}1\big|Z{=}(m, yes), \tau{=}(95,100)\big)\\
&= P\big(X{=}1\big|Z{=}(m, yes), \tau{=}(100,105)\big) \;=\; 5/8.
\end{aligned}
$$

Finally, the analog procedure yields

$$
\begin{aligned}
P\big(X{=}1\big|Z{=}(f, yes)\big) &= P\big(X{=}1\big|Z{=}(f, yes), \tau{=}(106,114)\big)\\
&= P\big(X{=}1\big|Z{=}(f, yes), \tau{=}(116,130)\big) \;=\; 2/8.
\end{aligned}
$$

This proves that Proposition (7.74), and with it, Equation (7.72) — which is equivalent to $\tau \perp\!\!\!\perp X \,|\, Z$ — holds in this example.

7.5 Summary and Conclusions

In this chapter, we treated some causality conditions, all of which involve the true outcome variables τ_x. The simple (i. e., the unconditional) ones are listed in Box 7.1, the conditional ones in Box 7.2. The implication relations among the simple causality conditions are listed in Table 7.1, whereas the implications among the conditional ones are found in Table 7.2. According to the last row of Table 7.2, Z-conditional independence of τ and X, the translation of Rosenbaum and Rubin's Z-conditional strong ignorability into true outcome theory, is the strongest, that is, the most restrictive condition among those causality conditions in which we condition on a covariate Z of X; it implies all other causality conditions listed in that table. The same applies to independence of τ and X, which implies all other causality conditions listed in Table 7.1.

Note that there are no implications between the simple causality conditions summarized in Table 7.1 and the conditional ones listed in Table 7.2. Even the strongest (most restrictive) condition $\tau \perp\!\!\!\perp X$ does not imply any of the causality conditions listed in Table 7.2 or vice versa. This implies, for example, that $\forall x$: $\tau_x \vDash 1_{X=x}$, which is equivalent to unbiasedness of $E(Y|X)$, does not imply $\forall x$: $\tau_x \vDash 1_{X=x} | Z$, which is equivalent to unbiasedness of $E(Y|X,Z)$ (see also the example described in section 6.6). Furthermore, $\tau \perp\!\!\!\perp X|Z$ does not imply $\tau \perp\!\!\!\perp X$ or any of the simple causality conditions listed in Box 7.1.

Limitations

The causality conditions treated in the present chapter have three important limitations: They are not generalizable, they can only indirectly be created via experimental design techniques, and they are not testable in empirical applications.

The term *not generalizable* refers to the fact mentioned above that, for example, $\tau \perp\!\!\!\perp X$ does not imply $\tau \perp\!\!\!\perp X|Z$, even not if Z is a covariate of X. This has serious disadvantages for the interpretation of the conditional expectation $E(Y|X,Z)$. Although $\tau \perp\!\!\!\perp X$ implies that $E(Y|X)$ is unbiased, we cannot conclude that $E(Y|X,Z)$ is unbiased as well. Suppose that the putative cause variable X is binary and Z represents the covariate *sex* with values m and f. Then, even if we know the causal total effects of treatment condition x compared to x' on the outcome Y, then we do not know anything about the conditional causal total effects for males and for females. This also means that the difference between the conditional expectation values $E(Y|X{=}1,Z{=}m)$ and $E(Y|X{=}0,Z{=}m)$ may not have a causal meaning, and the same applies to the corresponding difference for females.

The second limitation is that these conditions can only *indirectly* be created via experimental design techniques such as randomized assignment or conditionally randomized assignment of the observational unit (person) to a treatment condition. In chapters 8 to 10 we will treat other causality conditions that imply all conditions treated in the present chapter and that are created via randomized assignment of the observational unit to a treatment condition. Randomized assignment of a unit to a treatment is a way to create independence of the global potential confounder D_X and X, which implies $\tau \perp\!\!\!\perp X$ and $\tau \perp\!\!\!\perp X|Z$, and with them all causality conditions treated in the present chapter, including unbiasedness of $E(Y|X)$ and $E(Y|X,Z)$, provided that Z is a covariate of X (see Rem. 4.16). In this case, $E(Y|X{=}1) - E(Y|X{=}0)$ and $E(Y|X{=}1,Z{=}z) - E(Y|X{=}0,Z{=}z)$ are unbiased and have a causal interpretation if $P(X{=}x,Z{=}z) > 0$ for the values (x,z) of X and Z. In contrast, Z-conditionally randomized assignment of a unit to a treatment is a way to create Z-conditional independence of D_X and X, which implies $\tau \perp\!\!\!\perp X|Z$, and with it unbiasedness of $E(Y|X,Z)$, which is sufficient for computing conditional and average causal total effects (for more details see Box 6.3).

The third disadvantage mentioned above is *testability*. Just like unbiasedness, all causality conditions treated in this chapter, including Z-conditional independence of X and τ (i. e., strong ignorability), cannot be tested empirically without unrealistic additional assumptions because, unlike in the numerical examples treated in this chapter, in empirical applications, the values of the true outcome variables τ_x are unknown and cannot be estimated for more than one single value x of X (see the fundamental problem of causal inference described in the Preface and in Rem. 5.13). Therefore, in contrast to the causality conditions treated in the chapters to come, the causality conditions treated in the present chapter cannot be used for a selection of a covariate Z. Nevertheless, the Rosenbaum-Rubin conditions are of theoretical interest because they are implied by other causality conditions that *are* empirically testable (see chs. 8 to 10).

7.6 Proofs

Proof of Theorem 7.7

The assumptions include that τ_x is P-unique. Under this assumption,

$$
\begin{aligned}
\tau_x \perp\!\!\!\perp 1_{X=x} \quad &\Rightarrow \quad \tau_x \vDash 1_{X=x} && \text{[RS-Th. 4.40]}\\
&\Leftrightarrow \quad E(\tau_x|\, 1_{X=x}) \underset{P}{=} E(\tau_x) && \text{[RS-Def. 4.36]}\\
&\Leftrightarrow \quad E(Y|X{=}x) \vdash D_X. && \text{[Th. 6.9 (i)]}
\end{aligned}
$$

Proof of Theorem 7.8

Proposition (7.17). This proposition immediately follows from the fact that $\sigma(1_{X=x}) \subset \sigma(X)$ and RS-Box 2.1 (iv).

Propositions (7.18) to (7.20). With 7.1 (d) we assume that τ_x is P-unique. Under this assumption,

$$
\begin{aligned}
\tau_x \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau_x \vDash X && \text{[RS-Th. 4.40]}\\
&\Rightarrow \quad \tau_x \vDash 1_{X=x} && [\sigma(1_{X=x}) \subset \sigma(X),\ \text{RS-(4.45)}]\\
&\Leftrightarrow \quad E(Y \mid X{=}x) \vdash D_X . && [(7.16)]
\end{aligned}
$$

Proof of Theorem 7.17

Proposition (7.46). According to Assumption 7.1 (d), the true outcome variable τ_x is P-unique. Hence,

$$
\begin{aligned}
& \tau_x \vDash X \mid Z \\
\Leftrightarrow\quad & E(\tau_x \mid X, Z) \underset{P}{=} E(\tau_x \mid Z) && [(7.38)]\\
\Rightarrow\quad & E\big(E(\tau_x \mid X, Z) \,\big|\, 1_{X=x}, Z\big) \underset{P}{=} E\big(E(\tau_x \mid Z) \,\big|\, 1_{X=x}, Z\big) && \text{[RS-Box 4.1 (xiv)]}\\
\Leftrightarrow\quad & E(\tau_x \mid 1_{X=x}, Z) \underset{P}{=} E\big(E(\tau_x \mid Z) \,\big|\, 1_{X=x}, Z\big) \\
& && [\sigma(1_{X=x}, Z) \subset \sigma(X, Z),\ \text{RS-Box 4.1 (xiii)}]\\
\Leftrightarrow\quad & E(\tau_x \mid 1_{X=x}, Z) \underset{P}{=} E(\tau_x \mid Z) && [\sigma\big(E(\tau_x \mid Z)\big) \subset \sigma(1_{X=x}, Z),\ \text{RS-Box 4.1 (xi)}]\\
\Leftrightarrow\quad & \tau_x \vDash 1_{X=x} \mid Z . && [(7.37)]
\end{aligned}
$$

Proposition (7.47).

$$
\begin{aligned}
\tau_x \vDash 1_{X=x} \mid Z \quad &\Leftrightarrow \quad E(\tau_x \mid 1_{X=x}, Z) \underset{P}{=} E(\tau_x \mid Z) && [(7.37)]\\
&\Leftrightarrow \quad E^{X=x}(Y \mid Z) \vdash D_X . && \text{[Th. 6.20 (i)]}
\end{aligned}
$$

Proof of Corollary 7.18

Proposition (7.48).

$$
\forall\, x \in X(\Omega)\colon \tau_x \vDash X \mid Z \quad \Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} \mid Z . \qquad [(7.46)]
$$

Proposition (7.49).

$$
\begin{aligned}
\forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} \mid Z \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y \mid Z) \vdash D_X && [(7.47)]\\
&\Leftrightarrow \quad E(Y \mid X, Z) \vdash D_X . && \text{[Def. 6.18 (i)]}
\end{aligned}
$$

Proof of Theorem 7.19

Proposition (7.50). We assume that τ_x is P-unique. Hence,

$$\tau_x \perp\!\!\!\perp 1_{X=x}|Z \quad \Rightarrow \quad \tau_x \vDash 1_{X=x}|Z. \qquad \text{[RS-Box 4.1 (xiv), RS-Th. 6.8]}$$

Proposition (7.51). We assume that Z is a covariate of X. Hence,

$$\tau_x \vDash 1_{X=x}|Z \quad \Leftrightarrow \quad E^{X=x}(Y|Z) \vdash D_X. \qquad [(7.47)]$$

Proof of Theorem 7.20

Proposition (7.52). We assume that τ_x is P-unique. Hence,

$$\tau_x \perp\!\!\!\perp X|Z \quad \Rightarrow \quad \tau_x \vDash X|Z. \qquad \text{[RS-Th. 6.8]}$$

Propositions (7.53) and (7.54).

$$\begin{aligned} \tau_x \perp\!\!\!\perp X|Z \quad &\Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x}|Z && [\sigma(1_{X=x}) \subset \sigma(X), \text{ RS-Box 6.1 (vi)}] \\ &\Rightarrow \quad \tau_x \vDash 1_{X=x}|Z. && [\tau_x \text{ is } P\text{-unique, RS-Th. 6.8}] \end{aligned}$$

Proof of Theorem 7.21

Propositions (7.55) to (7.58). Under Assumptions 7.1 (a) to (c), (e), and (g),

$$\begin{aligned} & \tau \perp\!\!\!\perp X|Z \\ \Leftrightarrow \quad & \forall\, x \in X(\Omega):\ \tau \perp\!\!\!\perp 1_{X=x}|Z && \text{[RS-(6.8)]} \\ \Rightarrow \quad & \forall\, x, x' \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x'}|Z && [\sigma(\tau_x) \subset \sigma(\tau), \text{ RS-Box 6.1 (vi)}] \\ \Leftrightarrow \quad & \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp X|Z && \text{[RS-(6.8)]} \\ \Leftrightarrow \quad & \forall\, x, x' \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x'}|Z && \text{[RS-(6.8)]} \\ \Rightarrow \quad & \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x}|Z. && [x \in X(\Omega)] \end{aligned}$$

Propositions (7.59) and (7.60). Under the Assumptions 7.1 (a) to (g),

$$\begin{aligned} & \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp X|Z \\ \Rightarrow \quad & \forall\, x \in X(\Omega):\ \tau_x \vDash X|Z && [\tau_x \text{ is } P\text{-unique, RS-Th. 6.8}] \\ \Rightarrow \quad & \forall\, x \in X(\Omega):\ \tau_x \vDash 1_{X=x}|Z. && [\sigma(1_{X=x}) \subset \sigma(X), \text{ RS-(4.52)}] \end{aligned}$$

Proposition (7.61). Under the Assumptions 7.1 (a) to (g),

$$\begin{aligned} & \forall\, x \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x}|Z \\ \Rightarrow \quad & \forall\, x \in X(\Omega):\ \tau_x \vDash 1_{X=x}|Z. && [\tau_x \text{ is } P\text{-unique, RS-Th. 6.8}] \end{aligned}$$

Propositions (7.62) and (7.63). If we additionally assume that Z is a covariate of X, then

$$
\begin{aligned}
& \forall\, x \in X(\Omega):\ \tau_x \vDash 1_{X=x} \,|\, Z \\
\Leftrightarrow\quad & \forall\, x \in X(\Omega):\ E^{X=x}(Y|Z) \vdash D_X && \text{[Th. 6.20 (i), (7.38)]} \\
\Leftrightarrow\quad & E(Y|X,Z) \vdash D_X. && \text{[Def. 6.18 (i)]}
\end{aligned}
$$

7.7 Exercises

▷ **Exercise 7-1** Prove Propositions (7.5) and (7.6).

▷ **Exercise 7-2** Prove Propositions (7.7) to (7.9).

▷ **Exercise 7-3** Prove that Proposition (7.14) implies (7.13).

▷ **Exercise 7-4** Check the implications listed in Table 7.2 and find their proofs in this chapter.

▷ **Exercise 7-5** Change a single number in Table 7.3 so that in this modified example $\tau \perp\!\!\!\perp X \,|\, (Z{=}f)$ holds. Use the Causal Effects Xplorer to check this condition.

▷ **Exercise 7-6** Show that $P(X{=}x \,|\, D_X) \underset{P}{>} 0$ implies $P(X{=}x \,|\, Z) \underset{P}{>} 0$, if Z is a covariate of X.

Solutions

▷ **Solution 7-1** If the Assumptions 7.1 (a) to (d) hold, then

$$
\begin{aligned}
\tau_x \vDash X \quad & \Leftrightarrow\quad E(\tau_x|X) \underset{P}{=} E(\tau_x) && \text{[(7.2)]} \\
& \Leftrightarrow\quad \forall\, x' \in X(\Omega):\ E(\tau_x|\, 1_{X=x'}) \underset{P}{=} E(\tau_x) && \text{[RS-(4.40)]} \\
& \Leftrightarrow\quad \forall\, x' \in X(\Omega):\ \tau_x \vDash 1_{X=x'} && \text{[RS-Def. 4.36]} \\
& \Rightarrow\quad \tau_x \vDash 1_{X=x} && [x \in X(\Omega)] \\
& \Leftrightarrow\quad E(Y|X{=}x) \vdash D_X. && \text{[(6.5), Th. 6.9 (i)]}
\end{aligned}
$$

▷ **Solution 7-2** Under the Assumptions 7.1 (a) to (f), Propositions (7.7) to (7.8) immediately follow from Propositions (7.5) to (7.6), and Proposition (7.9) follows from Definition 6.3 (ii).

▷ **Solution 7-3**

$$
\begin{aligned}
\tau \perp\!\!\!\perp X \quad & \Leftrightarrow\quad \forall\, x \in X(\Omega):\ P(X{=}x \,|\, \tau) \underset{P}{=} P(X{=}x) && \text{[(7.14)]} \\
& \Leftrightarrow\quad \forall\, x' \in X(\Omega):\ E(1_{X=x'} \,|\, \tau) \underset{P}{=} E(1_{X=x'}) && \text{[RS-(4.10), RS-(3.9)]} \\
& \Rightarrow\quad \forall\, x, x' \in X(\Omega):\ E\big(E(1_{X=x'} \,|\, \tau) \,\big|\, \tau_x\big) \underset{P}{=} E\big(E(1_{X=x'}) \,\big|\, \tau_x\big) && \text{[RS-Box 4.1 (xiv)]} \\
& \Leftrightarrow\quad \forall\, x, x' \in X(\Omega):\ E(1_{X=x'} \,|\, \tau_x) \underset{P}{=} E(1_{X=x'}) && \text{[RS-Box 4.1 (xiii), (i)]} \\
& \Leftrightarrow\quad \forall\, x, x' \in X(\Omega):\ P(X{=}x' \,|\, \tau_x) \underset{P}{=} P(X{=}x') && \text{[RS-(4.10), RS-(3.9)]} \\
& \Leftrightarrow\quad \forall x:\ \tau_x \perp\!\!\!\perp X. && \text{[(7.13)]}
\end{aligned}
$$

▷ **Solution 7-4** The propositions of Table 7.2 are considered row wise.

Row 1

(1) Under the Assumptions 7.1 (a) to (d) and (g): $\tau_x \vDash X \mid Z \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This is Proposition (7.46).

Row 2

(2) Under the Assumptions 7.1 (a) to (d) and (g): $\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This is Proposition (7.50).

Row 3

(3) Under the Assumptions 7.1 (a) to (d) and (g): $\tau_x \perp\!\!\!\perp X \mid Z \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This is Proposition (7.54).

(4) Under the Assumptions 7.1 (a) to (d) and (g): $\tau_x \perp\!\!\!\perp X \mid Z \;\Rightarrow\; \tau_x \vDash X \mid Z$.
This is Proposition (7.52).

(5) Under the Assumptions 7.1 (a) to (c) and (g): $\tau_x \perp\!\!\!\perp X \mid Z \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$.
This is Proposition (7.53).

Row 4

(6) Under the Assumptions 7.1 (a) to (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} \mid Z\big) \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This proposition immediately follows from $x \in X(\Omega)$.

Row 5

(7) Under the Assumptions 7.1 (a) to (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \vDash X \mid Z\big) \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This immediately follows from Proposition (7.60).

(8) Under the Assumptions 7.1 (a) to (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \vDash X \mid Z\big) \;\Rightarrow\; \tau_x \vDash X \mid Z$.
This proposition immediately follows from $x \in X(\Omega)$.

Row 6

(9) Under the Assumptions 7.1 (a) to (e) and (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z\big) \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This immediately follows from Proposition (7.61).

(10) Under the Assumptions 7.1 (a) to (c), (e) and (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z\big) \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$.
This proposition immediately follows from $x \in X(\Omega)$.

(11) Under the Assumptions 7.1 (a) to (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z\big) \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} \mid Z\big)$.
This follows from Proposition (7.54).

Row 7

(12) Under the Assumptions 7.1 (a) to (e) and (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \mid Z\big) \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This follows from Propositions (7.59) and (7.60).

(13) Under the Assumptions 7.1 (a) to (e) and (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \mid Z\big) \;\Rightarrow\; \tau_x \vDash X \mid Z$.
This immediately follows from Proposition (7.59).

(14) Under the Assumptions 7.1 (a) to (c), (e) and (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \mid Z\big) \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$.
This follows from Proposition (7.53).

(15) Under the Assumptions 7.1 (a) to (c), (e) and (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \mid Z\big) \;\Rightarrow\; \tau_x \perp\!\!\!\perp X \mid Z$.
This proposition immediately follows from $x \in X(\Omega)$.

(16) Under the Assumptions 7.1 (a) to (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \mid Z\big) \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} \mid Z\big)$.
This follows from Propositions (7.53) and (7.54).

(17) Under the Assumptions 7.1 (a) to (g): $\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \mid Z\big) \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \vDash X \mid Z\big)$.
This follows from Proposition (7.52).

(18) Under the Assumptions 7.1 (a) to (c), (e) and (g):
$\big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \mid Z\big) \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z\big)$. This follows from Proposition (7.53).

Row 8

(19) Under the Assumptions 7.1 (a) to (e) and (g): $\tau \perp\!\!\!\perp X \mid Z \;\Rightarrow\; \tau_x \vDash 1_{X=x} \mid Z$.
This follows from Propositions (7.55) to (7.60).

(20) Under the Assumptions 7.1 (a) to (e) and (g): $\tau \perp\!\!\!\perp X \mid Z \;\Rightarrow\; \tau_x \vDash X \mid Z$.
This immediately follows from Propositions (7.55) to (7.59).

(21) Under the Assumptions 7.1 (a) to (c), (e) and (g): $\tau \perp\!\!\!\perp X|Z \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x}|Z$.
This follows from Propositions (7.55) and (7.58).

(22) Under the Assumptions 7.1 (a) to (c), (e) and (g): $\tau \perp\!\!\!\perp X|Z \;\Rightarrow\; \tau_x \perp\!\!\!\perp X|Z$.
This follows from Propositions (7.55) to (7.57).

(23) Under the Assumptions 7.1 (a) to (g): $\tau \perp\!\!\!\perp X|Z \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x}|\,Z\big)$.
This follows from Propositions (7.55) to (7.60).

(24) Under the Assumptions 7.1 (a) to (g): $\tau \perp\!\!\!\perp X|Z \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \vDash X|\,Z\big)$.
This immediately follows from Propositions (7.55) to (7.59).

(25) Under the Assumptions 7.1 (a) to (c), (e) and (g): $\tau \perp\!\!\!\perp X|Z \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp 1_{X=x}|Z\big)$.
This follows from Propositions (7.55) to (7.58).

(26) Under the Assumptions 7.1 (a) to (c), (e) and (g): $\tau \perp\!\!\!\perp X|Z \;\Rightarrow\; \big(\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X|Z\big)$.
This follows from Propositions (7.55) to (7.57).

▷ **Solution 7-5** Change $E^{X=2}(Y\,|\,U=Eva)$ from 126 to 150.

▷ **Solution 7-6** Remember, if Z is a covariate of Z, then $\sigma(Z) \subset \sigma(D_X)$ (see Rem. 4.16). Hence,

$$
\begin{aligned}
P(X=x\,|\,D_X) \underset{P}{>} 0 \quad &\Rightarrow \quad E\big(P(X=x\,|\,D_X)\,\big|\,Z\big) \underset{P}{>} 0 && \text{[RS-(4.13)]}\\
&\Leftrightarrow \quad E\big(E(1_{X=x}|D_X)\,\big|\,Z\big) \underset{P}{>} 0 && \text{[RS-(4.10)]}\\
&\Leftrightarrow \quad E(1_{X=x}|Z) \underset{P}{>} 0 && \text{[SN-(2.40), } \sigma(Z) \subset \sigma(D_X)\text{, RS-Box 4.1 (xiii)]}\\
&\Leftrightarrow \quad P(X=x\,|\,Z) \underset{P}{>} 0. && \text{[RS-(4.10)]}
\end{aligned}
$$

Chapter 8
Fisher Conditions

In chapter 6, we treated *unbiasedness* of the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$, and $E(Y|X,Z)$ and unbiasedness of the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x, Z{=}z)$. Unbiasedness of these terms is a first kind of causality conditions justifying causal interpretations of the dependencies described by conditional expectations and conditional expectation values. For example, if $E(Y|X{=}x)$ and $E(Y|X{=}x')$ are unbiased, then the prima facie effect $PFE_{xx'} = E(Y|X{=}x) - E(Y|X{=}x')$ is unbiased as well and it is identical to the *causal average (total) effect* $ATE_{xx'}$. Therefore, under unbiasedness, an estimate of $PFE_{xx'}$ is also an estimate of the $ATE_{xx'}$. Similarly, if $E(Y|X{=}x, Z{=}z)$ and $E(Y|X{=}x', Z{=}z)$ are unbiased, then the $(Z{=}z)$-conditional prima facie effect $PFE_{Z;xx'}(z)$ $= E(Y|X{=}x, Z{=}z) - E(Y|X{=}x', Z{=}z)$ is unbiased and $PFE_{Z;xx'}(z)$ is identical to the *causal* $(Z{=}z)$*-conditional (total) effect* $CTE_{Z;xx'}(z)$.

In chapter 7, we treated some other causality conditions, the Rosenbaum-Rubin conditions, which involve the true outcome variables τ_x and imply unbiasedness of the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$, $E(Y|X,Z)$, and their values. However, these conditions as well as unbiasedness itself cannot be tested empirically. Therefore we say that *they cannot be falsified*. Consequently, none of these conditions can be used for selecting the random variable Z with respect to which conditional independence of the multivariate true outcome variable τ and X (i. e., $\tau \perp\!\!\!\perp X|Z$) or conditional mean-independence of the true outcome variables τ_x and X (i. e., $\tau_x \vDash X|Z$) hold. Furthermore, unbiasedness can be *accidental* in the sense that it may hold for $E(Y|X)$ but not for $E(Y|X,Z)$, where Z is a covariate of X. This also applies, for example, to the strong ignorability condition $\tau \perp\!\!\!\perp X|Z$. If it holds for a specified covariate Z, then this does not imply that $\tau \perp\!\!\!\perp X|(Z,W)$ holds as well, even if Z and W are covariates of X. This is one of the reasons why, in the present and some of the next chapters, we study other causality conditions that are less volatile (i. e., that are generalizable) and falsifiable, therefore lending themselves for covariate selection in quasi-experiments and observational studies.

The conditions introduced in this chapter will be referred to as the *Fisher conditions* (for total effects). This name is chosen to recognize the contributions of Sir R. A. Fisher to understanding the relevance of the experimental design technique of *randomization* for causal inference (see, e. g., Fisher, 1925/1946). The Fisher conditions can be created by randomly assigning (e. g. by a coin flip) the observational unit to one of several treatment conditions represented by the values x of the putative cause variable X. In contrast to the causality conditions treated in the previous chapters, the Fisher conditions are also *falsifiable*. Hence, they can be tested in samples and can be used for covariate selection. Furthermore, they are not accidental in the sense described above, that is, they are *generalizable*.

Finally, the Fisher conditions may also hold if the putative cause variable X is a continuous random variable. Remember, all causality conditions described in chapters 6 and 7 are defined only if the values x of X have a positive probability $P(X{=}x)$. For example, if X is normally distributed, then $P(X{=}x) = 0$ for all values x of X, and this also applies if X has any other continuous distribution. Even in the definitions of causal effects treated in chapter 5 we presumed that X is discrete and has at least two values x and x' having a positive probability. Hence, although the Fisher conditions have many useful implications on unbiasedness and true outcome variables, they have far reaching consequences beyond true outcome theory.

Requirements

Reading this chapter we assume that the reader is familiar with the concepts treated in all chapters of Steyer (2024). Again, chapters 4 to 6 of that book are now crucial. They deal with the concepts of a conditional expectation, a conditional expectation with respect to a conditional probability measure, and conditional independence. Furthermore, we assume familiarity with chapters 4 to 7 of the present book.

In this chapter, we will often refer to the following assumptions and notation.

Notation and Assumptions 8.1

(a) *Let* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *be a regular probabilistic causality setup, let* D_X *denote a global potential confounder of* X*, and* $\mathscr{W}_X$ *the set of all potential confounders of* X*, that is, the set of all random variables on* $(\Omega, \mathscr{A}, P)$ *satisfying* $\sigma(W) \subset \sigma(D_X)$.

(b) *Let* $(\Omega'_X, \mathscr{A}'_X)$ *denote the value space of* X*, let* $x \in \Omega'_X$*, let* $\{x\} \in \mathscr{A}'_X$*, and let* $1_{X=x}$ *denote the indicator of the event* $\{X{=}x\} = \{\omega \in \Omega\colon X(\omega) = x\}$.

(c) *Let* Y *be real-valued with positive variance, assume* $0 < P(X{=}x) < 1$*, define the probability measure* $P^{X=x}\colon \mathscr{A} \to [0,1]$ *by* $P^{X=x}(A) = P(A \,|\, X{=}x)$*, for all* $A \in \mathscr{A}$*, and let* $\tau_x = E^{X=x}(Y \,|\, D_X)$ *denote a true outcome variable of* Y *given* x.

(d) *Let* Z *be a random variable on* $(\Omega, \mathscr{A}, P)$ *and let* $(\Omega'_Z, \mathscr{A}'_Z)$ *denote its value space.*

(e) *Let* $z \in \Omega'_Z$ *be a value of* Z*, let* $\{z\} \in \mathscr{A}'_Z$*, assume* $P(Z{=}z) > 0$*, and define the probability measure* $P^{Z=z}\colon \mathscr{A} \to [0,1]$ *by* $P^{Z=z}(A) = P(A \,|\, Z{=}z)$*, for all* $A \in \mathscr{A}$.

(f) *Let* $x' \in X(\Omega) = \{0, 1, \ldots, J\}$ *and* $\{x'\} \in \mathscr{A}'_X$*, let* $1_{X=x'}$ *denote the indicator of the event* $\{X{=}x'\} = \{\omega \in \Omega\colon X(\omega) = x'\}$*, and assume* $0 < P(X{=}x') < 1$*. Furthermore, let* $\tau_{x'} = E^{X=x'}(Y \,|\, D_X)$ *denote a true outcome variable of* Y *given* x'.

(g) *For all* $x \in X(\Omega) = \{0, 1, \ldots, J\}$*, let* $\{x\} \in \mathscr{A}'_X$ *and assume* $0 < P(X{=}x) < 1$*. Furthermore, let* $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$ *denote the* $(J+1)$*-variate random variable consisting of the true outcome variables* $\tau_0, \tau_1, \ldots, \tau_J$ *of* Y.

(h) *Let* Z *be a covariate of* X*, that is, let* $\sigma(Z) \subset \sigma(D_X)$.

8.1 F-Conditions

In this section, we present the causality conditions summarily referred to as the simple (or unconditional) and the conditional Fisher (F) conditions. In the simple F-conditions, we assume independence of a putative cause variable X and a global potential confounder D_X of X, or independence of D_X and an indicator variable $1_{X=x}$, where x denotes a value of

X. Among other things, these simple F-conditions imply unbiasedness of the conditional expectation $E(Y|X)$ and the conditional expectation value $E(Y|X{=}x)$, respectively. Such a simple F-condition is equivalent to independence of all potential confounders of X on one side and X (or $1_{X=x}$) on the other side. In contrast, the conditional F-conditions postulate Z-conditional independence of D_X and X (or $1_{X=x}$), where Z denotes a covariate of X. Among other things, these conditional F-conditions imply unbiasedness of $E(Y|X,Z)$ and $E^{X=x}(Y|Z)$, respectively. We also explicitly consider F-conditions in which we condition on a single value z of Z.

From a methodological point of view, it should be noted that the simple F-conditions are the mathematical foundation of the randomized experiment. In contrast, the conditional F-conditions are the mathematical foundation of the conditionally randomized experiment. However, the conditional F-conditions can also be used for covariate selection aiming at establishing conditional independence of D_X and X given the (possibly multivariate) covariate Z (for more details, see sect. 8.6).

8.1.1 Simple F-Conditions

In RS-section 2.4 we already introduced the concept of *independence of two random variables with respect to a probability measure P*. This concept is used repeatedly in Box 8.1, in which the definitions of all F-conditions are gathered. Reading the definitions in this box, remember that $\sigma(D_X)$ and $\sigma(X)$ denote the σ-algebras generated by the random variables D_X and X, respectively (see RS-Def. 2.12).

Remark 8.2 [Independence of D_X and $1_{X=x}$] Box 8.1 (i) presents the definition of *independence of a global potential confounder D_X of X and an indicator variable $1_{X=x}$* for the event $\{X{=}x\}$. In this definition, we require the Assumptions 8.1 (a) and (b). The notation is $D_X \perp\!\!\!\perp 1_{X=x}$. According to RS-Corollary 6.18,

$$
\begin{aligned}
D_X \perp\!\!\!\perp 1_{X=x} \quad &\Leftrightarrow \quad P(X{=}x\,|\,D_X) \underset{P}{=} P(X{=}x) && (8.1)\\
&\Leftrightarrow \quad P(1_{X=x}{=}1\,|\,D_X) \underset{P}{=} P(1_{X=x}{=}1) && (8.2)\\
&\Leftrightarrow \quad P(1_{X=x}{=}0\,|\,D_X) \underset{P}{=} P(1_{X=x}{=}0) && (8.3)\\
&\Leftrightarrow \quad P(X{\neq}x\,|\,D_X) \underset{P}{=} P(X{\neq}x). && (8.4)
\end{aligned}
$$

Note that Propositions (8.1) to (8.4) also hold if D_X is neither discrete nor numerical and if $P(X{=}x) = 0$. Hence, the condition $D_X \perp\!\!\!\perp 1_{X=x}$ can be used beyond true outcome theory of causal effects (see ch. 5), and this applies to all other conditions presented in Box 8.1. ◁

Remark 8.3 [Independence of D_X and X] The second definition [see Box 8.1 (ii)] is *independence of the putative cause variable X and a global potential confounder D_X of X*. The notation is $D_X \perp\!\!\!\perp X$. This concept also applies if neither X nor D_X are discrete or numerical. We only require Assumptions 8.1 (a). However, if X is discrete (see RS-Def. 2.62), then

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; P(X{=}x\,|\,D_X) \underset{P}{=} P(X{=}x) && (8.5)\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; D_X \perp\!\!\!\perp 1_{X=x} && (8.6)
\end{aligned}
$$

(see RS-Cor. 6.17). Hence, if X is discrete, then according to Proposition (8.5), independence of D_X and X (with respect to the probability measure P) implies that the D_X-conditional probabilities of the events $\{X{=}x\} = \{\omega \in \Omega\colon X(\omega) = x\}$ do not depend on the global

Box 8.1 F-conditions

Simple F-conditions

$D_X \perp\!\!\!\perp 1_{X=x}$ *Independence of D_X and $1_{X=x}$*. Under Assumptions 8.1 (a) and (b), it is defined by

$$\forall (A, B) \in \sigma(D_X) \times \sigma(1_{X=x}): \; P(A \cap B) = P(A) \cdot P(B). \tag{i}$$

Under Assumptions 8.1 (a) to (c), it implies $E(Y \,|\, X{=}x) \vdash D_X$.

$D_X \perp\!\!\!\perp X$ *Independence of D_X and X*. Under Assumptions 8.1 (a), it is defined by

$$\forall (A, B) \in \sigma(D_X) \times \sigma(X): \; P(A \cap B) = P(A) \cdot P(B). \tag{ii}$$

Under Assumptions 8.1 (a), (c), and (g), it implies $E(Y \,|\, X) \vdash D_X$.

***Z*-conditional F-conditions**

$D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z$ *Z-conditional independence of D_X and $1_{X=x}$*. Under Assumptions 8.1 (a), (b), and (d), it is defined by

$$\forall (A, B) \in \sigma(D_X) \times \sigma(1_{X=x}): P(A \cap B \,|\, Z) \underset{P}{=} P(A \,|\, Z) \cdot P(B \,|\, Z). \tag{iii}$$

Under Assumptions 8.1 (a) to (d) and that Z is a covariate of X it implies $E^{X=x}(Y \,|\, Z) \vdash D_X$.

$D_X \perp\!\!\!\perp X \,|\, Z$ *Z-conditional independence of D_X and X*. Under Assumptions 8.1 (a) and (d) it is defined by

Note:

$$\forall (A, B) \in \sigma(D_X) \times \sigma(X): \; P(A \cap B \,|\, Z) \underset{P}{=} P(A \,|\, Z) \cdot P(B \,|\, Z). \tag{iv}$$

Under Assumptions 8.1 (a), (c), (d), (g), and that Z is a covariate of X it implies $E(Y \,|\, X, Z) \vdash D_X$.

($Z{=}z$)-conditional F-conditions

$D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$ *$(Z{=}z)$-conditional independence of D_X and $1_{X=x}$*. Under Assumptions 8.1 (a), (b), (d), and (e), it is defined by

$$\forall (A, B) \in \sigma(D_X) \times \sigma(1_{X=x}): P^{Z=z}(A \cap B) \underset{P^{Z=z}}{=} P^{Z=z}(A) \cdot P^{Z=z}(B). \tag{v}$$

Under Assumptions 8.1 (a) to (e), and that Z is a covariate of X it implies $E^{Z=z}(Y \,|\, X{=}x) \vdash D_X$.

$D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$ *$(Z{=}z)$-conditional independence of D_X and X*. Under Assumptions 8.1 (a), (d), and (e), it is defined by

$$\forall (A, B) \in \sigma(D_X) \times \sigma(X): P^{Z=z}(A \cap B) \underset{P^{Z=z}}{=} P^{Z=z}(A) \cdot P^{Z=z}(B). \tag{vi}$$

Under Assumptions 8.1 (a) to (e), (g), and that Z is a covariate of X it implies $E^{Z=z}(Y \,|\, X) \vdash D_X$.

The proofs that the six F-conditions imply unbiasedness of the specified conditional expectations are found in the theorems and corollaries of section 8.3.

potential confounder D_X. Furthermore, according to Proposition (8.6), if X is discrete, then $D_X \perp\!\!\!\perp X$ is also equivalent to independence of D_X and all indicator variables $1_{X=x}$, $x \in X(\Omega)$, where denotes $X(\Omega)$ the image of Ω under the map X. $\triangleleft$

Example 8.4 [First Examples of Independence of D_X and X] Examples of independence of a global potential confounder D_X and the putative cause variable X have already been presented in RS-Table 1.2 and in Table 6.3 of the present book. In both examples, the person variable U takes the role of a global potential confounder D_X. $\triangleleft$

Remark 8.5 [Randomized Assignment of a Unit to a Treatment] As already mentioned in the introduction of this chapter, $D_X \perp\!\!\!\perp X$ can be created by randomized assignment (e. g., via a coin flip) of a unit to a treatment because, (a) by definition of randomized assignment, X deterministically depends on the coin flip whose outcome does not depend on D_X, and (b) D_X is prior or simultaneous to the treatment variable X (see Th. 4.31). This means that X cannot cause D_X and there is no common cause of X and Y. Note that 'randomization creates $D_X \perp\!\!\!\perp X$' is a substantive theory. It involves a term that refers to a design technique that can be applied in empirical research. In contrast to 'independence of D_X and X', the term 'randomized assignment' is not a mathematical concept. $\triangleleft$

8.1.2 Z-Conditional F-Conditions

The general concept of conditional independence of two random variables given a random variable has been treated in some detail in RS-chapter 6. For a more detailed presentation of conditional independence of random variables see SN-chapter 16.

Remark 8.6 [Z-Conditional Independence of D_X and $1_{X=x}$] The third causality condition in Box 8.1 (iii), is *Z-conditional independence of D_X and $1_{X=x}$*, denoted $D_X \perp\!\!\!\perp 1_{X=x} \mid Z$. In this definition, in which we require the Assumptions 8.1 (a), (b), and (d), $P(A \mid Z) := E(1_A \mid Z)$ denotes the Z-conditional probability of the event A (see RS-Rem. 4.12). This F-condition is tantamount to assuming that $1_{X=x}$ on one side and all potential confounders on the other side are Z-conditionally independent. According to RS-Theorem 6.6,

$$
\begin{aligned}
D_X \perp\!\!\!\perp 1_{X=x} \mid Z \quad &\Leftrightarrow \quad P(X{=}x \mid D_X, Z) \underset{P}{=} P(X{=}x \mid Z) &(8.7)\\
&\Leftrightarrow \quad P(1_{X=x}{=}1 \mid D_X, Z) \underset{P}{=} P(1_{X=x}{=}1 \mid Z) &(8.8)\\
&\Leftrightarrow \quad P(1_{X=x}{=}0 \mid D_X, Z) \underset{P}{=} P(1_{X=x}{=}0 \mid Z) &(8.9)\\
&\Leftrightarrow \quad P(X{\neq}x \mid D_X, Z) \underset{P}{=} P(X{\neq}x \mid Z). &(8.10)
\end{aligned}
$$

Note that Propositions (8.7) to (8.10) still hold if $P(X{=}x) = 0$. $\triangleleft$

Remark 8.7 [Z-Conditional Independence of D_X and X] The fourth causality condition, defined in Box 8.1 (iv), is *Z-conditional independence of D_X and X*, denoted $D_X \perp\!\!\!\perp X \mid Z$. For this definition, we require the Assumptions 8.1 (a) and (d). This causality condition is tantamount to the assumption that X on one side and all potential confounders of X on the other side are Z-conditionally independent. $\triangleleft$

Remark 8.8 [The Putative Cause Variable Does Not Have to Be Discrete] Just as $D_X \perp\!\!\!\perp X$, the condition $D_X \perp\!\!\!\perp X \mid Z$ is also defined and may hold if X is continuous. In this case, true outcome theory of causal effects is not applicable because there we presume positive probabilities $P(X{=}x)$ for all $x \in X(\Omega)$, which does not hold if X is continuous. Nevertheless, if Z is a covariate of X, then $D_X \perp\!\!\!\perp X | Z$ is still a causality condition, implying that $E(Y|X,Z)$ describes a *causal* Z-conditional dependence of Y on X. However, in this volume we refrain from generalizing the theory to continuous putative cause variables. $\triangleleft$

Remark 8.9 [$D_X \perp\!\!\!\perp X | Z$ if X is Discrete] If X is discrete (see RS-Def. 2.62), then

$$D_X \perp\!\!\!\perp X | Z \quad \Leftrightarrow \quad \forall\, x \in X(\Omega):\ P(X{=}x \,|\, D_X, Z) \underset{P}{=} P(X{=}x \,|\, Z) \tag{8.11}$$

$$\Leftrightarrow \quad \forall\, x \in X(\Omega):\ D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \tag{8.12}$$

(see RS-Th. 6.5). To emphasize, the right-hand sides of these propositions are equivalent to $D_X \perp\!\!\!\perp X | Z$ only if X is discrete. However, in contrast to the definition presented in Box 8.1 (iv), they do not hold anymore, if X is continuous. Also note that Proposition (8.5) is a special case of (8.11) for Z being a constant map, that is, for $\sigma(Z) = \{\Omega, \emptyset\}$. $\triangleleft$

Remark 8.10 [Consequences of Z Being a Covariate of X] If we assume that Z is a covariate of X, then, according to Definition 4.11 (iv) and Remark 4.16, $\sigma(Z) \subset \sigma(D_X)$ holds for the σ-algebras generated by these two random variables. This implies $\sigma(D_X) = \sigma(Z, D_X)$ [see RS-Prop. (2.19)] and

$$P(X{=}x \,|\, D_X, Z) \underset{P}{=} P(X{=}x \,|\, D_X) \tag{8.13}$$

[see RS-Def. 4.4 and RS-Eq. (4.10)]. Hence, if X is discrete and Z is a covariate of X, then Proposition (8.11) can also be written as

$$D_X \perp\!\!\!\perp X | Z \quad \Leftrightarrow \quad \forall\, x \in X(\Omega):\ P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x \,|\, Z). \tag{8.14}$$

Correspondingly, Propositions (8.7) to (8.10) simplify to

$$D_X \perp\!\!\!\perp 1_{X=x} | Z \quad \Leftrightarrow \quad P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x \,|\, Z) \tag{8.15}$$

$$\Leftrightarrow \quad P(1_{X=x}{=}1 \,|\, D_X) \underset{P}{=} P(1_{X=x}{=}1 \,|\, Z) \tag{8.16}$$

$$\Leftrightarrow \quad P(1_{X=x}{=}0 \,|\, D_X) \underset{P}{=} P(1_{X=x}{=}0 \,|\, Z) \tag{8.17}$$

$$\Leftrightarrow \quad P(X{\neq}x \,|\, D_X) \underset{P}{=} P(X{\neq}x \,|\, Z). \tag{8.18}$$

$\triangleleft$

Example 8.11 [A First Example of Z-Conditional Independence] An example of Z-conditional independence of the putative cause variable X and a global potential confounder D_X of X has already been presented in Table 6.4. In this example, U takes the role of a global potential confounder of X. In section 8.5 we will treat several examples in more detail. $\triangleleft$

8.1.3 $(Z{=}z)$-Conditional F-Conditions

Now we turn to conditional independence of a putative cause variable X (or $1_{X=x}$) and a global potential confounder D_X of X *given that Z takes on the value z*. In these definitions, we assume $P(Z{=}z) > 0$ and refer to the probability measure $P^{Z=z}$ defined in Assumptions 8.1 (e).

Remark 8.12 [$(Z{=}z)$-**Conditional Independence of** D_X **And** X] The fifth causality condition [see Box 8.1 (v)] is *$(Z{=}z)$-conditional independence of D_X and $1_{X=x}$*. It is denoted by $D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z)$. The required assumptions are 8.1 (a), (b), (d), and (e). This condition means that $1_{X=x}$ on one side and all potential confounders of X on the other side are $(Z{=}z)$-conditionally independent. The definition shown in Box 8.1 (v) reveals that $(Z{=}z)$-conditional independence of D_X and $1_{X=x}$ is equivalent to *independence of D_X and $1_{X=x}$ with respect to the probability measure $P^{Z=z}$* [see Assumptions 8.1 (e)]. That is,

$$D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z) \quad :\Leftrightarrow \quad D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x}. \tag{8.19}$$

Furthermore, according to RS-Corollary 6.18,

$$\begin{aligned} D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z) \quad &\Leftrightarrow \quad P^{Z=z}(X{=}x | D_X) \underset{P^{Z=z}}{=} P^{Z=z}(X{=}x) && (8.20)\\ &\Leftrightarrow \quad P^{Z=z}(1_{X=x}{=}1 | D_X) \underset{P^{Z=z}}{=} P^{Z=z}(1_{X=x}{=}1) && (8.21)\\ &\Leftrightarrow \quad P^{Z=z}(1_{X=x}{=}0 | D_X) \underset{P^{Z=z}}{=} P^{Z=z}(1_{X=x}{=}0) && (8.22)\\ &\Leftrightarrow \quad P^{Z=z}(X{\neq}x | D_X) \underset{P^{Z=z}}{=} P^{Z=z}(X{\neq}x). && (8.23) \end{aligned}$$

$\lhd$

Remark 8.13 [$(Z{=}z)$-**Conditional Independence of** D_X **and** X] The sixth causality condition, introduced in Box 8.1 (vi), is *$(Z{=}z)$-conditional independence of D_X and X*, denoted $D_X \perp\!\!\!\perp X | (Z{=}z)$. The required assumptions are 8.1 (a), (d), and (e). This condition means that the putative cause variable X on one side and all potential confounders of X on the other side are $(Z{=}z)$-conditionally independent. According to the definition in Box 8.1 (vi), $(Z{=}z)$-conditional independence of D_X and X is equivalent to *independence of D_X and X with respect to the probability measure $P^{Z=z}$*. That is,

$$D_X \perp\!\!\!\perp X | (Z{=}z) \quad :\Leftrightarrow \quad D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} X. \tag{8.24}$$

$\lhd$

Remark 8.14 [$(Z{=}z)$-**Conditional Independence of** D_X **and a Finite** X] Suppose that X is finite or countable. Then, according to Remark 8.3 and Proposition (8.24),

$$\begin{aligned} D_X \perp\!\!\!\perp X | (Z{=}z) \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ P^{Z=z}(X{=}x | D_X) \underset{P^{Z=z}}{=} P^{Z=z}(X{=}x) && (8.25)\\ &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\ D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z). && (8.26) \end{aligned}$$

Hence, if X is finite or countable, then $(Z{=}z)$-conditional independence of D_X and X is equivalent to $(Z{=}z)$-conditional independence of D_X and $1_{X=x}$ for all values x of X. $\lhd$

Remark 8.15 [Z-**Conditional Randomized Assignment**] If Z is prior in $(\mathscr{F}_t)_{t \in T}$ to X [see Def. 3.3 (ii)] and therefore a covariate of X (see Cor. 4.33), then we can conduct experiments in which we create $D_X \perp\!\!\!\perp X | Z$ without $D_X \perp\!\!\!\perp X$. For example, conditional independence of a treatment variable X and a global potential confounder D_X of X given a covariate Z of X may be created via Z-conditional randomization. In this technique of experimental design, assignment to treatment is arranged such that the conditional probabilities $P(X{=}x | Z{=}z)$ may differ for different values z of Z (representing, e. g., different degrees of severity of the disorder before treatment), but $D_X \perp\!\!\!\perp X \,|\, Z$ still holds. The reason is that, for all $x \in X(\Omega)$, this assignment procedure secures that no other potential confounder than Z determines the probabilities of being assigned to treatment x (see Rem. 8.59 or more details). ◁

8.2 Implications Among F-Conditions

In this section we study the implication structure among the F-conditions. A summary of all these implications is presented in Box 8.1. Proofs are found in the theorems treated in the present section and in the solution to Exercise 8-2.

Remark 8.16 [First Consequences of $D_X \perp\!\!\!\perp X$ and $D_X \perp\!\!\!\perp X \,|\, Z$] Under the Assumptions 8.1 (a) and (b),

$$D_X \perp\!\!\!\perp X \quad \Rightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} \tag{8.27}$$

because $\sigma(1_{X=x}) \subset \sigma(X)$ [see RS-Box 2.1 (iv)]. Correspondingly, and for the same reason, under the Assumptions 8.1 (a), (b), and (d),

$$D_X \perp\!\!\!\perp X \,|\, Z \quad \Rightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \tag{8.28}$$

[see RS-Box 6.1 (vi)]. ◁

In the following theorem, we consider an implication that we call *generalizability* of independence of a putative cause variable X and a global potential confounder D_X of X.

Theorem 8.17 [Generalizability of $D_X \perp\!\!\!\perp X$]
Let the Assumptions 8.1 (a) hold.

(i) Then

$$D_X \perp\!\!\!\perp X \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp X \,|\, W. \tag{8.29}$$

(ii) If we additionally assume 8.1 (b), then

$$D_X \perp\!\!\!\perp 1_{X=x} \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp 1_{X=x} | W. \tag{8.30}$$

(Proof p. 256)

Hence, under the Assumptions 8.1 (a) and that W is a potential confounder of X, independence of a putative cause variable X and a global potential confounder D_X of X implies W-conditional independence of X and D_X. This proposition is called *generalizability* of $D_X \perp\!\!\!\perp X$. Its methodological implications are discussed in more detail in Remark 8.63. Note that assuming W to be a potential confounder of X is crucial for this implication. Also note that, under the Assumptions 8.1 (a) and (b), Propositions (8.27) and (8.30) imply

$$D_X \perp\!\!\!\perp X \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp 1_{X=x} | W. \tag{8.31}$$

In the following theorem, we extend Theorem 8.17 considering generalizability of Z-conditional independence of a putative cause variable X and a global potential confounder D_X of X.

Table 8.1. Implications among the F-conditions

	$D_X \perp\!\!\!\perp 1_{X=x} \mid (Z{=}z)$	$D_X \perp\!\!\!\perp X \mid (Z{=}z)$	$D_X \perp\!\!\!\perp 1_{X=x} \mid Z$	$D_X \perp\!\!\!\perp X \mid Z$	$D_X \perp\!\!\!\perp 1_{X=x}$
$D_X \perp\!\!\!\perp X \mid (Z{=}z)$	(a), (b), (d), (e)				
$D_X \perp\!\!\!\perp 1_{X=x} \mid Z$	(a), (b), (d), (e)				
$D_X \perp\!\!\!\perp X \mid Z$	(a), (b), (d), (e)	(a), (d), (e)	(a), (b), (d)		
$D_X \perp\!\!\!\perp 1_{X=x}$	(a), (b), (d), (e), (h)		(a), (b), (h)		
$D_X \perp\!\!\!\perp X$	(a), (b), (d), (e), (h)	(a), (d), (e), (h)	(a), (b), (h)	(a), (h)	(a), (b)

Note: An entry such as (a), (b) means that the condition in the row implies the condition in the column, provided that the Assumptions 8.1 (a) and (b) hold. Trivial equivalences such as $D_X \perp\!\!\!\perp 1_{X=x} \Leftrightarrow D_X \perp\!\!\!\perp 1_{X=x}$ are omitted.

Theorem 8.18 [Generalizability of $D_X \perp\!\!\!\perp X \mid Z$]
Let the Assumptions 8.1 (a) and (d) hold.

(i) Then

$$D_X \perp\!\!\!\perp X \mid Z \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp X \mid (Z, W)\,. \tag{8.32}$$

(ii) If we additionally assume 8.1 (b), then

$$D_X \perp\!\!\!\perp 1_{X=x} \mid Z \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp 1_{X=x} \mid (Z, W)\,. \tag{8.33}$$

(Proof p. 257)

Hence, under the assumptions of Theorem 8.18, Z-conditional independence of a putative cause variable X and a global potential confounder D_X of X implies (Z, W)-conditional independence of X and D_X, provided that W is a potential confounder of X. This is what we call *generalizability* of $D_X \perp\!\!\!\perp X \mid Z$ (see Rem. 8.64) for a discussion of the methodological implications). Again, assuming W to be a potential confounder of X is crucial for this implication. Also note that, under the Assumption 8.1 (a), (b), (d), Propositions (8.28) and (8.33) imply

$$D_X \perp\!\!\!\perp X \mid Z \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp 1_{X=x} \mid (Z, W)\,. \tag{8.34}$$

In the next corollary we consider a consequence of $D_X \perp\!\!\!\perp X \mid Z$ that involves conditioning on an event $\{Z{=}z\} = \{\omega \in \Omega\colon Z(\omega) = z\}$ for which we assume $P(Z{=}z) > 0$. This corollary immediately follows from RS-Propositions (6.47) and (6.48).

Corollary 8.19 [$D_X \perp\!\!\!\perp X \mid Z$ Implies $(Z{=}z)$-Conditional Independence of D_X and X]
If the Assumptions 8.1 (a), (d), and (e) hold, then

$$D_X \perp\!\!\!\perp X \mid Z \quad \Rightarrow \quad D_X \perp\!\!\!\perp X \mid (Z{=}z)\,. \tag{8.35}$$

Hence, under the assumptions of Corollary 8.19, Z-conditional independence of a putative cause variable X and a global potential confounder of X implies $(Z{=}z)$-conditional independence of X and D_X.

Propositions (8.29) and (8.35) immediately yield the following corollary.

Corollary 8.20 [$D_X \perp\!\!\!\perp X$ **Implies** $(Z{=}z)$**-Conditional Independence of** D_X **and** X]
Let the Assumptions 8.1 (a), (d), and (e) hold, and let Z be a covariate of X. Then

$$D_X \perp\!\!\!\perp X \quad \Rightarrow \quad D_X \perp\!\!\!\perp X \mid (Z{=}z)\,. \tag{8.36}$$

Hence, under the assumptions of this corollary, independence of a putative cause variable X and a global potential confounder of X implies $(Z{=}z)$-conditional independence of X and D_X.

Remark 8.21 [A Consequence of $D_X \perp\!\!\!\perp X \mid (Z{=}z)$**]** Under the Assumptions 8.1 (a), (b), (d), and (e),

$$D_X \perp\!\!\!\perp X \mid (Z{=}z) \quad \Rightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} \mid (Z{=}z) \tag{8.37}$$

[see solution (1) to Exercise 8-2]. ◁

Note that the assumptions of the theorems and the other propositions stated above neither include $P(X{=}x) > 0$ nor that Y is real-valued, which are prerequisites for the true outcome variables τ_x to be defined. Hence, these propositions also hold beyond true outcome theory. This remark also applies to all other propositions summarized in Table 8.1. (For the proofs of these propositions, see the solution to Exercise 8-2.)

8.3 Implications of F-Conditions on RR-Conditions and Unbiasedness

In this section we treat the consequences of the Fisher conditions on the Rosenbaum-Rubin conditions and on unbiasedness. Now our assumptions include that the putative cause variable X is finite with values in the set $X(\Omega) = \{0, 1, \ldots, J\}$, $P(X{=}x) > 0$ for all values of X, and that Y is real-valued. These assumptions are prerequisites for the definitions of the true outcome variables τ_x, $x \in X(\Omega)$.

Table 8.2 summarizes the consequences of the Fisher conditions on the Rosenbaum-Rubin conditions. These consequences are proved in the theorems treated in this section or in Exercise 8-3. We also prove the implications of the Fisher conditions on unbiasedness.

8.3.1 Consequences of Independence of D_X and X

In the first theorem of this section, we consider consequences of $D_X \perp\!\!\!\perp X$ for the multivariate true outcome variable $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$ and unbiasedness of the conditional expectation $E(Y|X)$ and its values $E(Y|X{=}x)$. In this theorem, we confine ourselves to those consequences that do not involve conditioning on a covariate of X.

Theorem 8.22 [Consequences of $D_X \perp\!\!\!\perp X$ for True Outcomes and Unbiasedness]
Let the Assumptions 8.1 (a), (c), and (g) hold. Then $D_X \perp\!\!\!\perp X$ implies

(i) $\tau \perp\!\!\!\perp X$
(ii) $\forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x}$
(iii) $\forall\, x \in X(\Omega)\colon \tau_x$ *is P-unique*
(iv) $\forall\, x \in X(\Omega)\colon E(Y|X{=}x) \vdash D_X$
(v) $E(Y|X) \vdash D_X$.

(Proof p. 257)

Remark 8.23 [$D_X \perp\!\!\!\perp X$ Implies That All τ_x Are P-Unique] Under the assumptions of Theorem 8.22, $D_X \perp\!\!\!\perp X$ implies that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique [see Prop. (iii) of Th. 8.22]. Hence, P-uniqueness of the true outcome variables τ_x is not an additional assumption in Proposition (v) of Theorem 8.22, according to which independence of D_X and X implies unbiasedness of $E(Y|X)$. ◁

Remark 8.24 [$D_X \perp\!\!\!\perp X$ Implies the Rosenbaum-Rubin Conditions] Note that $D_X \perp\!\!\!\perp X$ does not only imply $\tau \perp\!\!\!\perp X$ and P-uniqueness of all true outcome variables τ_x [see Props. (i) and (iii) of Th. 8.22], but also all other simple Rosenbaum-Rubin conditions (see the last row of Table 7.1). ◁

In the next theorem we consider some consequences of independence of an indicator $1_{X=x}$ for a value x of the putative cause variable X and a global potential confounder D_X of X. Reading this theorem, remember that $E^{X=x}(Y|D_X)$ is the more explicit notation of a true outcome variable τ_x (see Def. 5.4).

Theorem 8.25 [Consequences of $D_X \perp\!\!\!\perp 1_{X=x}$]
Let the Assumptions 8.1 (a) to (c) hold. Then $D_X \perp\!\!\!\perp 1_{X=x}$ implies

(i) $E^{X=x}(Y|D_X) \perp\!\!\!\perp 1_{X=x}$ *and* $E^{X\neq x}(Y|D_X) \perp\!\!\!\perp 1_{X\neq x}$
(ii) $E^{X=x}(Y|D_X)$ *and* $E^{X\neq x}(Y|D_X)$ *are P-unique*
(iii) $E(Y|X{=}x) \vdash D_X$ *and* $E(Y|X{\neq}x) \vdash D_X$
(iv) $E(Y|\,1_{X=x}) \vdash D_X$.

(Proof p. 257)

Hence, according to this theorem, $D_X \perp\!\!\!\perp 1_{X=x}$ implies that the true outcome variable $E^{X=x}(Y|D_X)$ and the indicator $1_{X=x}$ are independent. Correspondingly, $D_X \perp\!\!\!\perp 1_{X=x}$ implies that $E^{X\neq x}(Y|D_X)$ and the indicator $1_{X\neq x}$ are independent. It also implies that the true outcome variables $E^{X=x}(Y|D_X)$ and $E^{X\neq x}(Y|D_X)$ are P-unique. Finally, $D_X \perp\!\!\!\perp 1_{X=x}$ implies that $E(Y|\,1_{X=x})$ and its values $E(Y|X{=}x)$ and $E(Y|X{\neq}x)$ are unbiased.

Table 8.2. Implications of the Fisher conditions on some Rosenbaum-Rubin conditions

	$\tau \perp\!\!\!\perp 1_{X=x} \mid (Z=z)$	$\tau \perp\!\!\!\perp X \mid (Z=z)$	$\tau \perp\!\!\!\perp 1_{X=x} \mid Z$	$\tau \perp\!\!\!\perp X \mid Z$	$\tau \perp\!\!\!\perp 1_{X=x}$	$\tau \perp\!\!\!\perp X$
$D_X \perp\!\!\!\perp 1_{X=x} \mid (Z=z)$	(a)-(e),(g)					
$D_X \perp\!\!\!\perp X \mid (Z=z)$	(a)-(e),(g)	(a)-(e),(g)				
$D_X \perp\!\!\!\perp 1_{X=x} \mid Z$	(a)-(e),(g)		(a)-(d),(g)			
$D_X \perp\!\!\!\perp X \mid Z$	(a)-(g)	(a)-(e),(g)	(a)-(d),(g)	(a)-(d),(g)		
$D_X \perp\!\!\!\perp 1_{X=x}$	(a)-(g),(h)		(a)-(d),(g),(h)		(a),(c),(g)	
$D_X \perp\!\!\!\perp X$	(a)-(g),(h)	(a)-(g),(h)	(a)-(d),(g),(h)	(a)-(d),(g),(h)	(a),(c),(g)	(a),(c),(g)

Note: An entry such as (a)-(g) means that the condition in the row implies the condition in the column, provided that the Assumptions 8.1 (a) to (g) hold. Consequences of the Rosenbaum-Rubin conditions displayed in the columns of the table are found in Tables 7.1 and 7.2.

Remark 8.26 [Further Consequences of $D_X \perp\!\!\!\perp 1_{X=x}$] Because $\sigma(1_{X=x}) = \sigma(1_{X \neq x})$, Theorem 8.25 (i) immediately yields

$$E^{X=x}(Y \mid D_X) \perp\!\!\!\perp 1_{X \neq x} \qquad \text{and} \qquad E^{X \neq x}(Y \mid D_X) \perp\!\!\!\perp 1_{X=x}\,. \tag{8.38}$$

$\triangleleft$

In the next theorem we consider some consequences of $D_X \perp\!\!\!\perp X$ involving conditioning on a covariate Z of X. These consequences include unbiasedness of the conditional expectations $E^{X=x}(Y \mid D_X)$, $x \in X(\Omega)$, and $E(Y \mid X, Z)$.

Theorem 8.27 [Consequences of $D_X \perp\!\!\!\perp X$ Involving Conditioning on a Covariate]
Let the Assumptions 8.1 (a) to (c) hold and Z be a covariate of X. Then $D_X \perp\!\!\!\perp X$ implies

(i) $\tau \perp\!\!\!\perp X \mid Z$
(ii) $\forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} \mid Z$
(iii) $\forall\, x \in X(\Omega)\colon E^{X=x}(Y \mid Z) \vdash D_X$
(iv) $E(Y \mid X, Z) \vdash D_X$.

(Proof p. 258)

Remark 8.28 [$D_X \perp\!\!\!\perp X$ Implies $\tau \perp\!\!\!\perp X \mid Z$] Under the assumptions of Theorem 8.27, $D_X \perp\!\!\!\perp X$ implies Z-conditional independence of the multivariate true outcome variable τ and the putative cause variable X. Remember, according to Theorem 8.22 (iii), $D_X \perp\!\!\!\perp X$ implies P-uniqueness of all true outcome variables τ_x, $x \in X(\Omega)$. Under the assumptions of Theorem 8.27, $D_X \perp\!\!\!\perp X$ also implies all consequences of $\tau \perp\!\!\!\perp X \mid Z$ listed in the last row of Table 7.2. Hence, under these assumptions, $D_X \perp\!\!\!\perp X$ does not only imply that $E(Y \mid X)$ and each $E(Y \mid X{=}x)$, $x \in X(\Omega)$, are unbiased, but it also implies unbiasedness of $E(Y \mid X, Z)$ and each

Z-conditional expectation $E^{X=x}(Y|Z)$, $x \in X(\Omega)$. To emphasize, under these assumptions and $D_X \perp\!\!\!\perp X$, an additional assumption of P-uniqueness of the true outcome variables τ_x is not necessary because this property already follows from $D_X \perp\!\!\!\perp X$. ◁

Now we turn to some implications of $D_X \perp\!\!\!\perp X$ that involve conditioning on a value z of Z, that is, conditioning on the event $\{Z{=}z\}$. According to the next theorem, $D_X \perp\!\!\!\perp X$ implies $(Z{=}z)$-conditional independence of τ and X, provided that Z is a covariate of X. Assuming $P(Z{=}z) > 0$ and $P(X{=}x) > 0$, for all values x of X, conditional independence of D_X and X also implies unbiasedness of each conditional expectation value $E^{Z=z}(Y|X{=}x) = E(Y|X{=}x, Z{=}z)$, $x \in X(\Omega)$, and unbiasedness of the X-conditional expectation $E^{Z=z}(Y|X)$.

Theorem 8.29 [$D_X \perp\!\!\!\perp X$ **Implies** $(Z{=}z)$**-Conditional Independence of** τ **and** X]
Let the Assumptions 8.1 (a) to (g) hold and assume that Z is a covariate of X. Then $D_X \perp\!\!\!\perp X$ implies

(i) $\forall\, x \in X(\Omega){:}\; P(X{=}x, Z{=}z) > 0$

(ii) $\tau \perp\!\!\!\perp X \,|\, (Z{=}z)$

(iii) $\forall\, x \in X(\Omega){:}\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$

(iv) $\forall\, x \in X(\Omega){:}\; E^{Z=z}(Y|X{=}x) \vdash D_X$

(v) $E^{Z=z}(Y|X) \vdash D_X$.

(Proof p. 259)

Remark 8.30 [Other Consequences of $D_X \perp\!\!\!\perp X$ **Involving an Event** $\{Z{=}z\}$**]** Hence, under the assumptions of Theorem 8.29, $D_X \perp\!\!\!\perp X$ implies that τ and X are $(Z{=}z)$-conditionally independent. Corollary 7.14 lists a number of consequences that follow from $\tau \perp\!\!\!\perp X \,|\, (Z{=}z)$, provided that all true outcome variables τ_x are $P^{Z=z}$-unique. Remember that $D_X \perp\!\!\!\perp X$ implies that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique [see Th. 8.22 (iii)], which in turn implies that they are also $P^{Z=z}$-unique [see RS-Box 5.1 (v)]. Hence, because of Propositions (i) and (ii) of Theorem 8.29, all the consequences listed in Corollary 7.14 also follow from $D_X \perp\!\!\!\perp X$, provided that Z is a covariate of X. No additional assumption of P-uniqueness of the true outcome variables τ_x, $x \in X(\Omega)$, is required because it already follows from $D_X \perp\!\!\!\perp X$. The most important consequences of $D_X \perp\!\!\!\perp X$ are unbiasedness of $E^{Z=z}(Y|X)$ and the conditional expectation values $E^{Z=z}(Y|X{=}x)$, $x \in X(\Omega)$, provided, of course, that the assumptions of Theorem 8.29 hold. ◁

8.3.2 Consequences of Z-Conditional Independence of D_X and X

Now we turn to the consequences of Z-conditional independence of D_X and X. Reading the second part of this theorem, note that P-uniqueness of a true outcome variable τ_x follows from the conjunction of $D_X \perp\!\!\!\perp X \,|\, Z$ and $P(X{=}x \,|\, Z) \underset{P}{>} 0$ (see Th. 8.34).

Theorem 8.31 [Consequences of $D_X \perp\!\!\!\perp X \,|\, Z$**]**
Let the Assumptions 8.1 (a) to (d) and (g) hold. Then $D_X \perp\!\!\!\perp X \,|\, Z$ implies

(i) $\tau \perp\!\!\!\perp X \,|\, Z$

(ii) $\forall\, x \in X(\Omega)\colon\ \tau \perp\!\!\!\perp 1_{X=x} | Z$

(iii) $\forall\, x, x' \in X(\Omega)\colon\ \tau_x \perp\!\!\!\perp 1_{X=x'} | Z$.

If, additionally, all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique and Z is a covariate of X, then $D_X \perp\!\!\!\perp X | Z$ also implies

(iv) $\forall\, x \in X(\Omega)\colon\ E^{X=x}(Y|Z) \vdash D_X$

(v) $E(Y|X,Z) \vdash D_X$.

(Proof p. 260)

To emphasize, Z-conditional independence of D_X and X alone is not sufficient for unbiasedness of $E(Y|X,Z)$, even if Z is a covariate of X. This will be exemplified in Example 8.33. However, in Theorem 8.34 we show that the conjunction of $D_X \perp\!\!\!\perp X | Z$ and $P(X{=}x|Z) \underset{P}{>} 0$, for all $x \in X(\Omega)$, implies P-uniqueness of all true outcome variables τ_x, $x \in X(\Omega)$.

Remark 8.32 [$D_X \perp\!\!\!\perp X | Z$ **Implies the Rosenbaum-Rubin Conditions]** Note that $D_X \perp\!\!\!\perp X | Z$ does not only imply $\tau \perp\!\!\!\perp X | Z$ but, if all τ_x, $x \in X(\Omega)$, are P-unique, also all other conditional Rosenbaum-Rubin conditions [see Prop. (i) of Th. 8.31 and the last row of Table 7.2]. Furthermore, if Z is a covariate of X, then $D_X \perp\!\!\!\perp X | Z$ also implies unbiasedness of $E(Y|X,Z)$ and all conditional expectations $E^{X=x}(Y|Z)$, $x \in X(\Omega)$. ◁

Example 8.33 [No Treatment For Males] Table 8.3 displays the parameters of a random experiment in which males have a zero probability to be treated. We assume that this random experiment has the same structure as the random experiments treated in section 6.5. That is,

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big),$$

as specified in section 6.5, is the regular probabilistic causality setup. Again, U takes the role of a global potential confounder D_X of X and $Z = sex$ is a covariate of X because $\sigma(Z) \subset \sigma(U)$. This implies

$$P(X{=}1|U,Z) \underset{P}{=} P(X{=}1|U)$$

because $\sigma(U,Z) = \sigma(U)$ [see RS-Prop. (2.19), RS-Def. 4.4, and RS-Rem. 4.12].

In this example, there is Z-conditional independence of U and X but the true outcome variable τ_1 is not P-unique. Therefore, unbiasedness of $E^{X=1}(Y|Z)$ and unbiasedness of $E(Y|X,Z)$ are not defined in this example.

We start checking if $U \perp\!\!\!\perp X | Z$ holds. According to RS-Theorem 6.6, if X is binary, then

$$U \perp\!\!\!\perp X | Z \quad \Leftrightarrow \quad P(X{=}1|U,Z) \underset{P}{=} P(X{=}1|Z).$$

Because, in this example, $\sigma(Z) \subset \sigma(U)$, this simplifies to

$$U \perp\!\!\!\perp X | Z \quad \Leftrightarrow \quad P(X{=}1|U) \underset{P}{=} P(X{=}1|Z).$$

Inspecting the column headed $P(X{=}1|U{=}u)$ in Table 8.3 shows that

Table 8.3. No treatment for males: Z-conditional independence of X and D_X

Fundamental parameters

Person u	Sex z	$P(U=u)$	$P(X=1 \mid U=u)$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$CTE_{U;10}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$
Joe	m	1/4	0	68	999^*	*ndef*	2/5	0
Jim	m	1/4	0	78	-999^*	*ndef*	2/5	0
Ann	f	1/4	3/4	106	114	8	1/10	1/2
Sue	f	1/4	3/4	116	130	14	1/10	1/2

	$x=0$	$x=1$	
$E(\tau_x)$:	92.0	111^*	$ATE_{10} =$ *ndef*
$E(Y \mid X=x)$:	80.6	122	$PFE_{10} = 41.4$
$E(\tau_x \mid Z=m)$:	73	0^*	$CTE_{Z;10}(m) =$ *ndef*
$E(Y \mid X=x, Z=m)$:	73	888^*	$PFE_{Z;10}(m) = 815^*$
$E(\tau_x \mid Z=f)$:	111	122	$CTE_{Z;10}(f) = 11$
$E(Y \mid X=x, Z=f)$:	111	122	$PFE_{Z;10}(f) = 11$

Note: An asterisk * indicates that this number is arbitrary and could be any other real number because, in this example, the corresponding term is not uniquely defined (see RS-sect. 4.4 for more details). Furthermore, *ndef* means that this term is not defined in this example.

$$P(X=1 \mid U)(\omega) \;=\; P(X=1 \mid Z)(\omega) \;=\; \begin{cases} 0, & \text{if } \omega \in \{Z=m\} \\ 3/4, & \text{if } \omega \in \{Z=f\}, \end{cases} \tag{8.39}$$

which proves $U \perp\!\!\!\perp X \mid Z$.

Now we show that, in this example, the true outcome variable τ_1 is not P-unique. Because U is a global potential confounder of X, P-uniqueness of τ_1 is equivalent to $P(X=1 \mid U) \underset{P}{>} 0$ (see RS-Th. 5.27), which in turn is equivalent to

$$P\big(\{\omega \in \Omega \colon P(X=1 \mid U)(\omega) > 0\}\big) \;=\; 1.$$

[see RS-Eq. (4.12)]. Looking at the columns headed $P(U=u)$ and $P(X=1 \mid U=u)$ in Table 8.3 shows that

$$P\big(\{\omega \in \Omega \colon P(X=1 \mid U)(\omega) > 0\}\big) \;=\; 1/4 + 1/4 \;=\; 1/2.$$

Therefore, $P(X=1 \mid U) \underset{P}{>} 0$ and with it P-uniqueness of $\tau_1 = E^{X=1}(Y \mid U)$ does not hold (see Exercise 8-4). Hence, there is at least one other version τ_1^* of the U-conditional expectation of Y with respect to the measure $P^{X=1}$ for which $\tau_1 \underset{P}{=} \tau_1^*$ does not hold (see Exercise 8-5). $\triangleleft$

For the second part of Theorem 8.31 we need the additional assumption that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique. In the following theorem we present a condition under which Z-conditional independence of D_X and X implies that a true outcome variable τ_x is P-unique.

Theorem 8.34 [A Condition Under Which $D_X \perp\!\!\!\perp X \,|\, Z$ Implies P-Uniqueness of τ_x]
Let the Assumptions 8.1 (a) to (d) hold and assume that Z is a covariate of X. Then

$$D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \;\wedge\; P(X{=}x \,|\, Z) \underset{P}{>} 0 \quad\Rightarrow\quad \tau_x \textit{ is } P\textit{-unique.} \tag{8.40}$$

If we additionally assume 8.1 (g), then

$$D_X \perp\!\!\!\perp X \,|\, Z \;\wedge\; \big(\forall\, x \in X(\Omega)\colon P(X{=}x \,|\, Z) \underset{P}{>} 0\big) \quad\Rightarrow\quad \forall\, x \in X(\Omega)\colon \tau_x \textit{ is } P\textit{-unique.} \tag{8.41}$$

(Proof p. 260)

Remark 8.35 [$D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z$ Transfers Positivity From $P(X{=}x \,|\, Z)$ to $P(X{=}x \,|\, D_X)$] According to RS-Theorem 5.27,

$$P(X{=}x \,|\, Z) \underset{P}{>} 0 \quad\Leftrightarrow\quad E^{X=x}(Y \,|\, Z) \text{ is } P\text{-unique.}$$

However, according to Proposition (8.40), if Z is a covariate of X, then the conjunction $D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \wedge P(X{=}x \,|\, Z) \underset{P}{>} 0$ implies that not only $E^{X=x}(Y \,|\, Z)$ is P-unique, but also a true outcome variable $E^{X=x}(Y \,|\, D_X) = \tau_x$. RS-Theorem 5.27 provides some other conditions that are equivalent to $P(X{=}x \,|\, Z) \underset{P}{>} 0$. ◁

Theorems 8.31 and 8.34 immediately imply the following corollary.

Corollary 8.36 [Consequences of $D_X \perp\!\!\!\perp X \,|\, Z$ For Unbiasedness of $E(Y \,|\, X, Z)$]
Let the Assumptions 8.1 (a) to (d) and (g) hold, and assume that Z is a covariate of X. Then

$$\begin{aligned} & D_X \perp\!\!\!\perp X \,|\, Z \;\wedge\; \big(\forall\, x \in X(\Omega)\colon P(X{=}x \,|\, Z) \underset{P}{>} 0\big) \\ \Rightarrow\quad & \forall\, x \in X(\Omega)\colon E^{X=x}(Y \,|\, Z) \vdash D_X && (8.42) \\ \Leftrightarrow\quad & E(Y \,|\, X, Z) \vdash D_X. && (8.43) \end{aligned}$$

Note that $D_X \perp\!\!\!\perp X \,|\, Z$ is not sufficient for unbiasedness of the conditional expectations $E^{X=x}(Y \,|\, Z)$ and $E(Y \,|\, X, Z)$. However, in conjunction with $P(X{=}x \,|\, Z) \underset{P}{>} 0, \forall\, x \in X(\Omega)$, it is, provided that Z is a covariate of X. In contrast, if Z is a covariate of X and $P(X{=}x) > 0$ for all $x \in X(\Omega)$, then $D_X \perp\!\!\!\perp X$ is sufficient for unbiasedness of the conditional expectations $E^{X=x}(Y \,|\, Z)$, $x \in X(\Omega)$, and $E(Y \,|\, X, Z)$ (see also Exercises 8-6 and 8-7).

In the following theorem, we consider consequences of Z-conditional independence of a putative cause variable X and a global potential confounder D_X of X on $(Z{=}z)$-conditional independence of X and the multivariate true outcome variable τ.

Theorem 8.37 [$D_X \perp\!\!\!\perp X | Z$ **Implies** $(Z{=}z)$**-Conditional Independence of** τ **And** X]
Let the Assumptions 8.1 (a) to (e) and (g) hold. Then $D_X \perp\!\!\!\perp X | Z$ *implies*

(i) $\forall\, x \in X(\Omega)$: τ_x *is* $P^{Z=z}$*-unique*
(ii) $\tau \perp\!\!\!\perp X | (Z{=}z)$
(iii) $\forall\, x \in X(\Omega)$: $\tau \perp\!\!\!\perp 1_{X=x} | (Z{=}z)$

If we additionally assume that Z *is a covariate of* X*, then*

(iv) $\forall\, x \in X(\Omega)$: $E^{Z=z}(Y | X{=}x) \vdash D_X$
(v) $E^{Z=z}(Y | X) \vdash D_X$.

(Proof p. 260)

Hence, under the assumptions of Theorem 8.37, $D_X \perp\!\!\!\perp X | Z$ implies that all true outcome variables τ_x, $x \in X(\Omega)$, are $P^{Z=z}$-unique and that $\tau = (\tau_0, \tau_1, \dots, \tau_J)$ and X are $(Z{=}z)$-conditionally independent. Furthermore, under the additional assumption that Z is a covariate of X, $D_X \perp\!\!\!\perp X | Z$ also implies unbiasedness of all conditional expectation values $E^{Z=z}(Y | X{=}x)$, $x \in X(\Omega)$, and of the X-conditional expectation $E^{Z=z}(Y | X)$ of Y with respect to the measure $P^{Z=z}$.

Remark 8.38 [Consequences of $\tau \perp\!\!\!\perp X | (Z{=}z)$**]** Corollary 7.14 lists some other consequences that follow from the conjunction of $\tau \perp\!\!\!\perp X | (Z{=}z)$ and $P^{Z=z}$-uniqueness of the true outcome variables. Because of Proposition (ii) of Theorem 8.37, these consequences also follow from $D_X \perp\!\!\!\perp X | Z$, provided that the assumptions of Theorem 8.37 hold. ◁

In Theorem 8.37 and Remark 8.38, we only consider a single value z of the random variable Z and assume that $P(Z{=}z) > 0$. In contrast, in the following theorem we assume $P(X{=}x, Z{=}z) > 0$ *for all pairs* (x, z) of values of X and Z, which implies $P(X{=}x) > 0$ for all values x of X and $P(Z{=}z) > 0$ for all values z of Z.

Theorem 8.39 [More Consequences of $D_X \perp\!\!\!\perp X | Z$**]**
Let the Assumptions 8.1 (a) to (e) hold and assume that $P(X{=}x, Z{=}z) > 0$ *for all pairs* $(x, z) \in X(\Omega) \times Z(\Omega)$. *Then* $D_X \perp\!\!\!\perp X | Z$ *implies*

(i) $\forall\, x \in X(\Omega)$: τ_x *is* P*-unique.*

If we additionally assume that Z *is a covariate of* X*, then*

(ii) $\forall\, x \in X(\Omega)$: $E^{X=x}(Y | Z) \vdash D_X$
(iii) $E(Y | X, Z) \vdash D_X$.

(Proof p. 261)

Hence, under the assumptions of Theorem 8.39, we can conclude that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique and, if Z is a covariate of X, then it also follows that all conditional expectations $E^{X=x}(Y | Z)$, $x \in X(\Omega)$, as well as $E(Y | X, Z)$ are unbiased.

Remark 8.40 [Methodological Consequences] Theorem 8.39 has important methodological consequences. The two crucial assumptions under which the conditional expectation

$E(Y|X,Z)$ is unbiased can often be created by the experimenter. The first is the assumption that $P(X{=}x, Z{=}z) > 0$ for all pairs (x,z) of values of X and Z. If, for example, X is a binary treatment variable and Z the binary covariate sex, then the experimenter simply has to make sure that there is a positive probability for each of the four pairs (x,z) of values of X and Z under which the outcome variable Y is observed. $\triangleleft$

8.3.3 Consequences of $(Z{=}z)$-Conditional Independence of D_X and X

Now we turn to some consequences of $(Z{=}z)$-conditional independence of D_X and X, which has been introduced in Remark 8.12 by independence of D_X and X with respect to the probability measure $P^{Z=z}$.

Theorem 8.41 [Consequences of $D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$]
Let the Assumptions 8.1 (a) to (e) and (g) hold. Then $D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$ implies

(i) $\tau \perp\!\!\!\perp X \,|\, (Z{=}z)$

(ii) $\forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.

If we additionally assume $P^{Z=z}(X{=}x) > 0$ for all $x \in X(\Omega)$, then $D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$ implies

(iii) $\forall\, x \in X(\Omega)\colon \tau_x$ *is* $P^{Z=z}$*-unique.*

If we additionally assume $P^{Z=z}(X{=}x) > 0$ for all $x \in X(\Omega)$ and Z is a covariate of X, then $D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$ also implies

(iv) $\forall\, x \in X(\Omega)\colon E^{Z=z}(Y \,|\, X{=}x) \vdash D_X$

(v) $E^{Z=z}(Y \,|\, X) \vdash D_X$.

(Proof p. 261)

Hence, under the assumptions of Theorem 8.41, $D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$ implies that the multivariate true outcome variable τ and X are $(Z{=}z)$-conditionally independent. This implies that τ and each indicator variable $1_{X=x}$, $x \in X(\Omega)$, are $(Z{=}z)$-conditionally independent. If we additionally assume $P^{Z=z}(X{=}x) > 0$ for all $x \in X(\Omega)$, then $D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$ also implies that all true outcome variables τ_x, $x \in X(\Omega)$, are $P^{Z=z}$-unique. Finally, if Z is a covariate of X, then $D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$ also implies and that the conditional expectation $E^{Z=z}(Y|X)$ and all the conditional expectation values $E^{Z=z}(Y|X{=}x) = E(Y|X{=}x, Z{=}z)$, $x \in X(\Omega)$, are unbiased. Other implications can be found in the last row of Table 7.2.

In the next theorem we consider some consequences of $(Z{=}z)$-conditional independence a global potential confounder D_X of X and an indicator $1_{X=x}$ for a value x of the putative cause variable X.

Theorem 8.42 [Some Consequences of $D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$]
Let the Assumptions 8.1 (a) to (e) hold. Then $D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$ implies

(i) $E^{X=x}(Y \,|\, D_X) \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$ *and* $E^{X\neq x}(Y \,|\, D_X) \perp\!\!\!\perp 1_{X\neq x} \,|\, (Z{=}z)$.

If we additionally assume $0 < P^{Z=z}(X{=}x) < 1$, then $D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$ implies

(ii) $E^{X=x}(Y \,|\, D_X)$ *and* $E^{X\neq x}(Y \,|\, D_X)$ *are* $P^{Z=z}$*-unique.*

If, additionally, $0 < P^{Z=z}(X=x) < 1$ *and* Z *is a covariate of* X*, then* $D_X \perp\!\!\!\perp 1_{X=x} | (Z=z)$ *also implies*

(iii) $E^{Z=z}(Y|X=x) \vdash D_X$ *and* $E^{Z=z}(Y|X\neq x) \vdash D_X$

(iv) $E^{Z=z}(Y|1_{X=x}) \vdash D_X$.

(Proof p. 262)

Hence, according to Theorem 8.42, $(Z=z)$-conditional independence of the indicator $1_{X=x}$ and a global potential confounder D_X of X implies that the true outcome variable $E^{X=x}(Y|D_X)$ and the indicator $1_{X=x}$ as well as $E^{X\neq x}(Y|D_X)$ and the indicator $1_{X\neq x}$ are $(Z=z)$-conditionally independent. If we additionally assume $0 < P^{Z=z}(X=x) < 1$, then $D_X \perp\!\!\!\perp 1_{X=x} | (Z=z)$ also implies that $E^{X=x}(Y|D_X)$ and $E^{X\neq x}(Y|D_X)$ are $P^{Z=z}$-unique. Finally, we also add the assumption that Z is a covariate of X, then $D_X \perp\!\!\!\perp 1_{X=x} | (Z=z)$ implies that $E^{Z=z}(Y|1_{X=x})$ and the conditional expectation values $E^{Z=z}(Y|X=x)$ and $E^{Z=z}(Y|X\neq x)$ are unbiased.

8.4 Unbiasedness of Prima Facie Effects and Effect Functions

In section 6.4 we showed how to identify average and conditional causal total effects and effect functions from unbiased prima facie effects and unbiased prima facie effect functions. Now we turn to sufficient conditions for unbiasedness of conditional and unconditional prima facie effects and effect functions. Hence, this section is also crucial for the identification of conditional and average causal total effects.

8.4.1 Unbiasedness of the Prima Facie Effect

In the following theorem we specify the conditions under which the prima facie effect

$$PFE_{xx'} \;=\; E(Y|X=x) - E(Y|X=x') \tag{8.44}$$

is unbiased, that is, under which $PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}$. Hence, according to Definition 6.23 (i), we specify conditions under which τ_x and $\tau_{x'}$ are P-unique and the prima facie effect $PFE_{xx'}$ is identical to the average causal total effect

$$ATE_{xx'} \;=\; E(\tau_x - \tau_{x'}) \;=\; E(\tau_x) - E(\tau_{x'}). \tag{8.45}$$

Theorem 8.43 [An F-Condition Implying Unbiasedness of the Prima Facie Effect]
Let the Assumptions 8.1 (a) to (c), and (f) hold. Then

$$\big(D_X \perp\!\!\!\perp 1_{X=x} \wedge D_X \perp\!\!\!\perp 1_{X=x'}\big) \quad\Rightarrow\quad PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}. \tag{8.46}$$

(Proof p. 263)

Remark 8.44 [Identification of $ATE_{xx'}$] Hence, under the assumptions of Theorem 8.43, the conjunction of (a) independence of the indicator $1_{X=x}$ and a global potential confounder D_X of X and (b) independence of the indicator $1_{X=x'}$ and D_X implies P-uniqueness of τ_x and $\tau_{x'}$ as well as

$$PFE_{xx'} = E(Y|X{=}x) - E(Y|X{=}x') = ATE_{xx'} = E(\tau_x) - E(\tau_{x'}). \tag{8.47}$$

That is, the conjunction $D_X \perp\!\!\!\perp 1_{X=x} \wedge D_X \perp\!\!\!\perp 1_{X=x'}$ implies that the difference between the conditional expectation values $E(Y|X{=}x)$ and $E(Y|X{=}x')$ is identical to the causal average total effect of x compared to x'. In this context we also say that the $ATE_{xx'}$ is *identified* by that difference. Note that P-uniqueness of τ_x follows from $D_X \perp\!\!\!\perp 1_{X=x}$ and P-uniqueness of $\tau_{x'}$ follows from $D_X \perp\!\!\!\perp 1_{X=x'}$. It is not an additional assumption. ◁

Remark 8.45 [$D_X \perp\!\!\!\perp X$ Implies $PFE_{xx'} = ATE_{xx'}$] Also note that

$$D_X \perp\!\!\!\perp X \quad \Rightarrow \quad \big(D_X \perp\!\!\!\perp 1_{X=x} \wedge D_X \perp\!\!\!\perp 1_{X=x'}\big), \tag{8.48}$$

which follows from $\sigma(1_{X=x}), \sigma(1_{X=x'}) \subset \sigma(X)$ and RS-Box 2.1 (iv). Hence, under the assumptions of Theorem 8.43,

$$D_X \perp\!\!\!\perp X \quad \Rightarrow \quad PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}} \tag{8.49}$$

$$\Rightarrow \quad PFE_{xx'} = ATE_{xx'}. \tag{8.50}$$

This means, under $D_X \perp\!\!\!\perp X$, the average causal total effect $ATE_{xx'}$ is *identified* by the prima facie effect $ATE_{xx'}$. ◁

8.4.2 Unbiasedness of a Z-Conditional Prima Facie Effect Function

In the next theorem we specify two conditions under each of which a Z-conditional prima facie effect variable

$$PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z) \tag{8.51}$$

is unbiased. Remember, in Definition 6.23 (ii), we introduced the notation

$$PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}} \quad \Leftrightarrow \quad \big(\tau_x \text{ and } \tau_{x'} \text{ are } P\text{-unique} \wedge PFE_{Z;xx'}(Z) \underset{P}{=} CTE_{Z;xx'}(Z)\big) \tag{8.52}$$

for *unbiasedness of $PFE_{Z;xx'}(Z)$*, where

$$CTE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x - \tau_{x'}|Z) \tag{8.53}$$

denotes a causal Z-conditional total effect variable.

Theorem 8.46 [Unbiasedness of a Z-Conditional Prima Facie Effect Variable]
If the Assumptions 8.1 (a) to (d), and (f) hold, $P(X{=}x|Z), P(X{=}x'|Z) \underset{P}{>} 0$, and Z is a covariate of X, then each of the following conditions implies $PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$:

(i) $D_X \perp\!\!\!\perp 1_{X=x}|Z \wedge D_X \perp\!\!\!\perp 1_{X=x'}|Z$

(ii) $D_X \perp\!\!\!\perp X|Z$.

(Proof p. 264)

Remark 8.47 [Identification of $CTE_{Z;xx'}(Z)$] Hence, under the assumptions of Theorem 8.46, the conjunction of Z-conditional independence of a global potential confounder D_X

of X and the indicator $1_{X=x}$ and of D_X and $1_{X=x'}$ implies that the Z-conditional prima facie effect variable $PFE_{Z;xx'}(Z)$ is unbiased.

Under the same assumptions, the same conclusion can be drawn from Z-conditional independence of D_X and X. Because unbiasedness of $PFE_{Z;xx'}(Z)$ implies

$$PFE_{Z;xx'}(Z) \underset{P}{=} CTE_{Z;xx'}(Z)$$

[see Prop. (8.52)], we also say that the causal Z-conditional total effect variable $CTE_{Z;xx'}(Z)$ *is identified by* $PFE_{Z;xx'}(Z)$, provided that the assumptions of Theorem 8.46 hold. ◁

Remark 8.48 [$PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$ **And the Identification of the Causal Average Total Effect**] If $PFE_{Z;xx'}$ is unbiased, then we can not only identify the causal Z-conditional total effect variable $CTE_{Z;xx'}(Z)$ (see Rem. 8.47), but also the causal average total effect $ATE_{xx'}$, *even if* $D_X \perp\!\!\!\perp X$ *and Equation* (8.47) *do not hold.* For details see Theorem 6.34 and the example presented in Table 6.4. ◁

According to the following theorem, independence of X and a global potential confounder D_X of X is a sufficient condition for unbiasedness of a Z-conditional prima facie effect function $PFE_{Z;xx'}$, and it does not require the additional assumptions $P(X{=}x\,|\,Z) \underset{P}{>} 0$ and $P(X{=}x'|\,Z) \underset{P}{>} 0$ that are indispensable in Theorem 8.46. Instead, these additional assumption already follow from $D_X \perp\!\!\!\perp X$.

Theorem 8.49 [Again Unbiasedness of a Z-Conditional Prima Facie Effect Variable]
If the Assumptions 8.1 (a) to (d) and (f) hold, and Z is a covariate of X, then each of the following conditions implies $PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}$:

(i) $D_X \perp\!\!\!\perp 1_{X=x} \;\wedge\; D_X \perp\!\!\!\perp 1_{X=x'}$

(ii) $D_X \perp\!\!\!\perp X$.

(Proof p. 264)

8.4.3 Unbiasedness of a $(Z{=}z)$-Conditional Prima Facie Effect

Now we consider a single value z of Z and specify conditions under which a $(Z{=}z)$-conditional prima facie effect

$$PFE_{Z;xx'}(z) \;=\; E^{X=x}(Y|Z{=}z) - E^{X=x'}(Y|Z{=}z) \tag{8.54}$$

is unbiased. Note that $PFE_{Z;xx'}(z)$ is the value of the effect function $PFE_{Z;xx'}\colon \Omega'_Z \to \mathbb{R}$. Also remember that $PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}$ denotes *unbiasedness of* $PFE_{Z;xx'}(z)$, which is defined by

$$PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}} \quad \Leftrightarrow \quad \big(PFE_{Z;xx'}(z) = CTE_{Z;xx'}(z) \wedge \tau_x, \tau_{x'} \text{ are } P^{Z=z}\text{-unique}\big) \tag{8.55}$$

[see Def. 6.23 (iii)]. Presuming $P^{Z=z}$-uniqueness of τ_x and $\tau_{x'}$,

$$CTE_{Z;xx'}(z) \;=\; E(\tau_x|Z{=}z) - E(\tau_{x'}|Z{=}z) \tag{8.56}$$

[see Eq. (5.31)] is the causal $(Z{=}z)$-conditional total effect and $CTE_{Z;xx'}(z)$ is the value of the causal Z-conditional total effect function $CTE_{Z;xx'}\colon \Omega'_Z \to \mathbb{R}$.

Theorem 8.50 [Unbiasedness of the $(Z{=}z)$-Conditional Prima Facie Effect]
Let the Assumptions 8.1 (a) to (f) hold, let $P(X{=}x \mid Z{=}z), P(X{=}x' \mid Z{=}z) > 0$, and let Z be a covariate of X. Then each of the following conditions implies $PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}$:

(i) $D_X \perp\!\!\!\perp 1_{X=x} \mid (Z{=}z) \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'} \mid (Z{=}z)$

(ii) $D_X \perp\!\!\!\perp 1_{X=x} \mid Z \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'} \mid Z$

(iii) $D_X \perp\!\!\!\perp X \mid Z$

(iv) $D_X \perp\!\!\!\perp X \mid (Z{=}z)$.

(Proof p. 264)

Remark 8.51 [Identification of $CTE_{Z;xx'}(z)$] Given the assumptions of Theorem 8.50, each of the four conditions listed in this theorem implies that the causal $(Z{=}z)$-conditional total effect is identical to the difference between the two conditional expectation values $E(Y \mid X{=}x, Z{=}z)$ and $E(Y \mid X{=}x', Z{=}z)$. That is, each of these four conditions implies

$$CTE_{Z;xx'}(z) \ = \ PFE_{Z;xx'}(z) \ = \ E(Y \mid X{=}x, Z{=}z) - E(Y \mid X{=}x', Z{=}z)\,. \tag{8.57}$$

In this context we also say that *$CTE_{Z;xx'}(z)$ is identified* by the $(Z{=}z)$-conditional prima facie effect $PFE_{Z;xx'}(z)$, that is, it is identified by the difference between the conditional expectation values $E(Y \mid X{=}x, Z{=}z)$ and $E(Y \mid X{=}x', Z{=}z)$. ◁

Remark 8.52 [Conditional Randomization] Note that $D_X \perp\!\!\!\perp X \mid Z$ [see Th. 8.50 (iii)] can be created by conditional randomization, that is, by randomized assignment of the unit to a treatment x, *conditional on the values z of a covariate Z*. In this case, the treatment probabilities $P(X{=}x \mid Z{=}z)$ can be fixed by the experimenter and may differ between different values z of the covariate Z (see Rem. 8.15). ◁

Remark 8.53 [Covariate Selection] Also note that we may try to select the (possibly multivariate) covariate $Z = (Z_1, \dots, Z_m)$ in such a way that $D_X \perp\!\!\!\perp X \mid Z$ holds. For instance, *severity of the disorder*, *knowing about the treatment*, and *availability of the treatment* often are candidates for such covariates. However, there is no guarantee that $D_X \perp\!\!\!\perp X \mid Z$ holds for a specified (univariate or multivariate) covariate Z. This is why the conditions specified in the following theorem are of utmost importance. ◁

Theorem 8.54 [Again Unbiasedness of a $(Z{=}z)$-Conditional Prima Facie Effect]
If the Assumptions 8.1 (a) to (f) hold and Z is a covariate of X, then each of the following conditions implies $PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}$:

(i) $D_X \perp\!\!\!\perp 1_{X=x} \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'}$

(ii) $D_X \perp\!\!\!\perp X$.

(Proof p. 265)

Remark 8.55 [Randomization] Note that $D_X \perp\!\!\!\perp X$ [see Th. 8.54 (ii)] can be created in an experiment by randomization, that is, by randomized assignment of the sampled unit to a treatment condition (represented by a value x of X). Also note that Z can be any covariate of X and remember, according to Theorem 4.31, Definition 4.11 (iv), and Remark 4.16, every random variable on $(\Omega, \mathscr{A}, P)$ that is prior in $(\mathscr{F}_t)_{t \in T}$ to X is a covariate of X. ◁

8.5 Examples

Now we illustrate the causality conditions treated in this chapter by some examples. In RS-chapter 1 we introduced an example with independence of X and a global potential confounder D_X (see RS-Table 1.2). A second example has already been presented in chapter 6 (see Table 6.3), and in the same chapter, there is also an example with Z-conditional independence of D_X and X (see Table 6.4). The structure of these random experiments has already been treated in section 6.5, that is,

$$\big((\Omega,\mathscr{A},P),(\mathscr{F}_t)_{t\in T},\mathscr{C},\mathscr{D}_{\mathscr{C}},X,Y\big),$$

as specified in section 6.5, is the regular probabilistic causality setup. Furthermore, the observational-unit variable U is a global potential confounder, that is, $D_X = U$, implying

$$E(Y|X,D_X) \underset{P}{=} E(Y|X,U)$$

and

$$P(X{=}1|D_X) \underset{P}{=} P(X{=}1|U).$$

In this equation, $P(X{=}1|U)$ denotes the individual treatment probability function, whose values are the individual treatment probabilities $P(X{=}1|U{=}u)$ (see RS-Remarks 4.12 and 4.25). If the values u of U represent the observational units *at the onset of treatment*, then $D_X = U$ will hold in empirical applications if

(a) no fallible covariate is assessed, and
(b) there is no other variable that is simultaneous to the treatment variable X (such as a second treatment variable).

If a fallible covariate is assessed, then this covariate, say Z, is not measurable with respect to U, which implies that it is not identical to a composition $f(U)$ of U and some map f. In this case, $D_X = (U,Z)$ is a global potential confounder of X, unless there are still other fallible covariates of X.

Example 8.56 [Independence of X and U] Table 6.3 displays an example in which $D_X \perp\!\!\!\perp X$ holds, where $D_X = U$. In this table, it is easily seen that the individual treatment probabilities are the same for all units, that is,

$$P(X{=}1|U) \;=\; P(X{=}1) \;=\; \frac{3}{4}.$$

Note that $U \perp\!\!\!\perp X$ implies independence of X and *all* random variables on $(\Omega,\mathscr{A},P)$ that are measurable with respect to U [see RS-Box 2.1 (iv)].

In the same example, $D_X \perp\!\!\!\perp X|Z$ holds as well, where $Z = sex$, because $D_X = U$ and Z is measurable with respect to U. Furthermore,

$$P(X{=}1|U,Z) \underset{P}{=} P(X{=}1|Z) \underset{P}{=} P(X{=}1) \;=\; \frac{3}{4}.$$

Note that $U \perp\!\!\!\perp X\,|\,Z$ implies that X and all random variables that are measurable with respect to U are also Z-conditionally independent [see RS-Box 6.1 (vi)].

If $D_X{=}U$, then, according to Theorem 8.22 (v), independence of U and X implies unbiasedness of the conditional expectation $E(Y|X)$. Furthermore, because, in this example,

X is binary and $P(X{=}0) > 0$, $P(X{=}1) > 0$, independence of X and U also implies unbiasedness of the conditional expectation values $E(Y|X{=}0)$ and $E(Y|X{=}1)$ [see Th. 8.22 (iv)] as well as unbiasedness of the prima facie effect

$$PFE_{10} := E(Y|X{=}1) - E(Y|X{=}0) = 102.333 - 92.333 = 10$$

(see Th. 8.43), and this implies $PFE_{10} = ATE_{10}$.

As already stated above, $U \perp\!\!\!\perp X \,|\, Z$ holds as well for Z = *sex*. Because $D_X{=}U$ and Z is U-measurable, according to Theorem 8.27 (iv), the conditional expectation $E(Y|X,Z)$ is unbiased. Furthermore, because $P(X{=}0, Z{=}z)$, $P(X{=}1, Z{=}z) > 0$ for both values z of Z, this implies unbiasedness of all conditional expectation values $E^{Z=z}(Y|X{=}x) = E(Y|X{=}x, Z{=}z)$ [see Th. 8.29 (iv)], and that the $(Z{=}z)$-conditional prima facie effects

$$PFE_{Z;10}(m) = E(Y|X{=}1, Z{=}m) - E(Y|X{=}0, Z{=}m) \approx 92.50 - 83.00 \approx 9.50$$

and

$$PFE_{Z;10}(f) = E(Y|X{=}1, Z{=}f) - E(Y|X{=}0, Z{=}f) = 122 - 111 = 11$$

are unbiased (see Th. 8.50).

Also note that

$$PFE_{10} = E[PFE_{Z;10}(Z)] = 9.50 \cdot \frac{4}{6} + 11 \cdot \frac{2}{6} = 10.$$

This means that the prima facie effect is the expectation of the corresponding conditional prima facie effect variable. And because $PFE_{Z;10}(Z)$ is unbiased, it is identical to the causal Z-conditional total effect variable $CTE_{Z;10}(Z)$, and its expectation is identical to the causal average total effect ATE_{10} (see Th. 6.34). $\triangleleft$

Example 8.57 [Z-Conditional Independence of X and U] In Table 6.4 we already treated an example in which $D_X \perp\!\!\!\perp X | Z$ holds, whereas $D_X \perp\!\!\!\perp X$ does not. Again, $D_X{=}U$. As is easily seen in this table, the individual treatment probabilities are the same for all units *within each of the two subsets of males and females*, that is,

$$\text{for each male unit } u: \quad P(X{=}1|U{=}u, Z{=}m) = P(X{=}1|Z{=}m) = \frac{3}{4}$$

and

$$\text{for each female unit } u: \quad P(X{=}1|U{=}u, Z{=}f) = P(X{=}1|Z{=}f) = \frac{1}{4}.$$

Hence, the individual treatment probabilities are 3/4 for males and 1/4 for females. These individual treatment probabilities differ for different values of the covariate Z, but they are invariant *given a value of the covariate* Z. Note that the conditional treatment probability $P(X{=}1|Z{=}m) = 3/4$ is also the individual treatment probability $P(X{=}1|U{=}u)$ for the four male units, and the conditional treatment probability $P(X{=}1|Z{=}f) = 1/4$ is also the individual treatment probability $P(X{=}1|U{=}u)$ for the two female units. This follows from the fact that $P(X{=}1|U,Z)$ and $P(X{=}1|Z)$ are conditional expectations [see RS-Eq. (4.10)] and

$$\begin{aligned}
P(X{=}1|U) &\underset{P}{=} E\big(P(X{=}1|U,Z)\,\big|\,U\big) && [\sigma(U,Z) \subset \sigma(U),\ \text{RS-Box 4.1 (xiii)}] \\
&\underset{P}{=} E\big(P(X{=}1|Z)\,\big|\,U\big) && [P(X{=}1|U,Z) \underset{P}{=} P(X{=}1|Z)] \\
&\underset{P}{=} P(X{=}1|Z). && \big[\sigma\big(P(X{=}1|Z)\big) \subset \sigma(U),\ \text{RS-Box 4.1 (xi)}\big]
\end{aligned}$$

In Table 6.4 there is only Z*-conditional* independence of U and the treatment variable X, that is, $U \perp\!\!\!\perp X \,|\, Z$, whereas $U \perp\!\!\!\perp X$ does not hold. Therefore, in this table, the prima facie effect

$$PFE_{10} \;=\; E(Y|X=1) - E(Y|X=0) \;\approx\; 96.715 - 99.800 \;\approx\; -3.085$$

is biased, because the average total effect in this example is $ATE_{10} = 10$ (see Exercise 8-9). However, the $(Z=z)$*-conditional* prima facie effects are unbiased. In fact, they are the same as in Table 6.3 (see Example 8.56). Hence, we can use the conditional prima facie effects to compute the average total effect. If the conditional prima facie effects are unbiased, that is, if they are equal to the corresponding causal conditional total effects, then the expectation of the conditional prima facie effect variable is equal to the causal average total effect (see Th. 6.34). In our example, this expectation is,

$$\begin{aligned} ATE_{10} \;=\; E\big(PFE_{Z;10}(Z)\big) \;&=\; PFE_{Z;10}(m)\cdot\frac{4}{6} + PFE_{Z;10}(f)\cdot\frac{2}{6} \\ &=\; 9.50\cdot\frac{4}{6} + 11\cdot\frac{2}{6} \;=\; 10.00. \end{aligned}$$

$\triangleleft$

8.6 Methodological Consequences

Now we discuss the conclusions from the theory treated in this chapter for the design and analysis of experiments and quasi-experiments. Theorems 8.22 and 8.27 are the theoretical foundation of the experimental design technique of *randomization* and of the analysis of causal conditional and causal average total treatment effects in experiments by comparing (unadjusted) means between treatment conditions. Theorem 8.31 is the theoretical foundation for the experimental design technique of *conditional randomization* and of the analysis of causal conditional total treatment effects. Together with Theorem 6.34, this theorem can also be used for the analysis of causal average total effects. Finally, Theorems 8.31 and 8.34, and Corollary 8.36 can be used in (nonrandomized) quasi-experiments for covariate selection and the analysis of causal conditional and causal average total effects. This will now be explained in more detail.

Remark 8.58 [Randomization] It is well-known at least since (Fisher, 1925/1946) that *randomization* plays a crucial role in ruling out bias in comparisons of means. In a randomized experiment, there is always an observational-unit variable U whose value u denotes the observational unit that is sampled if the random experiment considered is actually conducted. The definition of a global potential confounder D_X [see Def. 4.11 (iii)] guarantees that U is measurable with respect to a global potential confounder D_X. In fact, in a simple experiment, in which we do not observe any fallible pretests and in which there is only one single treatment variable X, the random variable U itself is a global potential confounder of X. Such a global potential confounder of X comprises all potential confounders of X in the sense that they are measurable with respect to the global potential confounder. In any case, if X represents a discrete treatment variable, then $D_X \perp\!\!\!\perp X$ is equivalent to

$$P(X=x|D_X) \underset{P}{=} P(X=x|U) \underset{P}{=} P(X=x), \quad \forall\, x \in X(\Omega)\,. \tag{8.58}$$

The last one of these two equations implies that all units $u \in U(\Omega)$ have the same probability $P(X{=}x \mid U{=}u) = P(X{=}x)$ to be assigned to treatment x, and this holds for all $x \in X(\Omega)$.

Remember that we are talking about a *single-unit trial*. A simple example of such a single-unit trial consists of sampling a single observational unit from a set of units, possibly assessing a number of covariates, assigning the unit (or observing its assignment) to one of the treatment conditions and observing the outcome variable (see ch. 2 for more details and other kinds of single-unit trials).

Note that Equation (8.58) does *not* imply that the probabilities $P(X{=}x)$ are the same for all treatment conditions x. If, for example, there are two treatment conditions, say 0 and 1, then the two treatment probabilities might be $P(X{=}1) = 1/4$ and $P(X{=}0) = 3/4$. However, $D_X \perp\!\!\!\perp X$ implies $P(X{=}1 \mid U) = P(X{=}1)$ because $\sigma(U) \subset \sigma(D_X)$. In other words, $D_X \perp\!\!\!\perp X$ implies that the individual treatment probabilities $P(X{=}x \mid U{=}u)$ are identical for all units, and they are equal to the (unconditional) treatment probability $P(X{=}x)$. Hence, in a perfect randomized experiment we ensure $D_X \perp\!\!\!\perp X$. If $P(X{=}x) > 0$ for all $x \in X(\Omega)$, then this condition implies that the conditional expectation values $E(Y \mid X{=}x)$, $x \in X(\Omega)$, are unbiased (see Th. 8.22) and that a prima facie effect $E(Y \mid X{=}x) - E(Y \mid X{=}x')$ is identical to the average causal total effect $ATE_{xx'}$ (see Th. 8.43).

It is important to note that through a perfect randomized experiment we do not only create $D_X \perp\!\!\!\perp X$ but also $D_X \perp\!\!\!\perp X \mid W$ for each potential confounder W of X (see Th. 8.17). If $P(X{=}x) > 0$ for all $x \in X(\Omega)$, then this condition implies $P(X{=}x \mid W) \underset{P}{>} 0$ for all $x \in X(\Omega)$, which in turn implies that each conditional expectation $E^{X=x}(Y \mid W)$, $x \in X(\Omega)$, the conditional expectation $E(Y \mid X,W)$, and each prima facie effect variable

$$PFE_{W;xx'}(W) \;=\; E^{X=x}(Y \mid W) - E^{X=x'}(Y \mid W), \quad x, x' \in X(\Omega),$$

are unbiased (see Cor. 8.36 and Th. 8.46). ◁

Remark 8.59 [Conditional Randomization] In the single-unit trial of an experiment or quasi-experiment, in which a unit u is sampled and a value z of a covariate Z is assessed prior to treatment, *conditional randomization* refers to randomized assignment of the unit u to the treatment condition x with probability $P(X{=}x \mid Z{=}z)$, $x \in X(\Omega)$. In this procedure we have to make sure that treatment assignment only depends on the values z of Z but not on any other attribute of the units or any other covariate. In more formal terms, we create $P(X{=}x \mid D_X) \underset{P}{=} P(X{=}x \mid Z)$ for all values $x \in X(\Omega)$, where $X(\Omega)$ denotes the image of Ω under X, that is, the set of all values of the treatment variable X.

We distinguish two cases. *First*, if Z is assessed without error, then each value z of Z represents an attribute of the observational units. In this case, there is a map f such that $Z = f(U)$ is the composition of U and f. This implies that U is D_X-measurable (see RS-Cor. 2.36). Table 6.4 contains an example with conditional independence of U and treatment variable X given that the sampled person is *male* or is *female*, the two values of Z. *Second*, if the covariate Z is assessed *with* error, that is, $Z{=}f(U){+}\varepsilon$ for some non-constant random variable ε, then a value z of Z does *not* represent an attribute of the units because z is not only determined by u but also by other factors and/or chance. Nevertheless, treatment can be assigned such that the treatment probability deterministically depends on the *observed score* z, and not on the attribute $f(u)$ of the unit u itself. In both cases, $Z{=}f(U)$ and $Z{=}f(U) + \varepsilon$, conditional randomized assignment of a unit u to one of the treatment conditions given a value z of Z secures that a global potential confounder D_X satisfies $D_X \perp\!\!\!\perp X \mid Z$, that is, Z-conditional independence of D_X and X. In the first case,

$D_X = U$ and in the second case $D_X = (U, Z)$, provided there are no other covariates of X than Z that are assessed with error. (Covariates that are assessed without error are maps of U and therefore U-measurable.)

Conditionally randomized assignment allows us that units with different values z_1 and z_2 of Z have different probabilities to be assigned to treatment x. Thus it is possible to assign a unit for which we observe a covariate value z_1 (e. g., high need) to a treatment x with a higher probability than a unit with covariate value z_2 (e. g., low need). In this way we may respect ethical and/or other requirements without compromising the validity of causal inference. Remember, unconditionally randomized assignment means that each unit has the *same* probability of being assigned to a given treatment, regardless of his or her need and any other attribute of the unit.

For simplicity, suppose there are just two treatment conditions, $X{=}0$ (control) and $X{=}1$ (treatment). Then conditional randomization consists of:

(a) fixing the conditional treatment probabilities $P(X{=}1 \mid Z{=}z)$ for all values z of the covariate Z of X,
(b) sampling a unit u and assessing the value z of the covariate Z, and
(c) randomly assigning the unit with probability $P(X{=}1 \mid Z{=}z)$ to treatment 1.

If, for example, the covariate has three values, say *high*, *medium*, and *low need*, we may toss a die and assign a unit with *high need* to treatment if the die shows less than six dots, and assign it to control otherwise. A unit with *medium need* might be assigned to treatment if the die shows less than four dots, and to control otherwise. Finally, a unit with *low need* might be assigned to treatment if the die shows one dot, and to control otherwise.

If the covariate Z has a finite number of values, then the conditional treatment probabilities $P(X{=}1 \mid Z{=}z)$ may be fixed in a table by assigning a treatment probability to each value z that seems appropriate with respect to ethical and other requirements. In the example above, these were the values 5/6 for *high need*, 3/6 for *medium need* and 1/6 for *low need*. If the covariate is univariate continuous, then we may also use a function, such as

$$P(X{=}1 \mid Z) \;=\; \frac{\exp(\lambda_0 + \lambda_1 \cdot Z)}{1 + \exp(\lambda_0 + \lambda_1 \cdot Z)}$$

with real-valued coefficients λ_0 and λ_1 that seem appropriate for the experiment considered.

If $P(X{=}x, Z{=}z), P(X{=}x', Z{=}z) > 0$, then Z-conditional randomization implies that the $(Z{=}z)$-conditional prima facie effect $PFE_{Z;xx'}(z) = E(Y \mid X{=}x, Z{=}z) - E(Y \mid X{=}x', Z{=}z)$ is identical to the causal conditional total effect given a value z of the covariate Z (see Th. 8.50 for the mathematical details). Hence, studying conditional prima facie effects in a perfect conditionally randomized experiment is tantamount to studying causal conditional total effects. And once we know the causal conditional total effects given the values z of a covariate Z, then we can also compute the causal average total effect (see Th. 6.34). $\lhd$

Remark 8.60 [Covariate Selection in Quasi-Experiments] Covariate selection is the first step in a number of techniques for the analysis of causal conditional and causal average total effects in *quasi-experiments*, in which by definition of the quasi-experiment, the experimenter cannot fix the true treatment probabilities himself. The steps to follow distinguish different techniques of analysis. By definition, in a quasi-experiment, there is no randomization and no conditional randomization. However, even under initial randomization, systematic attrition of subjects may invalidate the F-condition $D_X \perp\!\!\!\perp X$ (see, e. g., Abraham & Russell, 2004; Fichman & Cummings, 2003; Graham & Donaldson, 1993; Shadish et al., 2002). In this case, we will say that randomization failed and the initially randomized experiment turned into a quasi-experiment.

In a quasi-experiment, selecting the covariates in the covariate vector $Z := (Z_1, \ldots, Z_m)$ for which we can hope that $D_X \perp\!\!\!\perp X | Z$ holds is a useful strategy in the analysis of causal conditional and causal average total treatment effects. However, note that there might be many covariates determining the treatment probabilities. For instance, the *severity of the disorder*, *knowing about the treatment*, and *availability of the treatment* are candidates for such covariates. To emphasize, there is no guarantee that $D_X \perp\!\!\!\perp X | Z$ holds for a specified (univariate or multivariate) covariate Z, unless the conditional probabilities $P(X{=}x | Z)$, $x \in X(\Omega)$, are fixed by the experimenter. ◁

Remark 8.61 [Testability] As already mentioned before, in contrast to the causality conditions treated in chapters 6 and 7, the causality conditions $D_X \perp\!\!\!\perp X$ and $D_X \perp\!\!\!\perp X | Z$ can be tested in empirical applications, at least in the sense that some consequences of these conditions can be checked. If at least one of these consequences does not hold, then we say that the corresponding causality condition is *falsified*.

Let us briefly outline how we can test the assumption that $D_X \perp\!\!\!\perp X$ holds. Remember, if $X(\Omega)$ is finite or countable, then $D_X \perp\!\!\!\perp X$ is equivalent to

$$P(X{=}x | D_X) \underset{P}{=} P(X{=}x), \quad \forall\, x \in X(\Omega)\,. \tag{8.59}$$

This implies that

$$P(X{=}x | W) \underset{P}{=} P(X{=}x), \quad \forall\, x \in X(\Omega) \tag{8.60}$$

holds for each random variable W that is measurable with respect to D_X (see again Exercise 8-8). Examples for such random variables W are *sex*, *educational status*, but also fallible pretests, that is, variables that are assessed with measurement error before the onset of treatment.

Similarly, if $X(\Omega)$ is finite or countable, then $D_X \perp\!\!\!\perp X | Z$ is equivalent to

$$P(X{=}x | D_X) \underset{P}{=} P(X{=}x | Z), \quad \forall\, x \in X(\Omega)\,. \tag{8.61}$$

(Remember that Z denotes a covariate of X, which, by definition, is D_X-measurable.) This equation implies

$$P(X{=}x | W, Z) \underset{P}{=} P(X{=}x | Z), \quad \forall\, x \in X(\Omega), \tag{8.62}$$

for each potential confounder of X, that is, for each random variable W that is measurable with respect to D_X (see Exercise 8-13). Equations (8.60) and (8.62) can be tested using standard procedures for the analysis of conditional probabilities such as logistic regression (see, e. g., Agresti, 2007; Bonney, 1987; Green, 2003; Hosmer & Lemeshow, 2000) or probit regression (see, e. g., McCullagh & Nelder, 1989; Borooah, 2001; Liao, 1994). ◁

Remark 8.62 [The Covariate Selection Process] If Equation (8.62) does not hold for specified D_X-measurable random variables W and Z, then $D_X \perp\!\!\!\perp X | Z$, which implies Equation (8.62), cannot hold as well. In this case we may define $Z^* := (Z, W)$ and select a new D_X-measurable random variable W^*, and check if

$$P(X{=}x | W^*, Z^*) \underset{P}{=} P(X{=}x | Z^*), \quad \forall\, x \in X(\Omega), \tag{8.63}$$

holds. This process can be continued as long as there is doubt that

$$P(X{=}x\,|\,D_X) \underset{P}{=} P(X{=}x\,|\,Z^*), \quad \forall\, x \in X(\Omega)\,. \tag{8.64}$$

Of course, such a procedure does not guarantee that we find a (possibly multivariate) covariate Z^* such that Equation (8.64) holds. Instead, in a quasi-experiment, Equation (8.64) always remains an assumption. However, this assumption can always be falsified, which has a positive and a negative side. The negative side is that we can never be sure that this assumption holds. The positive side is that this assumption is empirically testable and, in this sense, it is not just a matter of belief (cf. Popper, 2005). ◁

Remark 8.63 [Generalizability of $D_X {\perp\!\!\!\perp} X$] If $D_X {\perp\!\!\!\perp} X$ holds, then, according to Theorem 8.17, $D_X {\perp\!\!\!\perp} X\,|\,W$ holds as well, provided that W is a potential confounder of X, or synonymously, a covariate of X. Under the assumptions of Theorem 8.22, this implies: If $D_X {\perp\!\!\!\perp} X$ holds, then $E(Y\,|\,X)$ is unbiased, and under the assumptions of Theorem 8.27, which include that Z is a covariate of X, it also implies unbiasedness of $E(Y\,|\,X,Z)$. Conditioning on a covariate Z of X may be meaningful in order to obtain a causal conditional total effect function $CTE_{Z;xx'}$ that is more fine-grained and therefore providing more specific information than the causal average total effect $ATE_{xx'}$.

If, for example, Z denotes the covariate *sex* with values m (*male*) and f (*female*), and $P(X{=}x, Z{=}z)$, $P(X{=}x', Z{=}z) > 0$ for both values z of Z, then $D_X {\perp\!\!\!\perp} X$ does not only imply $PFE_{xx'} = ATE_{xx'}$ but also

$$PFE_{Z;xx'}(z) \;=\; E(Y\,|\,X{=}x, Z{=}z) \;-\; E(Y\,|\,X{=}x', Z{=}z) \;=\; CTE_{Z;xx'}(z),$$

where $CTE_{Z;xx'}(z)$ denotes the causal $(Z{=}z)$-conditional total effect comparing x to x'. This applies to males $(z{=}m)$ and to females $(z{=}f)$. Even if $PFE_{xx'} = PFE_{Z;xx'}(m) = PFE_{Z;xx'}(f)$, then this would be an important information, extending the substantive interpretation of the causal average total effect $ATE_{xx'}$ because then it would be identical to the causal conditional total effects $CTE_{Z;xx'}(m)$ and $CTE_{Z;xx'}(f)$.

In order to emphasize the importance of this point for substantive research, imagine there would be an empirical study in which we can interpret $E(Y\,|\,X{=}x) - E(Y\,|\,X{=}x')$ as the causal average total treatment effect and at the same time the corresponding differences $E(Y\,|\,X{=}x, Z{=}z) - E(Y\,|\,X{=}x', Z{=}z)$ would have no causal interpretation! ◁

Remark 8.64 [Generalizability of $D_X {\perp\!\!\!\perp} X\,|\,Z$] Now assume that $D_X {\perp\!\!\!\perp} X\,|\,Z$ holds, where Z is a random variable on $(\Omega, \mathscr{A}, P)$, but not necessarily a covariate of X. If W is a potential confounder of X, then, according to Theorem 8.18, this implies that we also have $D_X {\perp\!\!\!\perp} X\,|\,(Z, W)$, that is, (Z, W)-conditional independence of D_X and X. Hence, if we condition on Z and $D_X {\perp\!\!\!\perp} X\,|\,Z$ holds, then there is no need to control for further potential confounders, at least not for the purpose of establishing unbiasedness. However, controlling for W may still be meaningful in order to obtain a more fine-grained total effect function $CTE_{Z,W;xx'}$ that contains more specific information than $CTE_{Z;xx'}$.

Under the Assumptions 8.1 (a) to (d), (g), $P(X{=}x\,|\,Z, W) \underset{P}{>} 0$ for all $x \in X(\Omega)$, and that Z is a covariate of X, the condition $D_X {\perp\!\!\!\perp} X\,|\,Z$ implies unbiasedness not only of $E(Y\,|\,X, Z)$ but, via $D_X {\perp\!\!\!\perp} X\,|\,(Z, W)$, also of $E(Y\,|\,X, Z, W)$ (see Cor. 8.36), no matter which other potential confounder W of X we consider. Controlling, additionally to Z, for another potential confounder W of X may still be meaningful in order to obtain a causal conditional total effect function $CTE_{Z,W;xx'}$ that is more fine-grained than $CTE_{Z;xx'}$. ◁

8.7 Summary and Conclusions

In chapter 6 we showed that unbiasedness of the conditional or unconditional prima facie effects is crucial for computing causal conditional and unconditional total effects from (the empircally estimable) conditional or unconditional prima facie effects. In chapter 7 we treated several causality conditions that imply unbiasedness, which cannot be tested in empirical applications. The most restrictive of these conditions is Rosenbaum and Rubin's strong ignorability. In the present chapter, we introduced a first kind of causality conditions that are empirically testable. These causality conditions, which are called Fisher conditions, are summarized in Box 8.1.

The implication relations among the causality conditions treated in this chapter are listed in Table 8.1 and the implication on the Rosenbaum-Rubin conditions are summarized in Table 8.2. This table also includes $\tau \perp\!\!\!\perp X$ and the strong ignorability condition $\tau \perp\!\!\!\perp X \,|\, Z$ that have been treated in chapter 7. The consequences of these two conditions on other causality conditions are summarized in the last rows of Table 7.1 and Table 7.2, respectively. Hence, putting these tables together yields long chains of implications. Obviously, $D_X \perp\!\!\!\perp X$ is the most powerful causality condition. If $P(X{=}x) > 0$ for all $x \in X(\Omega)$ and Z is a covariate of X, then it does not only imply unbiasedness of all $E(Y \,|\, X{=}x)$, $E^{X=x}(Y \,|\, Z)$, $x \in X(\Omega)$, as well as unbiasedness of $E(Y \,|\, X)$ and $E(Y \,|\, X, Z)$, but it also implies all other conditions treated until now (including those dealt with in ch. 7). It even implies that all true outcomes variables are P-unique. In contrast, assuming $D_X \perp\!\!\!\perp X \,|\, Z$ but not $D_X \perp\!\!\!\perp X$, we need the additional assumptions that Z is a covariate of X and $P(X{=}x \,|\, Z) \underset{P}{>} 0$ in order to show that all $E^{X=x}(Y \,|\, Z)$, $x \in X(\Omega)$, as well as $E(Y \,|\, X, Z)$ are unbiased.

In experiments, the condition $D_X \perp\!\!\!\perp X$ can be created by randomized assignment of the observational unit to one of the treatment conditions represented by a value x of X. Similarly, $D_X \perp\!\!\!\perp X \,|\, Z$ can be created by conditionally randomized assignment based on the values z of covariate Z, provided of course, that Z is a pretreatment variable.

Outside the randomized experiment, and even in cases in which X does not denote a treatment variable that is manipulated by an experimenter, $D_X \perp\!\!\!\perp X \,|\, Z$ may also be valid if the covariate Z is selected appropriately (see Rem. 8.60 to 8.62).

Unlike unbiasedness and the other causality conditions treated in chapter 7, independence of X and a global potential confounder D_X of X as well as Z-conditional independence of X and D_X are empirically testable. Suppose that X is discrete. In order to test $D_X \perp\!\!\!\perp X$, we simply have to select a potential confounder W of X and investigate if $P(X{=}x \,|\, W) \underset{P}{=} P(X{=}x)$ holds for all values x of X. Similarly, if X is discrete and Z is a covariate of X, then we can test $D_X \perp\!\!\!\perp X \,|\, Z$ by selecting a potential confounder W of X and see if $P(X{=}x \,|\, Z, W) \underset{P}{=} P(X{=}x \,|\, Z)$ holds for all values x of X.

8.8 Proofs

Proof of Theorem 8.17

Proposition (8.29). By definition of a potential confounder of X, $\sigma(W) \subset \sigma(D_X)$, and this implies $\sigma(W, D_X) = \sigma(D_X)$ [see RS-Prop. (2.19)]. Therefore,

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon (W, D_X) \perp\!\!\!\perp X && [\sigma(W, D_X) = \sigma(D_X)] \\
&\Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp X \,|\, W. && [\text{RS-Box 6.1 (ix)}]
\end{aligned}
$$

Proposition (8.30). Correspondingly,

$$\begin{aligned} D_X \perp\!\!\!\perp 1_{X=x} \quad &\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon (W, D_X) \perp\!\!\!\perp 1_{X=x} && [\sigma(W, D_X) = \sigma(D_X)] \\ &\Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp 1_{X=x} | W. && [\text{RS-Box 6.1 (ix)}] \end{aligned}$$

Proof of Theorem 8.18

Proposition (8.32). By definition of a potential confounder of X, $\sigma(W) \subset \sigma(D_X)$, and this implies $\sigma(W, D_X) = \sigma(D_X)$ [see RS-Prop. (2.19)]. Therefore,

$$\begin{aligned} D_X \perp\!\!\!\perp X | Z \quad &\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon (W, D_X) \perp\!\!\!\perp X \,|\, Z && [\sigma(W, D_X) = \sigma(D_X)] \\ &\Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp X \,|\, (Z, W)\,. && [\text{RS-Box 6.1 (viii)}] \end{aligned}$$

Proposition (8.33). Analogously,

$$\begin{aligned} D_X \perp\!\!\!\perp 1_{X=x} | Z \quad &\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon (W, D_X) \perp\!\!\!\perp 1_{X=x} | Z && [\sigma(W, D_X) = \sigma(D_X)] \\ &\Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon D_X \perp\!\!\!\perp 1_{X=x} | (Z, W)\,. && [\text{RS-Box 6.1 (viii)}] \end{aligned}$$

Proof of Theorem 8.22

Proposition (i). This follows from $\sigma(\tau) \subset \sigma(D_X)$ and RS-Box 2.1 (iv).
Proposition (ii).

$$\begin{aligned} D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X && [\text{(i)}] \\ &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x}\,. && [(7.23)] \end{aligned}$$

Proposition (iii).

$$\begin{aligned} D_X \perp\!\!\!\perp X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x) && [(8.5)] \\ &\Rightarrow \quad \forall\, x \in X(\Omega)\colon P(X{=}x \,|\, D_X) \underset{P}{>} 0 && [P(X{=}x) > 0, \text{SN-}(2.40)] \\ &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \text{ is } P\text{-unique.} && [\text{RS-Th. 5.27}] \end{aligned}$$

Propositions (iv) and (v).

$$\begin{aligned} D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X \wedge \big(\forall\, x \in X(\Omega)\colon \tau_x \text{ is } P\text{-unique}\big) && [\text{(i), (iii)}] \\ &\Rightarrow \quad \forall\, x \in X(\Omega)\colon E(Y \,|\, X{=}x) \vdash D_X && [\text{Th. 7.10}] \\ &\Leftrightarrow \quad E(Y \,|\, X) \vdash D_X\,. && [(7.28)] \end{aligned}$$

Proof of Theorem 8.25

Note that $0 < P(X{=}x) < 1$ implies $0 < P(X{\neq}x) < 1$ because $P(X{\neq}x) = 1 - P(X{=}x)$.

Proposition (i). Because $\sigma\big(E^{X=x}(Y|D_X)\big) \subset \sigma(D_X)$, RS-Box 2.1 (iv) implies

$$D_X \perp\!\!\!\perp 1_{X=x} \quad \Rightarrow \quad E^{X=x}(Y|D_X) \perp\!\!\!\perp 1_{X=x}\,.$$

Correspondingly, because $\sigma\big(E^{X\neq x}(Y|D_X)\big) \subset \sigma(D_X)$ and $\sigma(1_{X=x}) = \sigma(1_{X\neq x})$, RS-Box 2.1 (iv) yields

$$D_X \perp\!\!\!\perp 1_{X\neq x} \quad \Rightarrow \quad E^{X\neq x}(Y|D_X) \perp\!\!\!\perp 1_{X\neq x}.$$

Proposition (ii).

$$\begin{aligned}
D_X \perp\!\!\!\perp 1_{X=x} \quad &\Leftrightarrow \quad P(X{=}x|D_X) \underset{P}{=} P(X{=}x) && [(8.1)]\\
&\Rightarrow \quad P(X{=}x|D_X) \underset{P}{>} 0 && [P(X{=}x) > 0,\ \text{SN-}(2.40)]\\
&\Leftrightarrow \quad E^{X=x}(Y|D_X)\ \text{is } P\text{-unique.} && [\text{RS-Th. } 5.27]
\end{aligned}$$

Correspondingly,

$$\begin{aligned}
D_X \perp\!\!\!\perp 1_{X=x} \quad &\Leftrightarrow \quad P(X{\neq}x|D_X) \underset{P}{=} P(X{\neq}x) && [(8.4)]\\
&\Rightarrow \quad P(X{\neq}x|D_X) \underset{P}{>} 0 && [P(X{\neq}x) > 0,\ \text{SN-}(2.40)]\\
&\Leftrightarrow \quad E^{X\neq x}(Y|D_X)\ \text{is } P\text{-unique.} && [\text{SN-Cor. } 14.48]
\end{aligned}$$

Proposition (iii).

$$\begin{aligned}
&D_X \perp\!\!\!\perp 1_{X=x}\\
\Rightarrow\quad &E^{X=x}(Y|D_X) \perp\!\!\!\perp 1_{X=x} \wedge E^{X=x}(Y|D_X)\ \text{is } P\text{-unique} && [(i), (iii)]\\
\Rightarrow\quad &E(Y|X{=}x) \vdash D_X. && [E^{X=x}(Y|D_X) = \tau_x,\ \text{Th. } 7.8]
\end{aligned}$$

$$\begin{aligned}
&D_X \perp\!\!\!\perp 1_{X=x}\\
\Rightarrow\quad &E^{X\neq x}(Y|D_X) \perp\!\!\!\perp 1_{X\neq x} \wedge E^{X\neq x}(Y|D_X)\ \text{is } P\text{-unique} && [(i), (iii)]\\
\Rightarrow\quad &E^{X\neq x}(Y|D_X) \vDash 1_{X\neq x} \wedge E^{X\neq x}(Y|D_X)\ \text{is } P\text{-unique} && [\text{RS-Th. } 4.40]\\
\Rightarrow\quad &E\big(E^{X\neq x}(Y|D_X)\,\big|\,1_{X\neq x}\big) \underset{P}{=} E\big(E^{X\neq x}(Y|D_X)\big) \wedge E^{X\neq x}(Y|D_X)\ \text{is } P\text{-unique}\\
& && [\text{RS-Def. } 4.36]\\
\Leftrightarrow\quad &E(Y|X{\neq}x) \vdash D_X. && [\text{Th. } 6.9\ (i),\ (6.13)]
\end{aligned}$$

Proposition (iv). This immediately follows from Propositions (iii) and (6.13).

Proof of Theorem 8.27

Proposition (i).

$$\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad D_X \perp\!\!\!\perp X|Z && [\text{Th. } 8.17]\\
&\Rightarrow \quad \tau \perp\!\!\!\perp X|Z. && [\sigma(\tau) \subset \sigma(D_X),\ \text{RS-Box } 6.1\ (vi)]
\end{aligned}$$

Proposition (ii).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, Z && \text{[(i)]}\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} \,|\, Z\,. && \text{[RS-(6.8)]}
\end{aligned}
$$

Propositions (iii) and (iv).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, Z \wedge (\forall\, x \in X(\Omega)\colon \tau_x \text{ is } P\text{-unique}) && \text{[(i), Th. 8.22 (iii)]}\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y \,|\, Z) \vdash D_X && \text{[Th. 7.21]}\\
&\Leftrightarrow \quad E(Y \,|\, X, Z) \vdash D_X\,. && \text{[(7.63)]}
\end{aligned}
$$

Proof of Theorem 8.29

Proposition (i).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X\\
\Rightarrow \quad & Z \perp\!\!\!\perp X && [\sigma(Z) \subset \sigma(D_X), \text{ RS-Box 2.1 (iv)}]\\
\Rightarrow \quad & \forall\, x \in X(\Omega)\colon P(X{=}x, Z{=}z) = P(X{=}x) \cdot P(Z{=}z)\\
& && [\{X{=}x\} \in \sigma(X), \{Z{=}z\} \in \sigma(Z), \text{ Box 8.1 (ii)}]\\
\Rightarrow \quad & \forall\, x \in X(\Omega)\colon P(X{=}x, Z{=}z) > 0. && [P(Z{=}z) > 0, \forall\, x \in X(\Omega)\colon P(X{=}x) > 0]
\end{aligned}
$$

Proposition (ii).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, Z && \text{[Th. 8.27 (i)]}\\
&\Rightarrow \quad \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} X && [P(Z{=}z) > 0, \text{ RS-(6.37)}]\\
&\Leftrightarrow \quad \tau \perp\!\!\!\perp X \,|\, (Z{=}z)\,. && \text{[RS-(6.47)]}
\end{aligned}
$$

Proposition (iii).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, (Z{=}z) && \text{[(ii)]}\\
&\Leftrightarrow \quad \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} X && \text{[RS-(6.47)]}\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && \text{[RS-(6.27)]}\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)\,. && \text{[RS-(6.47)]}
\end{aligned}
$$

Propositions (iv) and (v).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, (Z{=}z) \wedge \big(\forall\, x \in X(\Omega)\colon \tau_x \text{ is } P\text{-unique}\big) && \text{[(ii), Th. 8.22 (iii)]}\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{Z=z}(Y \,|\, X{=}x) \vdash D_X && \text{[Cor. 7.14]}\\
&\Leftrightarrow \quad E^{Z=z}(Y \,|\, X) \vdash D_X\,. && \text{[(7.36)]}
\end{aligned}
$$

Proof of Theorem 8.31

Propositions (i) to (iii).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X \,|\, Z \\
\Rightarrow \quad & \tau \perp\!\!\!\perp X \,|\, Z && [\sigma(\tau) \subset \sigma(D_X),\ \text{RS-Box } 6.1\ (vi)] \\
\Leftrightarrow \quad & \forall\, x \in X(\Omega):\ \tau \perp\!\!\!\perp 1_{X=x} \,|\, Z && [\text{RS-Th. } 6.5] \\
\Rightarrow \quad & \forall\, x, x' \in X(\Omega):\ \tau_x \perp\!\!\!\perp 1_{X=x'} \,|\, Z\,. && [\sigma(\tau_x) \subset \sigma(\tau),\ \text{RS-Box } 6.1\ (vi)]
\end{aligned}
$$

Propositions (iv) and (v). These propositions follow from Proposition (i) and Theorem 7.21.

Proof of Theorem 8.34

Proposition (8.40).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \ \wedge\ P(X{=}x \,|\, Z) \underset{P}{>} 0 \\
\Leftrightarrow \quad & P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x \,|\, Z) \ \wedge\ P(X{=}x \,|\, Z) \underset{P}{>} 0 && [(8.7), (8.13)] \\
\Rightarrow \quad & P(X{=}x \,|\, D_X) \underset{P}{>} 0 && [\text{SN-}(2.40)] \\
\Leftrightarrow \quad & \tau_x \text{ is } P\text{-unique.} && [\text{RS-Th. } 5.27]
\end{aligned}
$$

Proposition (8.41).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X \,|\, Z \ \wedge\ \forall\, x \in X(\Omega):\ P(X{=}x \,|\, Z) \underset{P}{>} 0 \\
\Leftrightarrow \quad & \forall\, x \in X(\Omega):\ \big(D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \ \wedge\ P(X{=}x \,|\, Z) \underset{P}{>} 0\big) && [(8.12)] \\
\Rightarrow \quad & \forall\, x \in X(\Omega):\ \tau_x \text{ is } P\text{-unique.} && [(8.40)]
\end{aligned}
$$

Proof of Theorem 8.37

Proposition (i).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \,|\, Z \quad \Rightarrow \quad & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} X && [P(Z{=}z) > 0,\ \text{RS-}(6.37)] \\
\Rightarrow \quad & \forall\, x \in X(\Omega):\ \tau_x \text{ is } P^{Z=z}\text{-unique.} && [\text{Th. } 8.22\ (iii)]
\end{aligned}
$$

Proposition (ii).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \,|\, Z \quad \Rightarrow \quad & \tau \perp\!\!\!\perp X \,|\, Z && [\text{Th. } 8.31] \\
\Rightarrow \quad & \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} X && [P(Z{=}z) > 0,\ \text{RS-}(6.37)] \\
\Leftrightarrow \quad & \tau \perp\!\!\!\perp X \,|\, (Z{=}z)\,. && [\text{RS-}(6.47)]
\end{aligned}
$$

Proposition (iii).

$$
D_X \perp\!\!\!\perp X \,|\, Z \quad \Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, (Z{=}z) \qquad [(ii)]
$$

$$\begin{aligned}
&\Leftrightarrow \quad \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} X && [\text{RS-}(6.47)]\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && [(7.12), (7.14)]\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z=z)\,.
\end{aligned}$$

Proposition (iv). We assume that Z is a covariate of X, that is, $\sigma(Z) \subset \sigma(D_X)$.

$$\begin{aligned}
&\quad D_X \perp\!\!\!\perp X \,|\, Z\\
&\Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, (Z=z) \ \wedge \ \big(\forall\, x \in X(\Omega)\colon\ \tau_x \text{ is } P^{Z=z}\text{-unique}\big) && [\text{(i), (ii)}]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon\ \big(\tau_x \perp\!\!\!\perp X \,|\, (Z=z) \ \wedge \ \tau_x \text{ is } P^{Z=z}\text{-unique}\big) && [(7.32)]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{Z=z}(Y \,|\, X=x) \vdash D_X\,. && [\sigma(Z) \subset \sigma(D_X), (7.35)]
\end{aligned}$$

Proposition (v). Again, we assume $\sigma(Z) \subset \sigma(D_X)$. Hence,

$$\begin{aligned}
D_X \perp\!\!\!\perp X \,|\, Z \quad &\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{Z=z}(Y \,|\, X=x) \vdash D_X && [\text{(iv)}]\\
&\Leftrightarrow \quad E^{Z=z}(Y \,|\, X) \vdash D_X\,. && [\text{Def. 6.18 (ii)}]
\end{aligned}$$

Proof of Theorem 8.39

Proposition (i).

$$\begin{aligned}
&\quad D_X \perp\!\!\!\perp X \,|\, Z \ \wedge \ \big(\forall\, (x,z) \in X(\Omega) \times Z(\Omega)\colon P(X=x, Z=z) > 0\big)\\
&\Leftrightarrow \quad D_X \perp\!\!\!\perp X \,|\, Z \ \wedge \ \big(\forall\, (x,z) \in X(\Omega) \times Z(\Omega)\colon P^{Z=z}(X=x) > 0\big)\\
& && [\text{def. of } P^{Z=z}(X=x)]\\
&\Rightarrow \quad D_X \perp\!\!\!\perp X \,|\, Z \ \wedge \ \big(\forall\, x \in X(\Omega)\colon P(X=x \,|\, Z) \underset{P}{>} 0\big) && [\forall\, z \in Z(\Omega)\colon P(Z=z) > 0]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon \tau_x \text{ is } P\text{-unique.} && [\text{Th. 8.34}]
\end{aligned}$$

Proposition (ii). We additionally assume that Z is a covariate of X. Hence,

$$\begin{aligned}
D_X \perp\!\!\!\perp X \,|\, Z \quad &\Rightarrow \quad \forall\, (x,z) \in X(\Omega) \times Z(\Omega)\colon E^{Z=z}(Y \,|\, X=x) \vdash D_X && [\text{Th. 8.37 (iv)}]\\
&\Leftrightarrow \quad \forall\, (x,z) \in X(\Omega) \times Z(\Omega)\colon E^{X=x}(Y \,|\, Z=z) \vdash D_X && [(6.19)]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y \,|\, Z) \vdash D_X\,. && [\text{(i), Th. 6.17}]
\end{aligned}$$

Proposition (iii). This follows from Proposition (ii) and Definition 6.18 (i).

Proof of Theorem 8.41

Proposition (i).

$$\begin{aligned}
D_X \perp\!\!\!\perp X \,|\, (Z=z) \quad &\Leftrightarrow \quad D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} X && [(8.24)]\\
&\Rightarrow \quad \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} X && [\sigma(\tau) \subset \sigma(D_X), \text{RS-Box 2.1 (iv)}]\\
&\Leftrightarrow \quad \tau \perp\!\!\!\perp X \,|\, (Z=z)\,. && [\text{RS-}(6.47)]
\end{aligned}$$

Proposition (ii).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X \,|\, (Z{=}z) \\
\Rightarrow \quad & \tau \perp\!\!\!\perp X \,|\, (Z{=}z) && \text{[(i)]} \\
\Leftrightarrow \quad & \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} X && \text{[RS-(6.47)]} \\
\Rightarrow \quad & \forall\, x \in X(\Omega)\colon \tau \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && [\forall\, x \in X(\Omega)\colon \sigma(1_{X=x}) \subset \sigma(X), \text{ RS-Box 2.1 (iv)}] \\
\Leftrightarrow \quad & \forall\, x \in X(\Omega)\colon \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z). && \text{[RS-(6.47)]}
\end{aligned}
$$

Proposition (iii).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X \,|\, (Z{=}z) \\
\Leftrightarrow \quad & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} X && \text{[(8.24)]} \\
\Leftrightarrow \quad & \forall\, x \in X(\Omega)\colon P^{Z=z}(X{=}x \,|\, D_X) \underset{P^{Z=z}}{=} P^{Z=z}(X{=}x) && \text{[(8.25)]} \\
\Rightarrow \quad & \forall\, x \in X(\Omega)\colon P^{Z=z}(X{=}x \,|\, D_X) \underset{P^{Z=z}}{>} 0 && [P^{Z=z}(X{=}x) > 0, \text{ SN-(2.40)}] \\
\Leftrightarrow \quad & \forall\, x \in X(\Omega)\colon \tau_x \text{ is } P^{Z=z}\text{-unique.} && \text{[RS-Th. 5.27 (i), (ii)]}
\end{aligned}
$$

Propositions (iv) and (v). Now we also assume that Z is a covariate of X. Hence,

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \,|\, (Z{=}z) \quad \Rightarrow \quad & \tau \perp\!\!\!\perp X \,|\, (Z{=}z) && \text{[(i)]} \\
\Rightarrow \quad & \forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \,|\, (Z{=}z) && \text{[(7.32)]} \\
\Rightarrow \quad & \forall\, x \in X(\Omega)\colon E^{Z=z}(Y \,|\, X{=}x) \vdash D_X && \text{[(iii), (7.35)]} \\
\Rightarrow \quad & E^{Z=z}(Y \,|\, X) \vdash D_X. && \text{[(iii), (7.36)]}
\end{aligned}
$$

Proof of Theorem 8.42

The proof is analog to the proof of Theorem 8.25. We only have to replace the probability measure P on $(\Omega, \mathscr{A}, P)$ by the measure $P^{Z=z}$ defined in Assumption 8.1 (e). Also note that

$$
E^{X=x}(Y \,|\, D_X) \underset{P}{=} E^{1_{X=x}=1}(Y \,|\, D_X) \qquad \text{and} \qquad E^{X\neq x}(Y \,|\, D_X) \underset{P}{=} E^{1_{X=x}=0}(Y \,|\, D_X).
$$

Hence, $E^{X=x}(Y \,|\, D_X)$ and $E^{X\neq x}(Y \,|\, D_X)$ are versions of the two true outcome variables pertaining to the indicator variable $1_{X=x}$ (see Def. 5.4).

Proposition (i).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z) \\
\Leftrightarrow \quad & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && \text{[(8.19)]} \\
\Rightarrow \quad & E^{X=x}(Y \,|\, D_X) \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && [\sigma\big(E^{X=x}(Y \,|\, D_X)\big) \subset \sigma(D_X), \text{ RS-Box 2.1 (iv)}] \\
\Leftrightarrow \quad & E^{X=x}(Y \,|\, D_X) \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z). && \text{[RS-Prop. (6.47)]}
\end{aligned}
$$

The proof of $E^{X\neq x}(Y|D_X) \perp\!\!\!\perp 1_{X\neq x}|(Z=z)$ is analog. Just replace $E^{X=x}(Y|D_X)$ by $E^{X\neq x}(Y|D_X)$ and $1_{X=x}$ by $1_{X\neq x}$.

Proposition (ii).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x}|(Z=z) \\
\Leftrightarrow\quad & P^{Z=z}(X=x|D_X) \underset{P^{Z=z}}{=} P^{Z=z}(X=x) && [(8.20)] \\
\Rightarrow\quad & P^{Z=z}(X=x|D_X) \underset{P^{Z=z}}{>} 0 && [P^{Z=z}(X=x) > 0, \text{SN-}(2.40)] \\
\Leftrightarrow\quad & E^{X=x}(Y|D_X) \text{ is } P^{Z=z}\text{-unique.} && [\text{RS-Th. } 5.27]
\end{aligned}
$$

Analogously,

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x}|(Z=z) \\
\Leftrightarrow\quad & P^{Z=z}(X\neq x|D_X) \underset{P^{Z=z}}{=} P^{Z=z}(X\neq x) && [(8.23)] \\
\Rightarrow\quad & P^{Z=z}(X\neq x|D_X) \underset{P^{Z=z}}{>} 0 && [P^{Z=z}(X\neq x) > 0, \text{SN-}(2.40)] \\
\Leftrightarrow\quad & E^{X\neq x}(Y|D_X) \text{ is } P^{Z=z}\text{-unique.} && [\text{SN-Cor. } 14.48]
\end{aligned}
$$

Proposition (iii). Now we additionally assume that Z is a covariate of X.

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x}|(Z=z) \\
\Leftrightarrow\quad & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && [(8.19)] \\
\Rightarrow\quad & E^{X=x}(Y|D_X) \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} \wedge E^{X=x}(Y|D_X) \text{ is } P^{Z=z}\text{-unique} && [\text{(i), (iii)}] \\
\Rightarrow\quad & E^{Z=z}(Y|X=x) \vdash D_X. && [E^{X=x}(Y|D_X) = \tau_x, \text{Th. } 7.8, (7.20)]
\end{aligned}
$$

The proof of $E^{Z=z}(Y|X\neq x) \vdash D_X$ is almost analog.

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x}|(Z=z) \\
\Leftrightarrow\quad & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && [(8.19)] \\
\Leftrightarrow\quad & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X\neq x} && [\sigma(1_{X=x}) = \sigma(1_{X\neq x}), \text{RS-Def. } 2.59] \\
\Rightarrow\quad & E^{X\neq x}(Y|D_X) \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X\neq x} \wedge E^{X\neq x}(Y|D_X) \text{ is } P^{Z=z}\text{-unique} && [\text{(i), (iii)}] \\
\Rightarrow\quad & E^{Z=z}(Y|X\neq x) \vdash D_X. && [\text{Th. } 7.8, (7.20)]
\end{aligned}
$$

Proposition (iv). This immediately follows from Propositions (iii) and (6.13).

Proof of Theorem 8.43

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \wedge D_X \perp\!\!\!\perp 1_{X=x'} \\
\Rightarrow\quad & E(Y|X=x) \vdash D_X \wedge E(Y|X=x') \vdash D_X && [\text{Th. } 8.25\ \text{(iii)}]
\end{aligned}
$$

$$\Leftrightarrow \quad E(Y|X{=}x) = E(\tau_x) \;\wedge\; E(Y|X{=}x') = E(\tau_{x'}) \;\wedge\; \tau_x, \tau_{x'} \text{ are } P\text{-unique}$$
[Def. 6.3 (i)]

$$\Rightarrow \quad E(Y|X{=}x) - E(Y|X{=}x') = E(\tau_x) - E(\tau_{x'}) \;\wedge\; \tau_x, \tau_{x'} \text{ are } P\text{-unique}$$

$$\Leftrightarrow \quad PFE_{xx'} \vdash \mathscr{D}_{\mathscr{C}}\,.$$
[Def. 6.23 (i)]

Proof of Theorem 8.46

Proposition (i).

$$D_X \perp\!\!\!\perp 1_{X=x} | Z \;\wedge\; D_X \perp\!\!\!\perp 1_{X=x'} | Z$$

$$\Rightarrow \quad E^{1_{X=x}=1}(Y\,|\,Z) \vdash \mathscr{D}_{\mathscr{C}} \;\wedge\; E^{1_{X=x'}=1}(Y\,|\,Z) \vdash \mathscr{D}_{\mathscr{C}}$$
[Th. 8.31 (iv), (8.40)]

$$\Leftrightarrow \quad E^{X=x}(Y|Z) \vdash D_X \;\wedge\; E^{X=x'}(Y|Z) \vdash D_X$$
$[P(1_{X=x} = 1) = P(X{=}x),\, P(1_{X=x'} = 1) = P(X{=}x')]$

$$\Rightarrow \quad PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}\,.$$
[Th. 6.25]

Proposition (ii). This follows from Propositions (i) and (8.12).

Proof of Theorem 8.49

Proposition (i).

$$D_X \perp\!\!\!\perp 1_{X=x} \;\wedge\; D_X \perp\!\!\!\perp 1_{X=x'}$$

$$\Rightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} | Z \;\wedge\; D_X \perp\!\!\!\perp 1_{X=x'} | Z \;\wedge\; \tau_x, \tau_{x'} \text{ are } P\text{-unique}$$
[Th. 8.17 (ii), Th. 8.25 (ii)]

$$\Rightarrow \quad E^{1_{X=x}=1}(Y\,|\,Z) \vdash \mathscr{D}_{\mathscr{C}} \;\wedge\; E^{1_{X=x'}=1}(Y\,|\,Z) \vdash \mathscr{D}_{\mathscr{C}}$$
[Th. 8.31 (iv)]

$$\Leftrightarrow \quad E^{X=x}(Y|Z) \vdash D_X \;\wedge\; E^{X=x'}(Y|Z) \vdash D_X$$
$[P(1_{X=x} = 1) = P(X{=}x),\, P(1_{X=x'} = 1) = P(X{=}x')]$

$$\Rightarrow \quad PFE_{Z;xx'} \vdash \mathscr{D}_{\mathscr{C}}\,.$$
[Th. 6.25]

Proposition (ii). This follows from Propositions (i) and (8.6).

Proof of Theorem 8.50

Proposition (i).

$$D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z) \;\wedge\; D_X \perp\!\!\!\perp 1_{X=x'} | (Z{=}z)$$

$$\Rightarrow \quad E^{Z=z}(Y|X{=}x) \vdash D_X \;\wedge\; E^{Z=z}(Y|X{=}x') \vdash D_X$$
[Th. 8.42 (iii)]

$$\Leftrightarrow \quad E^{X=x}(Y|Z{=}z) \vdash D_X \;\wedge\; E^{X=x'}(Y|Z{=}z) \vdash D_X$$
[(6.19)]

$$\Rightarrow \quad PFE_{Z;xx'}(z) \vdash \mathscr{D}_{\mathscr{C}}\,.$$
[Th. 6.26]

Proposition (ii).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} | Z \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'} | Z \\
\Rightarrow\ & D_X \perp\!\!\!\perp 1_{X=x} | (Z=z) \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'} | (Z=z) && \text{[RS-(6.37), RS-(6.47)]} \\
\Rightarrow\ & PFE_{Z;xx'}(z) \vdash \mathcal{D}_{\mathscr{C}}. && \text{[(i)]}
\end{aligned}
$$

Proposition (iii).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X | Z \\
\Rightarrow\ & D_X \perp\!\!\!\perp 1_{X=x} | Z \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'} | Z && [\sigma(1_{X=x}), \sigma(1_{X=x'}) \subset \sigma(X), \text{RS-Box 6.1 (vi)}] \\
\Rightarrow\ & PFE_{Z;xx'}(z) \vdash \mathcal{D}_{\mathscr{C}}. && \text{[(ii)]}
\end{aligned}
$$

Proposition (iv).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X | (Z=z) \\
\Leftrightarrow\ & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} X \\
\Rightarrow\ & D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} \ \wedge\ D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x'} && [\sigma(1_{X=x}), \sigma(1_{X=x'}) \subset \sigma(X),\ \text{RS-Box 2.1 (iv)}] \\
\Leftrightarrow\ & D_X \perp\!\!\!\perp 1_{X=x} | (Z=z) \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'} | (Z=z) && \text{[RS-(6.47)]} \\
\Rightarrow\ & PFE_{Z;xx'}(z) \vdash \mathcal{D}_{\mathscr{C}}. && \text{[(i)]}
\end{aligned}
$$

Proof of Theorem 8.54

Proposition (i).

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \ \wedge\ D_X \perp\!\!\!\perp 1_{X=x'} \\
\Rightarrow\ & E^{Z=z}(Y | X=x) \vdash D_X \ \wedge\ E^{Z=z}(Y | X=x') \vdash D_X && \text{[Th. 8.29 (iv)]} \\
\Rightarrow\ & PFE_{Z;xx'}(z) \vdash \mathcal{D}_{\mathscr{C}}. && \text{[Th. 6.26]}
\end{aligned}
$$

Proposition (ii). This follows from Propositions (i) and (8.6).

8.9 Exercises

▷ **Exercise 8-1** Prove: If D_X is a global potential confounder and W a potential confounder of X, then $D_X \perp\!\!\!\perp X$ implies $W \perp\!\!\!\perp X$.

▷ **Exercise 8-2** Check if and where the implications listed in Table 8.1 have been proven in this chapter and prove those that have not. Use and specify the appropriate choice of assumptions listed in the Assumptions 8.1.

▷ **Exercise 8-3** Check if and where the implications listed in Table 8.2 have been proven in this chapter and prove those that have not. Use and specify the appropriate choice of assumptions listed in the Assumptions 8.1.

▷ **Exercise 8-4** Consider Table 8.3 and show that $P(X=1|U)(\omega) = P(X=1|Z)(\omega)$, for all $\omega \in \{Z=m\} \cup \{Z=f\}$ implies $P(X=1|U) \underset{P}{=} P(X=1|Z)$.

▷ **Exercise 8-5** Consider Table 8.3 and specify the four values of a second version of the U-conditional expectation of Y with respect to the measure $P^{X=1}$.

▷ **Exercise 8-6** Let the Assumptions 8.1 (a) to (c) and (g) hold and let Z be a covariate of X. Which terms are then unbiased if $D_X \perp\!\!\!\perp X$ holds?

▷ **Exercise 8-7** Let the Assumptions 8.1 (a) to (d) and (g) hold. Which terms are then unbiased if we also assume $D_X \perp\!\!\!\perp X | Z$ and $P(X{=}x | Z) \underset{P}{>} 0$?

▷ **Exercise 8-8** Show that $P(X{=}x | D_X) \underset{P}{=} P(X{=}x)$ implies

$$P(X{=}x | W) \underset{P}{=} P(X{=}x),$$

provided that the random variable W is measurable with respect to D_X. Assume that $P(X{=}x) > 0$.

▷ **Exercise 8-9** Compute the conditional expectation value $E(Y | X{=}0)$ and the expectation $E(\tau_0)$ in the example presented in Table 6.4 and compare these numbers to each other.

▷ **Exercise 8-10** Describe randomized assignment of a unit to one of two treatment conditions!

▷ **Exercise 8-11** Describe *conditionally* randomized assignment of a unit to one of two treatment conditions given a covariate Z! For simplicity, assume that Z is finite with $P(Z{=}z) > 0$ for all its values $z \in Z(\Omega)$ and that X is dichotomous with values 0 and 1.

▷ **Exercise 8-12** Show that $E[P(X{=}x | U)] = P(X{=}x)$.

▷ **Exercise 8-13** Show that $P(X{=}x | D_X) \underset{P}{=} P(X{=}x | Z)$, implies

$$P(X{=}x | Z, W) \underset{P}{=} P(X{=}x | Z),$$

if the random variable W is measurable with respect to D_X. Assume that $P(X{=}x) > 0$.

Solutions

▷ **Solution 8-1** Definitions 4.11 (iv) and (iii) of a potential confounder and a global potential confounder of X imply $\sigma(W) \subset \sigma(D_X)$. Therefore, RS-Proposition (6.5) yields the proposition.

▷ **Solution 8-2** We check the implications listed in Table 8.1 row-wise.

(1) Under the Assumptions 8.1, (a), (b), (d), and (e): $D_X \perp\!\!\!\perp X | (Z{=}z) \Rightarrow D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z)$.

$$\begin{aligned} D_X \perp\!\!\!\perp X | (Z{=}z) \quad &\Leftrightarrow \quad D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} X && [(8.24)] \\ &\Rightarrow \quad D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && [(8.6)] \\ &\Leftrightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z). && [(8.19)] \end{aligned}$$

(2) Under the Assumptions 8.1 (a), (b), (d), and (e): $D_X \perp\!\!\!\perp 1_{X=x} | Z \Rightarrow D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z)$. This immediately follows from RS-Propositions (6.47) and (6.48).

(3) Under the Assumptions 8.1 (a), (b), (d), and (e): $D_X \perp\!\!\!\perp X | Z \Rightarrow D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z)$.

$$\begin{aligned} D_X \perp\!\!\!\perp X | Z \quad &\Rightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} | Z && [\sigma(1_{X=x}) \subset \sigma(X), \text{RS-Box 6.1 (vi)}] \\ &\Rightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z). && [(2)] \end{aligned}$$

(4) Under the Assumptions 8.1 (a), (d), and (e): $D_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$.
This is Proposition (8.35).

(5) Under the Assumptions 8.1 (a), (b), and (d): $D_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z$.
This immediately follows from $\sigma(1_{X=x}) \subset \sigma(X)$ and RS-Box 6.1 (vi).

(6) Under the Assumptions 8.1 (a), (b), (d), (e), and (h): $D_X \perp\!\!\!\perp 1_{X=x} \;\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.

$$\begin{aligned} D_X \perp\!\!\!\perp 1_{X=x} \;&\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z && [(8.30)] \\ &\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)\,. && [(2)] \end{aligned}$$

(7) Under the Assumptions 8.1 (a), (b), and (h): $D_X \perp\!\!\!\perp 1_{X=x} \;\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z$.
This is Proposition (8.30).

(8) Under the Assumptions 8.1 (a), (b), (d), (e), and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.

$$\begin{aligned} D_X \perp\!\!\!\perp X \;&\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} && [(8.27)] \\ &\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)\,. && [(6)] \end{aligned}$$

(9) Under the Assumptions 8.1 (a), (d), (e), and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; D_X \perp\!\!\!\perp X \,|\, (Z{=}z)$.
This is Proposition (8.36).

(10) Under the Assumptions 8.1 (a), (b), and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z$.
This is Proposition (8.31).

(11) Under the Assumptions 8.1 (a) and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; D_X \perp\!\!\!\perp X \,|\, Z$.
This is Proposition (8.29).

(12) Under the Assumptions 8.1 (a) and (b): $D_X \perp\!\!\!\perp X \;\Rightarrow\; D_X \perp\!\!\!\perp 1_{X=x}$.
This is Proposition (8.27).

▷ **Solution 8-3** We check the implications listed in Table 8.2 row-wise.

(1) Under the Assumptions 8.1 (a) to (e) and (g): $D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z) \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.
This is Theorem 8.41 (i) for $1_{X=x}$ taking the role of X.

(2) Under the Assumptions 8.1 (a) to (e) and (g): $D_X \perp\!\!\!\perp X \,|\, (Z{=}z) \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.
This immediately follows from Theorem 8.41 (ii).

(3) Under the Assumptions 8.1 (a) to (e) and (g): $D_X \perp\!\!\!\perp X \,|\, (Z{=}z) \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, (Z{=}z)$.
This is Theorem 8.41 (i).

(4) Under the Assumptions 8.1 (a) to (e) and (g): $D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.
This is Theorem 8.37 (ii) for $1_{X=x}$ taking the role of X.

(5) Under the Assumptions 8.1 (a) to (d) and (g) : $D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, Z$.
This is Theorem 8.31 (i) for $1_{X=x}$ taking the role of X.

(6) Under the Assumptions 8.1 (a) to (g): $D_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.
This immediately follows from Theorem 8.37 (iii).

(7) Under the Assumptions 8.1 (a) to (e) and (g): $D_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, (Z{=}z)$.
This is Theorem 8.37 (ii).

(8) Under the Assumptions 8.1 (a) to (d) and (g) : $D_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, Z$.
This is Theorem 8.31 (ii).

(9) Under the Assumptions 8.1 (a) to (d) and (g) : $D_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, Z$.
This is Theorem 8.31 (i).

(10) Under the Assumptions 8.1 (a) to (g) and (h): $D_X \perp\!\!\!\perp 1_{X=x} \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z)$.
This is Theorem 8.29 (ii) for $1_{X=x}$ taking the role of X.

(11) Under the Assumptions 8.1 (a) to (d), (g), and (h): $D_X \perp\!\!\!\perp 1_{X=x} \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} \,|\, Z$.
This is Theorem 8.27 (i) for $1_{X=x}$ taking the role of X.

(12) Under the Assumptions 8.1 (a), (c), and (g): $D_X \perp\!\!\!\perp \mathit{1}_{X=x} \;\Rightarrow\; \tau \perp\!\!\!\perp \mathit{1}_{X=x}$.
This is Theorem 8.22 (i) for $\mathit{1}_{X=x}$ taking the role of X.

(13) Under the Assumptions 8.1 (a) to (g) and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp \mathit{1}_{X=x} | (Z=z)$.
This immediately follows from Theorem 8.29 (iii).

(14) Under the Assumptions 8.1 (a) to (g) and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp X | (Z=z)$.
This is Theorem 8.29 (ii).

(15) Under the Assumptions 8.1 (a) to (d), (g), and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp \mathit{1}_{X=x} | Z$.
This immediately follows from Theorem 8.27 (ii).

(16) Under the Assumptions 8.1 (a) to (d), (g), and (h): $D_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp X | Z$.
This is Theorem 8.27 (i).

(17) Under the Assumptions 8.1 (a), (c), and (g) : $D_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp \mathit{1}_{X=x}$.
This immediately follows from Theorem 8.22 (ii).

(18) Under the Assumptions 8.1 (a), (c), and (g) : $D_X \perp\!\!\!\perp X \;\Rightarrow\; \tau \perp\!\!\!\perp X$.
This is Theorem 8.22 (i).

▷ **Solution 8-4** In Example 8.33 we specified the regular causality setup to be the same as in section 6.5. Hence, $\Omega = \Omega_U \times \Omega_X \times \mathbb{R}$ [see Eq. (6.74)]. Furthermore, according to Equation (8.39), the values of $P(X{=}1|U)$ and $P(X{=}1|Z)$ are identical for all elements of the set

$$\{Z{=}m\} \cup \{Z{=}f\} \;=\; \{Joe, Jim\} \times \Omega_X \times \mathbb{R} \;\cup\; \{Ann, Sue\} \times \Omega_X \times \mathbb{R} \;=\; \Omega.$$

Hence, the set A used in RS-Definition 2.46 is Ø, the empty set. Because $P(\text{Ø}) = 0$ and

$$\forall\, \omega \in \Omega \setminus \text{Ø}\colon\; P(X{=}1|U)(\omega) = P(X{=}1|Z)(\omega)$$

holds [see RS-Eq. (2.61)], this proves $P(X{=}1|U) \underset{P}{=} P(X{=}1|Z)$.

▷ **Solution 8-5** The four values of a second version τ_1^* of the U-conditional expectation of Y with respect to the measure $P^{X=1}$ are

$$\begin{aligned}
\tau_1^*(\omega) &= 1000, \quad \text{if } \omega \in \{U{=}Joe\}\\
\tau_1^*(\omega) &= 2000, \quad \text{if } \omega \in \{U{=}Jim\}\\
\tau_1^*(\omega) &= 114, \quad \text{if } \omega \in \{U{=}Ann\}\\
\tau_1^*(\omega) &= 130, \quad \text{if } \omega \in \{U{=}Sue\}.
\end{aligned}$$

Note that, instead of 1000 and 2000, we could have chosen any other two numbers. The crucial point is: $\tau_1^* \underset{P^{X=1}}{=} \tau_1$ because $P^{X=1}(U{=}Joe) = P^{X=1}(U{=}Jim) = 0$ (see Table 8.3).

▷ **Solution 8-6** Under these assumptions $D_X \perp\!\!\!\perp X$ implies that $E(Y|X)$, $E(Y|X,Z)$, all $E(Y|X{=}x)$, and all $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, are unbiased. This also implies that the prima facie effects $PFE_{xx'}$ and the prima facie effect variables $PFE_{Z;xx'}(Z)$, $x, x' \in X(\Omega)$, are unbiased as well.

▷ **Solution 8-7** Under these assumptions $D_X \perp\!\!\!\perp X | Z$ implies that $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, are unbiased. It also implies that the prima facie effect variables $PFE_{Z;xx'}(Z)$, $x, x' \in X(\Omega)$, are unbiased as well. Also note that the causal average total effects $ATE_{xx'}$ can be computed from $PFE_{Z;xx'}(Z)$ (see Th. 6.34).

▷ **Solution 8-8** If $P(X{=}x|D_X) \underset{P}{=} P(X{=}x)$, then

$$\begin{aligned}
P(X{=}x|W) &\underset{P}{=} E(\mathit{1}_{X=x}|W) && \text{[RS-(4.10)]}\\
&\underset{P}{=} E\big(E(\mathit{1}_{X=x}|D_X)\,\big|\,W\big) && \text{[RS-Box 4.1 (xiii)]}\\
&\underset{P}{=} E\big(P(X{=}x|D_X)\,\big|\,W\big) && \text{[RS-(4.10)]}\\
&\underset{P}{=} E\big(P(X{=}x)\,|\,W\big) && [P(X{=}x|D_X) \underset{P}{=} P(X{=}x)\text{, RS-Box 4.1 (xiv)}]\\
&\underset{P}{=} P(X{=}x). && \text{[RS-Box 4.1 (i)]}
\end{aligned}$$

▹ **Solution 8-9** According to RS-Box 3.2 (ii), the conditional expectation value $E(Y|X{=}0)$ can be computed via

$$\begin{aligned} E(Y|X{=}0) &= \sum_u E(Y|X{=}0, U{=}u) \cdot P(U{=}u|X{=}0) \\ &= (68+78+88+98) \cdot \frac{1}{10} + (106+116) \cdot \frac{3}{10} = 99.8. \end{aligned}$$

In contrast, according to RS-(3.13) and Equations (6.77), (6.78), the expectation $E(\tau_0)$ can be computed via

$$\begin{aligned} E(\tau_0) = E[E^{X=0}(Y|U)] = E\big(g_0(U)\big) &= \sum_u g_0(u) \cdot P(U{=}u) \\ &= \sum_u E(Y|X{=}0, U{=}u) \cdot P(U{=}u) \\ &= (68+78+88+98+106+116) \cdot \frac{1}{6} = 92.3333. \end{aligned}$$

Comparing $E(Y|X{=}0) = 99.8$ to $E(\tau_0) = 92.3333$ shows that $E(Y|X{=}0)$ is strongly biased [see Def. 6.3 (i)].

▹ **Solution 8-10** In a random experiment, in which a unit u is sampled and assigned to one of two treatment conditions, we may assign the unit by coin toss, for instance. This ensures

$$P(X{=}1|D_X) \underset{P}{=} P(X{=}1),$$

that is, the treatment probabilities do not depend on the global potential confounder D_X of X [see Prop. (8.5)] and therefore not on any potential confounder of X (see Rem. 8.61). If D_X is specified such that U is measurable with respect to D_X, then $P(X{=}1|D_X) \underset{P}{=} P(X{=}1)$ implies that each unit u has the same probability $P(X{=}1|U{=}u) = P(X{=}1)$ of being assigned to treatment 1. If we assume that X is dichotomous with values 0 and 1, then this also implies that each unit u has the same probability $P(X{=}0)$ to be assigned to treatment 0 as well because

$$P(X{=}0|U{=}u) = 1 - P(X{=}1|U{=}u) = 1 - P(X{=}1) = P(X{=}0).$$

▹ **Solution 8-11** We consider a random experiment, in which a unit u is sampled and a value z of the covariate Z is assessed before the unit is assigned to one of the two treatment conditions. We also assume $P(Z{=}z) > 0$ for all values $z \in Z(\Omega)$ and $0 < P(X{=}1|Z) < 1$. Then *Z-conditional randomized assignment* of a unit to one of the two treatment conditions refers to assigning the unit u to treatment 1 with probability

$$P^{Z=z}(X{=}1|D_X) \underset{P^{Z=z}}{=} P^{Z=z}(X{=}1) = P(X{=}1|Z{=}z).$$

This equation implies

$$\begin{aligned} P^{Z=z}(X{=}1|U) &\underset{P^{Z=z}}{=} E^{Z=z}\big(P^{Z=z}(X{=}1|D_X)\,\big|\,U\big) && \text{[RS-Box 4.1 (xiii)]} \\ &\underset{P^{Z=z}}{=} E^{Z=z}\big(P^{Z=z}(X{=}1)\,\big|\,U\big) \\ &\underset{P^{Z=z}}{=} P^{Z=z}(X{=}1) && \text{[RS-Box 4.1 (i)]} \\ &= P(X{=}1|Z{=}z) && \text{[RS-(3.24), RS-(3.26)]} \end{aligned}$$

and

$$P^{Z=z}(X{=}1|U)(\omega) = P^{Z=z}(X{=}1) = P(X{=}1|Z{=}z), \quad \text{if } \omega \in \{Z{=}z\}$$

[see RS-Eq. (4.18)]. Hence, all units with the same value z of the covariate Z have the same positive probability to be assigned to treatment 1. This assignment procedure allows for different treatment probabilities for units for which we observe different values z of the covariate Z. It creates $D_X {\perp\!\!\!\perp} X|Z$ and $0 < P(X{=}1|Z) < 1$, which implies that $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, are unbiased (see Cor. 8.36).

▷ **Solution 8-12**

$$\begin{aligned} E\big(P(X{=}x\,|\,U)\big) &= E\big(E(1_{X=x}\,|U)\big) && [\mathit{RS}\text{-}(4.10)]\\ &= E(1_{X=x}) && [\mathit{RS}\text{-}\mathit{Box}\ 4.1\ (\mathrm{iv})]\\ &= P(X{=}x). && [\mathit{RS}\text{-}(3.9)] \end{aligned}$$

▷ **Solution 8-13** If $P(X{=}x\,|\,D_X) \underset{P}{=} P(X{=}x\,|\,Z)$, then

$$\begin{aligned} P(X{=}x\,|\,Z,W) &\underset{P}{=} E(1_{X=x}|Z,W) && [\text{RS-}(4.10)]\\ &\underset{P}{=} E\big(E(1_{X=x}|D_X)\,\big|\,Z,W\big) && [\text{RS-Box } 4.1\ (\text{xiii})]\\ &\underset{P}{=} E\big(P(X{=}x\,|\,D_X)\,\big|\,Z,W\big) && [\text{RS-}(4.10)]\\ &\underset{P}{=} E\big(P(X{=}x\,|\,Z)\,|\,Z,W\big)\\ &\underset{P}{=} P(X{=}x\,|\,Z). && [\text{RS-Box } 4.1\ (\text{xi})] \end{aligned}$$

Chapter 9
Suppes-Reichenbach Conditions

In chapter 8, we introduced the Fisher conditions, a first class of empirically testable causality conditions. We emphasized that $D_X \perp\!\!\!\perp X$ implies $D_X \perp\!\!\!\perp X \,|\, Z$, provided that Z is a covariate of X. We also showed that $D_X \perp\!\!\!\perp X \,|\, Z$ implies the Rosenbaum-Rubin conditions, if we assume $P(X{=}x \,|\, Z) \underset{P}{>} 0$, for all $x \in X(\Omega)$. The Fisher conditions focus on (conditional) independence of the *putative cause variable* X and all potential confounders of X.

In contrast, the causality conditions introduced in the present chapter, focus on conditional independence or conditional mean-independence of the *outcome variable* Y and all potential confounders of X. These conditions will summarily be referred to as the *Suppes-Reichenbach conditions*, honoring two pioneers who made early contributions to probabilistic causality (see, e. g., Reichenbach, 1956; Suppes, 1970). These conditions are also empirically testable, some of them also apply if X is continuous, and if the true outcome variables are P-unique, then the Suppes-Reichenbach conditions have implications on the Rosenbaum-Rubin conditions.

We start introducing these causality conditions, then treat their consequences and discuss their methodological implications. In particular, we discuss their role in covariate selection in the empirical analysis of causal conditional and average total effects.

Requirements

Reading this chapter we assume that the reader is familiar with the concepts treated in all chapters of Steyer (2024). Again, chapters 4 to 6 of that book are now crucial, dealing with the concepts of a conditional expectation, a conditional expectation with respect to a conditional probability measure, and conditional independence. Furthermore, we assume familiarity with chapters 4 to 8 of the present book.

In this chapter we often refer to the following notation and assumptions.

Notation and Assumptions 9.1

(a) *Let* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *be a regular probabilistic causality setup, let* D_X *denote a global potential confounder of* X*, and* $\mathscr{W}_X$ *the set of all potential confounders of* X*, that is, the set of all random variables on* $(\Omega, \mathscr{A}, P)$ *satisfying* $\sigma(W) \subset \sigma(D_X)$.

(b) *Let* $X(\Omega) = \{0, 1, \dots, J\}$ *denote the image of* Ω *under* X*, let* $x \in X(\Omega)$*, let* $\{x\} \in \mathscr{A}'_X$*, and let* $1_{X=x}$ *be the indicator of the event* $\{X{=}x\} = \{\omega \in \Omega \colon X(\omega) = x\}$*. Furthermore, assume* $P(X{=}x) > 0$ *and define the probability measure* $P^{X=x} \colon \mathscr{A} \to [0, 1]$ *by* $P^{X=x}(A) = P(A \,|\, X{=}x)$*, for all* $A \in \mathscr{A}$.

(c) *Let* Y *be real-valued with positive variance.*

(d) *For* $x \in X(\Omega)$, *let* $\tau_x = E^{X=x}(Y \mid D_X)$ *denote a (version of the) true outcome variable of* Y *given (the value)* x *(of* X*).*
(e) *For all* $x \in X(\Omega) = \{0, 1, \dots, J\}$, *let* $\{x\} \in \mathscr{A}'_X$ *and assume* $0 < P(X{=}x) < 1$. *Furthermore, let* $\tau := (\tau_0, \tau_1, \dots, \tau_J)$ *denote the* $(J+1)$*-variate random variable consisting of the true outcome variables* τ_x, $x \in X(\Omega)$.
(f) *Let* $x' \in \Omega'_X$ *and* $\{x'\} \in \mathscr{A}'_X$, *let* $1_{X=x'}$ *denote the indicator of the event* $\{X{=}x'\} = \{\omega \in \Omega\colon X(\omega) = x'\}$ *and assume that* $0 < P(X{=}x') < 1$. *Furthermore, let* $\tau_{x'} = E^{X=x'}(Y \mid D_X)$ *denote a (version of the) true outcome variable of* Y *given* x'.
(g) *Let* Z *be a random variable on* $(\Omega, \mathscr{A}, P)$ *and let* $(\Omega'_Z, \mathscr{A}'_Z)$ *denote its value space.*
(h) *Let* Z *be a covariate of* X, *that is, let* Z *be a random variable on* $(\Omega, \mathscr{A}, P)$ *satisfying* $\sigma(Z) \subset \sigma(D_X)$.
(i) *Assume that* τ_x *is* P*-unique.*
(j) *Assume that all* τ_x, $x \in X(\Omega)$, *are* P*-unique.*
(k) *Assume that* $\tau_{x'}$ *is* P*-unique.*

9.1 SR-Conditions

In this section, we introduce eight causality conditions, summarily referred to as the Suppes-Reichenbach (SR) conditions. They include conditional independence and conditional mean-independence among certain random variables. Remember, conditional mean-independence of a numerical random variable from another random variable given a third random variable has be been treated in RS-section 4.7 and conditional independence of two random variables given a third one in RS-section 6.1.2. In this chapter, we use these concepts for the definitions and notation of all SR-conditions. Later, their most important implications are studied, including their implications on the Rosenbaum-Rubin conditions and on unbiasedness.

9.1.1 Simple SR-Conditions

Box 9.1 presents all eight Suppes-Reichenbach conditions discussed in this chapter. We start commenting on those conditions that only involve conditioning on X or on one of its values x, but not on another random variable Z.

Remark 9.2 [Independence of Y and D_X With Respect to $P^{X=x}$] The very first condition displayed in Box 9.1 (i), $Y \perp\!\!\!\perp D_X \mid (X{=}x)$, is equivalent to *independence of* Y *and* D_X *with respect to the measure* $P^{X=x}$. This measure has been specified in Assumptions 9.1 (b). Hence,

$$Y \perp\!\!\!\perp D_X \mid (X{=}x) \quad \Leftrightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X, \tag{9.1}$$

where

$$Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X \quad :\Leftrightarrow \quad \forall (A, B) \in \sigma(Y) \times \sigma(D_X)\colon\ P^{X=x}(A \cap B) = P^{X=x}(A) \cdot P^{X=x}(B). \tag{9.2}$$

This implies that we can utilize all properties of independence of two random variables with respect to a probability measure. In this case, it is the measure $P^{X=x}$ instead of P.

Box 9.1 Suppes-Reichenbach conditions

$Y \perp\!\!\!\perp D_X \,|\, (X{=}x)$ — $(X{=}x)$*-conditional independence of Y and* D_X. Under Assumptions 9.1 (a) and (b), it is defined by

$$\forall\,(A,B) \in \sigma(Y) \times \sigma(D_X):\; P(A \cap B \,|\, X{=}x) = P(A \,|\, X{=}x) \cdot P(B \,|\, X{=}x). \quad \text{(i)}$$

$Y \vDash D_X \,|\, (X{=}x)$ — $(X{=}x)$*-conditional mean-independence of Y from* D_X. Under Assumptions 9.1 (a) to (c), it is defined by

$$E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y). \quad \text{(ii)}$$

Under Assumptions 9.1 (a) to (d), each of the above two conditions implies $E(Y \,|\, X{=}x) \vdash D_X$. If, additionally, Z is a covariate of X, then each of them also implies $E^{X=x}(Y \,|\, Z) \vdash D_X$.

$Y \perp\!\!\!\perp D_X \,|\, X$ — X*-conditional independence of Y and* D_X. Under Assumptions 9.1 (a), it is defined by

$$\forall\,(A,B) \in \sigma(Y) \times \sigma(D_X):\; P(A \cap B \,|\, X) \underset{P}{=} P(A \,|\, X) \cdot P(B \,|\, X). \quad \text{(iii)}$$

$Y \vDash D_X \,|\, X$ — X*-conditional mean-independence of Y from* D_X. Under Assumptions 9.1 (a) and (c), it is defined by

$$E(Y \,|\, X, D_X) \underset{P}{=} E(Y \,|\, X)\,. \quad \text{(iv)}$$

Under Assumptions 9.1 (a) to (e), each of the last two conditions implies $E(Y \,|\, X) \vdash D_X$. If, additionally, Z is a covariate of X, then each of them also implies $E(Y \,|\, X, Z) \vdash D_X$.

$Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z)$ — $(X{=}x, Z)$*-conditional independence of Y and* D_X. Under Assumptions 9.1 (a), (b), and (g), it is defined by

$$\forall\,(A,B) \in \sigma(Y) \times \sigma(D_X):\; P^{X=x}(A \cap B \,|\, Z) \underset{P^{X=x}}{=} P^{X=x}(A \,|\, Z) \cdot P^{X=x}(B \,|\, Z). \quad \text{(v)}$$

$Y \vDash D_X \,|\, (X{=}x, Z)$ — $(X{=}x, Z)$*-conditional mean-independence of Y from* D_X. Under Assumptions 9.1 (a) to (c) and (g), it is defined by

$$E^{X=x}(Y \,|\, Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y \,|\, Z)\,. \quad \text{(vi)}$$

Under Assumptions 9.1 (a) to (d) and if Z is a covariate of X, then each of the last two conditions implies $E^{X=x}(Y \,|\, Z) \vdash D_X$.

$Y \perp\!\!\!\perp D_X \,|\, (X, Z)$ — (X, Z)*-conditional independence of Y and* D_X. Under Assumptions 9.1 (a) and (g), it is defined by

$$\forall\,(A,B) \in \sigma(Y) \times \sigma(D_X):\; P(A \cap B \,|\, X, Z) \underset{P}{=} P(A \,|\, X, Z) \cdot P(B \,|\, X, Z). \quad \text{(vii)}$$

$Y \vDash D_X \,|\, (X, Z)$ — (X, Z)*-conditional mean-independence of Y from* D_X. Under Assumptions 9.1 (a), (c), and (g), it is defined by

$$E(Y \,|\, X, Z, D_X) \underset{P}{=} E(Y \,|\, X, Z)\,. \quad \text{(viii)}$$

Under Assumptions 9.1 (a) to (e) and that Z is a covariate of X, each of the last two conditions implies $E(Y \,|\, X, Z) \vdash D_X$ and $E^{X=x}(Y \,|\, Z) \vdash D_X$ for all $x \in X(\Omega)$.

Note that the condition $Y \perp\!\!\!\perp D_X \,|\, (X{=}x)$ does not apply if X is continuous because we assume $P(X{=}x) > 0$ [see Assumptions 9.1 (b) and RS-Rem. 2.77]. Independence of random variables (with respect to a probability measure) and some of its properties have been treated in RS-section 2.4. (More details are provided, for example, in SN-chapter 16. In particular, equivalent conditions for $Y \perp\!\!\!\perp D_X \,|\, X$ in terms of conditional distributions are found in SN-section 17.6.) $\lhd$

Remark 9.3 [Mean-Independence of Y and D_X With Respect to $P^{X=x}$] The causality condition $Y \vDash D_X \,|\, (X{=}x)$ defined in Box 9.1 (ii) is equivalent to *mean-independence of Y from D_X with respect to $P^{X=x}$* (see RS-Def. 4.36), which is denoted $Y \underset{P^{X=x}}{\vDash} D_X$. Hence,

$$Y \vDash D_X \,|\, (X{=}x) \quad \Leftrightarrow \quad Y \underset{P^{X=x}}{\vDash} D_X, \tag{9.3}$$

where

$$Y \underset{P^{X=x}}{\vDash} D_X \quad :\Leftrightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y). \tag{9.4}$$

Again note that $Y \vDash D_X \,|\, (X{=}x)$ does not apply if X is continuous because it is defined only if $P(X{=}x) > 0$. $\lhd$

Remark 9.4 [X-Conditional Independence of Y and D_X] The third causality condition defined in Box 9.1 (iii) is *X-conditional independence of the outcome variable Y and D_X*, denoted by $Y \perp\!\!\!\perp D_X \,|\, X$. This condition implies that the *distribution* of the outcome variable Y does not depend on the global potential confounder D_X, once we condition on the putative cause variable X. That is, we implicitly postulate that $P_{Y|X}$ is a version of the conditional distribution $P_{Y|X,D_X}$ (see SN-ch. 17). This condition also applies if X is continuous. More details about conditional independence are found in RS-chapter 6 and about conditional distributions in SN-chapter 16. $\lhd$

Remark 9.5 [X-Conditional Mean-Independence of Y From D_X] The next causality condition, defined in Box 9.1 (iv), is *X-conditional mean-independence of Y and D_X*, denoted $Y \vDash D_X \,|\, X$. With this condition we postulate that the (X, D_X)-conditional expectation of Y actually does not depend on the global potential confounder D_X of X, once we condition on X. That is, we postulate that $E(Y \,|\, X)$ is a version of the conditional expectation $E(Y \,|\, X, D_X)$ (see RS-ch. 4). Again, note that $Y \vDash D_X \,|\, X$ also applies if X is continuous. $\lhd$

Remark 9.6 [A Caveat Concerning Mediators] Note that none of the causality conditions treated so far excludes that mediators (and other variables that are between X and Y) determine the distribution of Y. For example, $Y \perp\!\!\!\perp D_X \,|\, X$ only implies that the conditional distribution $P_{Y|X}$ is a version of the conditional distribution $P_{Y|X,W}$ whenever W is a potential confounder of X. Remember, a potential confounder of X is D_X-measurable and therefore prior or simultaneous in $(\mathscr{F}_t)_{t \in T}$ to X (see Cor. 4.33). $\lhd$

In the following theorem we present a condition that is equivalent to $Y \vDash D_X \,|\, X$, provided that, for all values x of X, $P(X{=}x) > 0$.

Theorem 9.7 [A Condition Equivalent to $Y \vDash D_X \,|\, X$ if X is Discrete]
Let the Assumptions 9.1 (a) to (c) hold and assume $P(X{=}x) > 0$ for all $x \in X(\Omega)$. Then

$$Y \vDash D_X \,|\, X \quad \Leftrightarrow \quad \forall\, x \in X(\Omega):\; E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y). \tag{9.5}$$

(Proof p. 292)

Remark 9.8 [Constant True Outcome Variables] Let the Assumptions 9.1 (a) to (e) hold, which includes that all τ_x, $x \in X(\Omega)$, are P-unique, then

$$Y \vDash D_X | X \quad \Leftrightarrow \quad \tau \underset{P}{=} \big(E^{X=0}(Y), E^{X=1}(Y), \ldots, E^{X=J}(Y)\big), \tag{9.6}$$

where $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$ is a multivariate random variable consisting of $J+1$ true outcome variables. Hence, under these assumptions $Y \vDash D_X | X$ holds if and only if

$$\tau_x \underset{P}{=} E^{X=x}(Y), \quad \forall\, x \in X(\Omega) = \{0, 1, \ldots, J\}. \tag{9.7}$$

Also remember, according to RS-Remark 3.20, if $P(X{=}x) > 0$, then $E^{X=x}(Y) = E(Y|X{=}x)$ [see RS-Eq. (3.24)]. $\triangleleft$

Intuitively speaking, according to Proposition (9.5), $Y \vDash D_X | X$ means that X is the only cause of Y. No other random variable that is prior or simultaneous to X determines the expectation of Y. Of course, this is unrealistically restrictive in many applications. In contrast, their generalizations presented in the next section, where we additionally condition on a covariate of X, *are* realistic.

9.1.2 Conditional SR-Conditions

Now we turn to those Suppes-Reichenbach conditions that involve conditioning on another random variable Z on $(\Omega, \mathscr{A}, P)$. Note that $Z = (Z_1, \ldots, Z_m)$ may be a multivariate random variable on $(\Omega, \mathscr{A}, P)$ consisting of m random variables.

Remark 9.9 [$(X{=}x, Z)$-Conditional Independence of Y and D_X] The first one, which is defined in Box 9.1 (v), is called *$(X{=}x, Z)$-conditional independence of Y and D_X*, and denoted $Y \perp\!\!\!\perp D_X | (X{=}x, Z)$ or $Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X | Z$. Hence,

$$Y \perp\!\!\!\perp D_X | (X{=}x, Z) \quad \Leftrightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X | Z, \tag{9.8}$$

that is, the two symbols denote the same kind of Z-conditional independence of Y from D_X given a value x of X. $\triangleleft$

Remark 9.10 [$(X{=}x, Z)$-Conditional Mean-Independence of Y from D_X] The next condition, defined in Box 9.1 (vi), is *$(X{=}x, Z)$-conditional mean-independence of the outcome variable Y from D_X*, which is denoted by $Y \vDash D_X | (X{=}x, Z)$. With this condition we postulate that the D_X-conditional expectation of the outcome variable Y with respect to the measure $P^{X=x}$ does not depend on the global potential confounder D_X, once we condition on the random variable Z. If we additionally assume that Z is a covariate of X, then $\sigma(Z) \subset \sigma(D_X)$ [see Def. 4.11 (iv) and Rem. 4.16] and $\sigma(D_X, Z) = \sigma(D_X)$ [see RS-Prop. (2.19)]. In this case,

$$E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z, D_X). \tag{9.9}$$

[see RS-Def. 5.4] and

$$Y \vDash D_X | (X{=}x, Z) \quad \Leftrightarrow \quad E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z). \tag{9.10}$$

If Z is discrete and $P^{X=x}(Z=z) > 0$ for all values $z \in Z(\Omega)$, then

$$Y \vDash D_X | (X=x, Z) \quad \Leftrightarrow \quad \forall\, \omega \in \Omega\colon E^{X=x}(Y|D_X)(\omega) = E^{X=x}(Y|Z=z), \quad \text{if } \omega \in \{Z=z\} \tag{9.11}$$

(see Exercise 9-5). ◁

Remark 9.11 [(X,Z)-Conditional Independence of Y and D_X] The last but one causality condition, which is defined in Box 9.1 (vii), is *(X,Z)-conditional independence of the outcome variable Y from D_X*. It is denoted by $Y \perp\!\!\!\perp D_X | (X,Z)$. With this condition we implicitly postulate that the (X, D_X)-conditional distribution of the outcome variable Y does not depend on the global potential confounder D_X once we condition on the putative cause variable X and the random variable Z. That is, we implicitly postulate that $P_{Y|X,Z}$ is a version of the conditional distribution $P_{Y|X,D_X}$ (see again SN-ch. 17). ◁

Remark 9.12 [(X,Z)-Conditional Mean-Independence of Y From D_X] The last causality condition defined in Box 9.1 (viii) is *(X,Z)-conditional mean-independence of Y from D_X*, denoted $Y \vDash D_X | (X,Z)$. With this condition we postulate that the (X, D_X)-conditional expectation of Y actually does not depend on the global potential confounder D_X of X, once we condition on X and Z. That is, we postulate that $E(Y|X,Z)$ is a version of the conditional expectation $E(Y|X,D_X)$ (see RS-ch. 4). ◁

Remark 9.13 [An Equivalent Formulation of $Y \vDash D_X | (X,Z)$] If we assume that Z is a covariate of X, then $\sigma(Z) \subset \sigma(D_X)$ [see Def. 4.11 (iv) and Rem. 4.16] and $\sigma(D_X) = \sigma(D_X, Z)$. This implies

$$\sigma(X, D_X, Z) \;=\; \sigma(X, D_X) \tag{9.12}$$

and

$$E(Y|X,Z,D_X) \underset{P}{=} E(Y|X,D_X) \tag{9.13}$$

(see RS-Def. 4.4). Hence, if Z is a covariate of X, then

$$Y \vDash D_X | (X,Z) \quad \Leftrightarrow \quad E(Y|X,D_X) \underset{P}{=} E(Y|X,Z)\,. \tag{9.14}$$

◁

In the following theorem we present a condition that is equivalent to $Y \vDash D_X | (X,Z)$ if X is discrete. This theorem is a generalization of Theorem 9.7.

Theorem 9.14 [A Condition Equivalent to $Y \vDash D_X | (X,Z)$ if X is Discrete]
Let the Assumptions 9.1 (a) to (c) and (g) hold and assume $P(X=x) > 0$ for all $x \in X(\Omega)$. Then

$$Y \vDash D_X | (X,Z) \quad \Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; E^{X=x}(Y|Z,D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z)\,. \tag{9.15}$$

(Proof p. 293)

Remark 9.15 [True Outcome Variables as Functions of Z] Under the Assumptions 9.1 (a) to (e), which includes P-uniqueness of the true outcome variables τ_x, $x \in X(\Omega)$,

$$Y \vDash D_X | (X,Z) \quad \Leftrightarrow \quad \tau \underset{P}{=} \big(E^{X=0}(Y|Z), E^{X=1}(Y|Z), \ldots, E^{X=J}(Y|Z)\big), \tag{9.16}$$

Table 9.1. (X, Z)-conditional mean-independence of Y from D_X

Person u	Sex z	$P(U=u)$	$P(X=1 \mid U=u)$	$E^{X=0}(Y \mid U=u)$	$E^{X=1}(Y \mid U=u)$	$CTE_{U;10}(u)$	$P(U=u \mid X=0)$	$P(U=u \mid X=1)$
Tom	m	1/6	6/8	90	95	5	2/19	6/29
Tim	m	1/6	7/8	90	95	5	1/19	7/29
Joe	m	1/6	6/8	90	95	5	2/19	6/29
Jim	m	1/6	7/8	90	95	5	1/19	7/29
Ann	f	1/6	1/8	100	110	10	7/19	1/29
Sue	f	1/6	2/8	100	110	10	6/19	2/29

	$x=0$	$x=1$	
$E(\tau_x)$:	93.333	100	$ATE_{10} = 6.667$
$E(Y \mid X=x)$:	96.842	96.552	$PFE_{10} = -.29$
$E(\tau_x \mid Z=m)$:	90	95	$CTE_{Z;10}(m) = 5$
$E(Y \mid X=x, Z=m)$:	90	95	$PFE_{Z;10}(m) = 5$
$E(\tau_x \mid Z=f)$:	100	110	$CTE_{Z;10}(f) = 10$
$E(Y \mid X=x, Z=f)$:	100	110	$PFE_{Z;10}(f) = 10$

where τ again denotes a multivariate random variable consisting of $J+1$ true outcome variables. Hence, if Z is a (possibly multivariate) covariate of X, then $Y \vDash D_X \mid (X, Z)$ means that Z contains all potential confounders of X that, aside from X, determine the (X, D_X)-conditional expectation of Y. To emphasize, it does *not mean* that there are no intermediate variables between X and Y that determine Y over and above X and Z. $\lhd$

9.1.3 *Example:* (X, Z)*-Conditional Mean-Independence of* Y *From* D_X

Consider the random experiment presented in Table 9.1. We assume that it is a simple experiment that has the same structure as the examples treated in section 6.5. That is,

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big),$$

as specified in section 6.5, is the regular probabilistic causality setup. Again, X represents a treatment variable and U is a global potential confounder of X.

In this example, $Y \vDash U \mid X$, which by definition is equivalent to

$$E(Y \mid X, U) \underset{P}{=} E(Y \mid X), \tag{9.17}$$

does *not hold*. In this example, $P(X=x, U=u) > 0$, for all pairs $(x, u) \in X(\Omega) \times U(\Omega)$. Therefore, in this example, Equation (9.17) is equivalent to

$$\forall (x,u) \in X(\Omega) \times U(\Omega): \ E(Y|X{=}x, U{=}u) \ = \ E(Y|X{=}x). \tag{9.18}$$

Hence, with Equation (9.17) we postulate that, for each value x of X, the conditional expectation values $E(Y|X{=}x, U{=}u)$ are identical for all units u. If $P(X{=}x, U{=}u) > 0$, then $E(Y|X{=}x, U{=}u) = E^{X=x}(Y|U{=}u)$. Hence, an inspection of Table 9.1 shows that Equation (9.18), and with it Equation (9.17), *do not hold*. These equations would hold if, for example, all six values $E^{X=0}(Y|U{=}u)$ of $\tau_0 = E^{X=0}(Y|U)$ would be equal to 90 and all six values $E^{X=1}(Y|U{=}u)$ of $\tau_1 = E^{X=1}(Y|U)$ would be equal to 95. Note that the numbers 90 and 95 are arbitrary examples. Every other pair of two numbers would do as well.

In contrast, in this example, $Y \vDash U\,|\,(X,Z)$ *does hold*. This symbol stands for

$$E(Y|X,U) \underset{P}{=} E(Y|X,Z) \tag{9.19}$$

[see Eqs. (9.13) and (9.14) for $D_X = U$]. Furthermore, because $P(X{=}x, U{=}u) > 0$ for all $(x,u) \in X(\Omega) \times U(\Omega)$, this equation is equivalent to

$$\forall\, (x,u) \in X(\Omega) \times U(\Omega): E(Y|X{=}x, U{=}u) \ = \ E(Y|X{=}x, Z{=}z). \tag{9.20}$$

Hence, the conjunction of $Y \vDash U\,|\,(X,Z)$, $D_X = U$, and $P(X{=}x, U{=}u) > 0$ implies that there are no differences between persons u with respect to their conditional expectation values of the outcome variable Y given treatment condition x and person u. Also remember, if $P(X{=}x, U{=}u) > 0$, then $E(Y|X{=}x, U{=}u) = E^{X=x}(Y|U{=}u)$. Hence, an inspection of Table 9.1 shows that Proposition (9.20) and with it Equation (9.19) hold. All four values $E^{X=0}(Y|U{=}u)$ of τ_0 of the male persons are identical (to 90), all four values $E^{X=1}(Y|U{=}u)$ of τ_1 of the males are identical (to 95), all two values $E^{X=0}(Y|U{=}u)$ of τ_0 of the females are identical (to 100), and all two values $E^{X=1}(Y|U{=}u)$ of τ_1 of the females are identical (to 110). Hence, we may also say that the true outcome variables τ_0 and τ_1 are *homogeneous* given sex.

9.2 Implications Among the Suppes-Reichenbach Conditions

Table 9.2 summarizes the implications of the Suppes-Reichenbach conditions (defined in Box 9.1) among each other and displays the assumptions under which these implications hold. The proofs are found in the theorems and corollaries treated in the present section and in the solution to Exercise 9-6.

A first property of the Suppes-Reichenbach conditions $Y \perp\!\!\!\perp D_X|X$ and $Y \perp\!\!\!\perp D_X|(X,Z)$ is that they imply conditional mean-independence. This is detailed in the following corollary, which immediately follows from RS-Theorem 6.8.

Corollary 9.16 [$Y \perp\!\!\!\perp D_X|X$ **and** $Y \perp\!\!\!\perp D_X|(X,Z)$ **Imply Mean-Independence**]
Let the Assumptions 9.1 (a) and (c) hold. Then

$$Y \perp\!\!\!\perp D_X|X \quad \Rightarrow \quad Y \vDash D_X|X. \tag{9.21}$$

If we additionally assume 9.1 (g), then

$$Y \perp\!\!\!\perp D_X|(X,Z) \quad \Rightarrow \quad Y \vDash D_X|(X,Z). \tag{9.22}$$

Table 9.2. Implications among the Suppes-Reichenbach conditions

	$Y \vDash D_X \mid (X{=}x, Z)$	$Y \vDash D_X \mid (X, Z)$	$Y \vDash D_X \mid (X{=}x)$	$Y \vDash D_X \mid X$	$Y \perp\!\!\!\perp D_X \mid (X{=}x, Z)$	$Y \perp\!\!\!\perp D_X \mid (X, Z)$	$Y \perp\!\!\!\perp D_X \mid (X{=}x)$
$Y \vDash D_X \mid (X, Z)$	(a)-(c),(g)						
$Y \vDash D_X \mid (X{=}x)$	(a)-(c),(h)						
$Y \vDash D_X \mid X$	(a)-(c),(h)	(a),(c),(h)	(a)-(c)				
$Y \perp\!\!\!\perp D_X \mid (X{=}x, Z)$	(a)-(c),(g)						
$Y \perp\!\!\!\perp D_X \mid (X, Z)$	(a)-(c),(g)	(a),(c),(g)			(a)-(c),(g)		
$Y \perp\!\!\!\perp D_X \mid (X{=}x)$	(a)-(c),(h)		(a)-(c)		(a),(b),(h)		
$Y \perp\!\!\!\perp D_X \mid X$	(a)-(c),(h)	(a),(c),(h)	(a)-(c)	(a),(c)	(a),(b),(h)	(a),(h)	(a),(b)

Note: An entry such as (a)-(c) means that the condition in the row implies the condition in the column, provided that the Assumptions 9.1 (a) to (c) hold. Equivalences are omitted.

Hence, under the Assumptions 9.1 (a) and (c), X-conditional independence of the outcome variable Y and a global potential confounder D_X of X implies X-conditional mean-independence of Y from D_X. Similarly, if we consider an additional random variable Z on $(\Omega, \mathscr{A}, P)$, then (X, Z)-conditional independence of Y and D_X implies (X, Z)-conditional mean-independence of Y from D_X.

In the following theorem we consider the implications of $(X{=}x)$-conditional independence of Y from D_X and of $(X{=}x, Z)$-conditional independence of Y from D_X on the corresponding mean-independencies.

Theorem 9.17 [$Y \perp\!\!\!\perp D_X \mid (X{=}x)$ **and** $Y \perp\!\!\!\perp D_X \mid (X{=}x, Z)$ **Imply Mean-Independence**]
Let the Assumptions 9.1 (a) to (c) hold. Then

$$Y \perp\!\!\!\perp D_X \mid (X{=}x) \quad \Rightarrow \quad Y \vDash D_X \mid (X{=}x). \tag{9.23}$$

If we additionally assume 9.1 (g), then

$$Y \perp\!\!\!\perp D_X \mid (X{=}x, Z) \quad \Rightarrow \quad Y \vDash D_X \mid (X{=}x, Z). \tag{9.24}$$

(Proof p. 293)

Hence, under the assumptions of Theorem 9.17, $(X{=}x)$-conditional independence of the outcome variable Y and the potential confounder D_X of X implies $(X{=}x)$-conditional mean-independence of Y from D_X. Furthermore, if we consider an additional random variable Z on $(\Omega, \mathscr{A}, P)$, then $(X{=}x, Z)$-conditional independence of Y and D_X implies $(X{=}x, Z)$-conditional mean-independence of Y from D_X.

In the following theorem we show that X-conditional mean-independence of the outcome variable Y from a global potential confounder D_X of X implies $(X{=}x)$-conditional mean-independence of Y from D_X. This implication also holds if we do not only condition on the putative cause variable X but additionally on a random variable Z.

Theorem 9.18 [Consequences of $Y \vDash D_X | X$ and $Y \vDash D_X | (X, Z)$]
Let the Assumptions 9.1 (a) to (c) hold. Then

$$Y \vDash D_X | X \quad \Rightarrow \quad Y \vDash D_X | (X{=}x). \tag{9.25}$$

If we additionally assume 9.1 (g), then

$$Y \vDash D_X | (X, Z) \quad \Rightarrow \quad Y \vDash D_X | (X{=}x, Z). \tag{9.26}$$

(Proof p. 293)

Now we turn to the issue of generalizability, which has already been discussed for the Fisher conditions (see Ths. 8.17, 8.18 and Remarks 8.63, 8.64). Note that the SR-conditions considered in the following theorem neither presume that X is discrete nor the existence of the true outcome variables. Reading this theorem, remember that $W \in \mathscr{W}_X$ means that $\sigma(W) \subset D_X$, that is, W is a potential confounder of X [see Def. 4.11 (iv)].

Theorem 9.19 [Generalizability of $Y \perp\!\!\!\perp D_X | X$ and $Y \perp\!\!\!\perp D_X | (X, Z)$]
Let the Assumptions 9.1 (a) hold. Then

$$Y \perp\!\!\!\perp D_X | X \;\Rightarrow\; \forall\, W \in \mathscr{W}_X\colon Y \perp\!\!\!\perp D_X | (X, W). \tag{9.27}$$

If we additionally assume 9.1 (g), then

$$Y \perp\!\!\!\perp D_X | (X, Z) \;\Rightarrow\; \forall\, W \in \mathscr{W}_X\colon Y \perp\!\!\!\perp D_X | (X, Z, W). \tag{9.28}$$

(Proof p. 294)

Remark 9.20 [Methodological Implications of Generalizability] The two properties of Theorem 9.19 are referred to as *generalizability* of $Y \perp\!\!\!\perp D_X | X$ and $Y \perp\!\!\!\perp D_X | (X, Z)$, respectively. According to Proposition (9.27), once we have X-conditional independence of Y and D_X, then we also have (X, W)-conditional independence of Y and D_X, if W is a potential confounder of X.

Similarly, according to Proposition (9.28), once we have (X, Z)-conditional independence of Y and D_X, then we also have (X, Z, W)-conditional independence of Y and D_X, if W is a potential confounder of X. From a methodological point of view, this consequence of $Y \perp\!\!\!\perp D_X | (X, Z)$ has two positive consequences:

(a) Under $Y \perp\!\!\!\perp D_X | X$, there is no need to control for further potential confounders of X, at least not for the purpose of justifying a causal interpretation of the W-conditional dependence of Y on X as described by a conditional distribution $P_{Y|X,W}$ or by a conditional expectation $E(Y|X,W)$.
(b) Under $Y \perp\!\!\!\perp D_X | (X, Z)$, a causal interpretation of the (Z, W)-conditional dependence of Y on X as described by $P_{Y|X,Z,W}$ or by $E(Y|X,Z,W)$ is also possible, provided that W is a potential confounder of X.

◁

Now we turn to generalizability of X-conditional mean-independence of Y from D_X and of (X,Z)-conditional mean-independence of Y from D_X.

Theorem 9.21 [Generalizability of $Y \vDash D_X | X$ and $Y \vDash D_X | (X,Z)$]
Let the Assumptions 9.1 (a) and (c) hold. Then

$$Y \vDash D_X | X \quad \Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon\, Y \vDash D_X | (X,W)\,. \tag{9.29}$$

If we additionally assume 9.1 (g), then

$$Y \vDash D_X | (X,Z) \quad \Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon\, Y \vDash D_X | (X,Z,W)\,. \tag{9.30}$$

(Proof p. 294)

Hence, under the assumptions of Theorem 9.21, X-conditional mean-independence of the outcome variable Y from a global potential confounder D_X of X implies (X,W)-conditional mean-independence of Y from D_X, whenever W is a potential confounder of X [see Prop. (9.29)]. This is what we mean saying that $Y \vDash D_X | X$ is *generalizable*.

Furthermore, (X,Z)-conditional mean-independence of Y from D_X implies (X,Z,W)-conditional mean-independence of Y from D_X [see Prop. (9.30)], whenever W is a potential confounder of X. This is what we mean saying that $Y \vDash D_X | (X,Z)$ is *generalizable*.

In the following theorem we show that $(X{=}x)$-conditional mean-independence of Y from D_X as well as $(X{=}x, Z)$-conditional mean-independence of Y from D_X are generalizable as well.

Theorem 9.22 [Generalizability of $Y \vDash D_X | (X{=}x)$ and $Y \vDash D_X | (X{=}x, Z)$]
Let the Assumptions 9.1 (a) and (c) hold. Then

$$Y \vDash D_X | (X{=}x) \quad \Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon\, Y \vDash D_X | (X{=}x, W)\,. \tag{9.31}$$

If we additionally assume 9.1 (g), then

$$Y \vDash D_X | (X{=}x, Z) \quad \Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon\, Y \vDash D_X | (X{=}x, Z, W)\,. \tag{9.32}$$

(Proof p. 294)

Remark 9.23 [Immediate Implications] Combining Propositions (9.23) and (9.31) yields

$$Y \perp\!\!\!\perp D_X | (X{=}x) \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon\, Y \vDash D_X | (X{=}x, W), \tag{9.33}$$

provided that the Assumptions 9.1 (a) to (c) hold. Correspondingly, if we additionally assume 9.1 (g), then combining Propositions (9.24) and (9.32) yields

$$Y \perp\!\!\!\perp D_X | (X{=}x, Z) \quad \Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon\, Y \vDash D_X | (X{=}x, Z, W)\,. \tag{9.34}$$

$\triangleleft$

9.3 Consequences for the Rosenbaum-Rubin Conditions

Now we study the implications of the Suppes-Reichenbach conditions on the Rosenbaum-Rubin conditions. Remember, in chapter 7 we showed that the Rosenbaum-Rubin conditions imply unbiasedness. Hence, if the Suppes-Reichenbach conditions imply the Rosenbaum-Rubin conditions, then the Suppes-Reichenbach conditions also imply unbiasedness.

Table 9.3 summarizes the implications of some Suppes-Reichenbach conditions (defined in Box 9.1) on some Rosenbaum-Rubin conditions and displays the assumptions under which these implications hold. The proofs are found in the theorems and corollaries of the present section and in the solution to Exercise 9-7.

Remark 9.24 [Consequences of the Suppes-Reichenbach Independence Conditions] Table 9.3 only lists the implications of the Suppes-Reichenbach mean-independence conditions. However, according to Table 9.2, these mean-independence conditions follow from their corresponding independence conditions. More precisely,

(i) Under Assumptions 9.1 (a) to (c), and (g), $Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z)$ implies $Y \vDash D_X \,|\, (X{=}x, Z)$
(ii) Under Assumptions 9.1 (a), (c), and (g), $Y \perp\!\!\!\perp D_X \,|\, (X, Z)$ implies $Y \vDash D_X \,|\, (X, Z)$
(iii) Under Assumptions 9.1 (a) to (c), $Y \perp\!\!\!\perp D_X \,|\, (X{=}x)$ implies $Y \vDash D_X \,|\, (X{=}x)$
(iv) Under Assumptions 9.1 (a) and (c), $Y \perp\!\!\!\perp D_X \,|\, X$ implies $Y \vDash D_X \,|\, X$.

Hence, we can replace the mean-independence conditions in the first column of Table 9.3 by their corresponding independence conditions, thus receiving a new Table of implications. ◁

9.3.1 Consequences of $Y \vDash D_X \,|\, (X{=}x)$ and $Y \vDash D_X \,|\, X$

In Theorem 9.25, we consider a single value x of X with $P(X{=}x) > 0$ and show that $(X{=}x)$-conditional mean-independence of the outcome variable Y from a global potential confounder of X implies independence of the true outcome variable τ_x and the putative cause variable X. A crucial assumption in this theorem is P-uniqueness of τ_x [see Ass. 9.1 (i)].

Theorem 9.25 [An Implication of $Y \vDash D_X \,|\, (X{=}x)$ on τ_x]
Let the Assumptions 9.1 (a) to (d) hold and assume that τ_x is P-unique. Then

$$Y \vDash D_X \,|\, (X{=}x) \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X. \tag{9.35}$$

(Proof p. 295)

Remark 9.26 [Further Consequences of $Y \vDash D_X \,|\, (X{=}x)$] Under the assumptions of Theorem 9.25, the consequences of $\tau_x \perp\!\!\!\perp X$ listed in Table 7.1 allow us to conclude that $(X{=}x)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X, that is, $Y \vDash D_X \,|\, (X{=}x)$, also implies

(i) $\tau_x \perp\!\!\!\perp 1_{X=x}$
(ii) $\tau_x \vDash X$
(iii) $\forall\, x' \in X(\Omega): \tau_x \vDash 1_{X=x'}$
(iv) $\tau_x \vDash 1_{X=x}$
(v) $E(Y \,|\, X{=}x) \vdash D_X$.

Table 9.3. Implications of the Suppes-Reichenbach conditions on some Rosenbaum-Rubin conditions

	$\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$	$\tau_x \perp\!\!\!\perp X \mid Z$	$\tau \perp\!\!\!\perp X \mid Z$	$\tau \perp\!\!\!\perp 1_{X=x} \mid Z$	$\tau_x \perp\!\!\!\perp 1_{X=x}$	$\tau_x \perp\!\!\!\perp X$	$\tau \perp\!\!\!\perp X$	$\tau \perp\!\!\!\perp 1_{X=x}$
$Y \vDash D_X \mid (X{=}x, Z)$	(a)-(d), (h), (i)	(a)-(d), (h), (i)						
$Y \vDash D_X \mid (X, Z)$	(a)-(d), (h), (i)	(a)-(d), (h), (i)	(a)-(e), (h), (j)	(a)-(e), (h), (j)				
$Y \vDash D_X \mid (X{=}x)$	(a)-(d), (g), (i)	(a)-(d), (g), (i)			(a)-(d), (i)	(a)-(d), (i)		
$Y \vDash D_X \mid X$	(a)-(d), (g), (i)	(a)-(d), (g), (i)	(a)-(e), (g), (j)	(a)-(e), (g), (j)	(a)-(d), (i)	(a)-(d), (i)	(a)-(e), (j)	(a)-(e), (j)

Note: An entry such as (a)-(d), (i) means that the condition in the row implies the condition in the column, provided that the Assumptions 9.1 (a) to (d) and (i) hold. Consequences of the Rosenbaum-Rubin conditions displayed in the columns of the table are found in Tables 7.1 and 7.2.

Remember, $E(Y|X{=}x) \vdash D_X$ denotes unbiasedness of $E(Y|X{=}x)$, which is defined by the conjunction of P-uniqueness of τ_x and $E(Y|X{=}x) = E(\tau_x)$ (see Def. 6.3). ◁

Remark 9.27 [Comparing $Y \vDash D_X | (X{=}x)$ to $D_X \perp\!\!\!\perp 1_{X=x}$] Note that $\tau_x \perp\!\!\!\perp 1_{X=x}$ does not only follow from $Y \vDash D_X | (X{=}x)$, but also from the Fisher condition $D_X \perp\!\!\!\perp 1_{X=x}$ because $\sigma(\tau_x) \subset D_X$. However, in contrast to $D_X \perp\!\!\!\perp 1_{X=x}$, assuming $Y \vDash D_X | (X{=}x)$ does not imply that $E^{X=x}(Y|D_X)$ is P-unique [cf. Th. 8.22 (iii)]. This is why in Theorem 9.25 we need the assumption that $\tau_x = E^{X=x}(Y|D_X)$ is P-unique [see Ass. 9.1 (i)]. ◁

According to the following theorem, $(X{=}x)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X also implies that the true outcome variable τ_x and the putative cause variable X are Z-*conditionally* independent.

Theorem 9.28 [Another Consequence of $Y \vDash D_X | (X{=}x)$]
Let the Assumptions 9.1 (a) to (d) and (g) hold, and assume that τ_x is P-unique. Then

$$Y \vDash D_X | (X{=}x) \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X \,|\, Z. \tag{9.36}$$

(Proof p. 295)

Hence, if we assume that the true outcome variable τ_x is P-unique, then $Y \vDash D_X | (X{=}x)$ implies that τ_x and the putative cause variable X are Z-conditionally independent. Note that Z can be any random variable on $(\Omega, \mathscr{A}, P)$. To emphasize, Z does not have to be a covariate of X.

Remark 9.29 [Further Consequences of $Y \vDash D_X | (X{=}x)$] Under the assumptions of Theorem 9.28, the consequences of $\tau_x \perp\!\!\!\perp X \,|\, Z$ listed in Table 7.2 allow us to conclude that $(X{=}x)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X also implies

(i) $\tau_x \perp\!\!\!\perp 1_{X=x} | Z$
(ii) $\tau_x \vDash X \,|\, Z$
(iii) $\tau_x \vDash 1_{X=x} | Z$.

If we additionally asssume that Z is a covariate of X, then

(iv) $E^{X=x}(Y|Z) \vdash D_X$.

Remember, $E^{X=x}(Y|Z) \vdash D_X$ denotes unbiasedness of $E^{X=x}(Y|Z)$, which is defined by the conjunction of P-uniqueness of τ_x and $E^{X=x}(Y|Z) = E(\tau_x|Z)$ [see Def. 6.13 (i)]. ◁

Instead of a single value x of the putative cause variable X, now we consider X itself and turn to X-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X, denoted $Y \vDash D_X | X$ and defined by $E(Y|X, D_X) \underset{P}{=} E(Y|X)$ [see Box 9.1 (iv)].

Theorem 9.30 [Consequences of $Y \vDash D_X | X$]
Let the Assumptions 9.1 (a) to (d) hold and assume that τ_x is P-unique. Then

$$Y \vDash D_X | X \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X. \tag{9.37}$$

If we additionally assume 9.1 (g), then

$$Y \vDash D_X|X \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X\,|\,Z. \tag{9.38}$$

(Proof p. 295)

Remark 9.31 [More Implications] Note that the consequence $\tau_x \perp\!\!\!\perp X$ of Proposition (9.37) has itself several consequences that already have been listed in Remark 9.26. This allows us to conclude, for example,

$$Y \vDash D_X|X \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x}, \tag{9.39}$$

provided that the Assumptions 9.1 (a) to (d) hold and that τ_x is P-unique.

Similarly, the consequence $\tau_x \perp\!\!\!\perp X\,|\,Z$ of Proposition (9.38) has itself several consequences listed in Remark 9.29, which, if we add Assumption 9.1 (g), allows us to conclude

$$Y \vDash D_X|X \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x}|Z. \tag{9.40}$$

$\lhd$

In Theorem 9.30, we considered a single true outcome variable τ_x. However, X-conditional mean-independence of Y from a global potential confounder D_X of X also has consequences for the multivariate true outcome variable $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$.

Theorem 9.32 [Consequences of $Y \vDash D_X|X$ for τ]
Let the Assumptions 9.1 (a) to (e) hold and assume that all τ_x, $x \in X(\Omega)$, are P-unique. Then

$$Y \vDash D_X|X \quad \Rightarrow \quad \tau \perp\!\!\!\perp X. \tag{9.41}$$

If we additionally assume 9.1 (g), then

$$Y \vDash D_X|X \quad \Rightarrow \quad \tau \perp\!\!\!\perp X\,|\,Z. \tag{9.42}$$

(Proof p. 296)

Remark 9.33 [Immediate Implications] Under the assumptions of Theorem 9.32,

$$Y \vDash D_X|X \quad \Rightarrow \quad \tau \perp\!\!\!\perp 1_{X=x}\,. \tag{9.43}$$

This immediately follows from Proposition (9.41), $\sigma(1_{X=x}) \subset \sigma(X)$, and RS-Box 2.1 (iv). Similarly, if we additionally consider a random variable Z on $(\Omega, \mathscr{A}, P)$, then

$$Y \vDash D_X|X \quad \Rightarrow \quad \tau \perp\!\!\!\perp 1_{X=x}|Z, \tag{9.44}$$

which immediately follows from Proposition (9.42), $\sigma(1_{X=x}) \subset \sigma(X)$, and RS-Box 6.1 (vi). $\lhd$

Remark 9.34 [Further Implications of $Y \vDash D_X|X$] Again under the assumptions of Theorem 9.32, according to Proposition (9.41) and Table 7.1, $Y \vDash D_X|X$ also implies all causality conditions listed in the last row Table 7.1, including unbiasedness of $E(Y|X)$. Furthermore, according to Proposition (9.42) and Table 7.2, $Y \vDash D_X|X$ also implies all other causality conditions listed in the last row of Table 7.2. Finally, if we additionally assume that Z is a covariate of X, then $Y \vDash D_X|X$ also implies that $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, are unbiased. $\lhd$

9.3.2 Consequences of $Y \vDash D_X | (X{=}x, Z)$ and $Y \vDash D_X | (X, Z)$

Now we extend the results of section 9.3.1 to the case in which we additionally condition on a covariate Z of X. We start considering a single conditional expectation $E^{X=x}(Y|Z)$ and a single conditional expectation value $E(Y|X{=}x, Z{=}z)$. Remember, that $E^{X=x}(Y|Z)$ denotes the Z-conditional expectation of Y with respect to the probability measure $P^{X=x}$. Finally, see Box 9.1 (vi) for the definition of the term $Y \vDash D_X | (X{=}x, Z)$, that is, of $(X{=}x, Z)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X.

Theorem 9.35 [Implications of $Y \vDash D_X | (X{=}x, Z)$]
Let the Assumptions 9.1 (a) to (d) and (h) hold, and assume that τ_x is P-unique. Then

$$Y \vDash D_X | (X{=}x, Z) \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X \,|\, Z. \tag{9.45}$$

(Proof p. 296)

Remark 9.36 [More Implications] Again note that the consequence $\tau_x \perp\!\!\!\perp X \,|\, Z$ in Proposition (9.45) has itself several consequences that have already been listed in Remark 9.29. If Z is a covariate of X, then according to this remark, we can conclude

$$Y \vDash D_X | (X{=}x, Z) \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x} \,|Z \tag{9.46}$$

and

$$Y \vDash D_X | (X{=}x, Z) \quad \Rightarrow \quad E^{X=x}(Y|Z) \vdash D_X. \tag{9.47}$$

Hence, under the assumptions of Theorem 9.35, $Y \vDash D_X | (X{=}x, Z)$ also implies unbiasedness of $E^{X=x}(Y|Z)$. ◁

In the following theorem, we consider a single value z of a covariate Z of X and a single value x of the putative cause variable X, assume $P(X{=}x, Z{=}z) > 0$, and use the probability measure $P^{Z=z}: \mathscr{A} \to [0,1]$ defined by

$$P^{Z=z}(A) \;=\; P(A \,|\, Z{=}z), \quad \forall A \in \mathscr{A}.$$

Reading this theorem note that P-uniqueness of τ_x implies that it is also $P^{Z=z}$-unique but not vice versa [see RS-Box 5.1 (v)]. In this sense, assuming $P^{Z=z}$-uniqueness of τ_x is less restrictive than P-uniqueness of τ_x.

Theorem 9.37 [Unbiasedness of $E(Y|X{=}x, Z{=}z)$]
Let the Assumptions 9.1 (a) to (d), and (h) hold, let $z \in Z(\Omega)$ be a value of Z such that $P(X{=}x, Z{=}z) > 0$, and assume that τ_x is $P^{Z=z}$-unique. Then

$$Y \vDash D_X | (X{=}x, Z) \quad \Rightarrow \quad E^{X=x}(Y|Z{=}z) \vdash D_X. \tag{9.48}$$

(Proof p. 296)

Hence, if the assumptions of Theorem 9.37 hold, which include that Z is a covariate of X and that τ_x is $P^{Z=z}$-unique, then $(X{=}x, Z)$-conditional mean-independence of Y

from a global potential confounder of X implies that the conditional expectation value $E(Y|X{=}x, Z{=}z)$ is unbiased [see Def. 6.13 (ii)].

Now we turn to an implication of (X, Z)-conditional mean-independence of the outcome variable Y from a global potential confounder of X. Again, we assume that Z is a covariate of X.

Theorem 9.38 [An Implication of $Y \vDash D_X | (X, Z)$]
Let the Assumptions 9.1 (a) to (e) and (h) hold and assume that all τ_x, $x \in X(\Omega)$, are P-unique. Then

$$Y \vDash D_X | (X, Z) \quad \Rightarrow \quad \tau \perp\!\!\!\perp X \,|\, Z \tag{9.49}$$

and

$$Y \vDash D_X | (X, Z) \quad \Rightarrow \quad \tau \perp\!\!\!\perp 1_{X=x} | Z. \tag{9.50}$$

(Proof p. 297)

Remark 9.39 [Further Implications of $Y \vDash D_X | (X, Z)$] Hence, if we assume that Z is a covariate of X and all τ_x are P-unique, then $Y \vDash D_X | (X, Z)$ implies $\tau \perp\!\!\!\perp X \,|\, Z$ and $\tau \perp\!\!\!\perp 1_{X=x} | Z$. According to Table 7.2, $\tau \perp\!\!\!\perp X \,|\, Z$ in turn implies that $E(Y | X, Z)$ is unbiased and that all other causality conditions indicated in the last row of that table hold. For example, under the Assumptions 9.1 (a) to (d), (h) and that τ_x is P-unique [i. e., Ass. 9.1 (i)], these implications include

$$Y \vDash D_X | (X, Z) \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp X \,|\, Z \tag{9.51}$$

and

$$Y \vDash D_X | (X, Z) \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x} | Z. \tag{9.52}$$

◁

According to the following theorem, if we presume that z is a value of a covariate Z of X such that $P(X{=}x, Z{=}z) > 0$, then, $Y \vDash D_X | (X, Z)$ also implies that $E^{X=x}(Y | Z{=}z)$ is unbiased, denoted $E^{X=x}(Y | Z{=}z) \vdash D_X$, provided that we assume that τ_x is (at least) $P^{Z=z}$-unique. [Remember, according to RS-Box 5.1 (v), P-uniqueness of $\tau_x = E^{X=x}(Y | D_X)$ implies that it is also $P^{Z=z}$-unique.]

Theorem 9.40 [$Y \vDash D_X | (X, Z)$ Implies Unbiasedness of $E^{X=x}(Y | Z{=}z)$]
Let the Assumptions 9.1 (a) to (d) and (h) hold, let $z \in Z(\Omega)$ be a value of Z such that $P(X{=}x, Z{=}z) > 0$, and assume that τ_x is $P^{Z=z}$-unique. Then

$$Y \vDash D_X | (X, Z) \quad \Rightarrow \quad E^{X=x}(Y | Z{=}z) \vdash D_X. \tag{9.53}$$

(Proof p. 297)

Hence, under the assumptions of Theorem 9.40, (X, Z)-conditional mean-independence of the outcome variable Y from a global potential confounder D_X of X implies that the conditional expectation value $E^{X=x}(Y | Z{=}z)$ is unbiased. Crucial assumptions for this implication are that Z is a covariate of X, the true outcome variable is P-unique, and $P(X{=}x, Z{=}z) > 0$.

Remark 9.41 [Covariate Selection] In contrast to the Fisher condition $D_X \perp\!\!\!\perp X \mid Z$ (see Th. 8.31), the condition $Y \vDash D_X \mid (X, Z)$ cannot be created by a design technique such as randomization. Instead we can only try to find a (possibly multivariate) covariate Z of X such that $Y \vDash D_X \mid (X, Z)$ holds. In other words, $Y \vDash D_X \mid (X, Z)$ can only be used for covariate selection (see section 9.4 for more details). ◁

9.3.3 Consequences for the Prima Facie Effect Variables

Remember, a Z-conditional prima facie effect variable is defined by

$$PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y \mid Z) - E^{X=x'}(Y \mid Z) \tag{9.54}$$

and a Z-conditional causal total effect variable by

$$CTE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x \mid Z) - E(\tau_{x'} \mid Z). \tag{9.55}$$

According to the following theorem, $Y \vDash D_X \mid (X, Z)$ has consequences for unbiasedness of the function $PFE_{Z;xx'}$.

Theorem 9.42 [Unbiasedness of a Conditional Prima Facie Effect Variable]
Let the Assumptions 9.1 (a) to (d), (f), and (j) hold, and assume that Z is a covariate of X. Then

$$Y \vDash D_X \mid (X, Z) \quad \Rightarrow \quad PFE_{Z;xx'}(Z) \underset{P}{=} CTE_{Z;xx'}(Z). \tag{9.56}$$

(Proof p. 298)

Hence, under the assumptions of Theorem 9.42, (X, Z)-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X implies that the Z-conditional prima facie effect function $PFE_{Z;xx'}(Z)$ is unbiased.

Now we consider a single value z of the covariate Z. According to the following theorem, $Y \vDash D_X \mid (X, Z)$ has consequences for unbiasedness of the $(Z{=}z)$-conditional prima facie effect

$$PFE_{Z;xx'}(z) = E^{X=x}(Y \mid Z{=}z) - E^{X=x'}(Y \mid Z{=}z), \tag{9.57}$$

which is uniquely defined if $P(X{=}x, Z{=}z), P(X{=}x', Z{=}z) > 0$. Also remember,

$$CTE_{Z;xx'}(z) = E(\tau_x \mid Z{=}z) - E(\tau_{x'} \mid Z{=}z) \tag{9.58}$$

[see Eq. (8.56)].

Theorem 9.43 [Unbiasedness of a $(Z{=}z)$-Conditional Prima Facie Effect]
Let the Assumptions 9.1 (a) to (d) and (f) hold, and assume that Z is a covariate of X. Furthermore, assume $P(X{=}x, Z{=}z)$, $P(X{=}x', Z{=}z) > 0$ and that τ_x and $\tau_{x'}$ are $P^{Z=z}$-unique. Then

$$Y \vDash D_X \mid (X, Z) \quad \Rightarrow \quad PFE_{Z;xx'}(z) = CTE_{Z;xx'}(z). \tag{9.59}$$

(Proof p. 298)

Hence, under the assumptions of Theorem 9.43, (X, Z)-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X implies that the prima facie effect $PFE_{Z;xx'}(z)$ is unbiased.

9.3.4 Example: (X, Z)-Conditional Mean-Independence of Y From D_X

As already shown in Example 9.1.3, Table 9.1 describes a random experiment in which $Y \vDash D_X|(X,Z)$ holds. In this example, the person variable U is a global potential confounder, that is, $D_X = U$ and Z (*sex*) is U-measurable, which implies

$$E(Y|X,Z,U) \underset{P}{=} E(Y|X,D_X) \underset{P}{=} E(Y|X,U).$$

Therefore, the values of the true outcome variables τ_x are the individual conditional expectation values $E^{X=x}(Y|U=u) = E(Y|X=x, U=u)$, $x = 0,1$. As is easily seen in Table 9.1, the true outcomes under control and the true total effects of treatment 1 vs. 0 are the same for all male persons and they are the same for all females. For all males, the values of the true outcomes *given control* are

$$\tau_0(\omega) \;=\; E^{X=0}(Y|Z=m) \;=\; 90, \quad \text{if } \omega \in \{Z=m\}$$

and for all females they are

$$\tau_0(\omega) \;=\; E^{X=0}(Y|Z=f) \;=\; 100, \quad \text{if } \omega \in \{Z=f\}.$$

Given treatment, the true outcomes for all males are

$$\tau_1(\omega) \;=\; E^{X=1}(Y|Z=m) \;=\; 95, \quad \text{if } \omega \in \{Z=m\}$$

and for all females they are

$$\tau_1(\omega) \;=\; E^{X=1}(Y|Z=f) \;=\; 110, \quad \text{if } \omega \in \{Z=f\}.$$

Hence,

$$\tau_x \;=\; E^{X=x}(Y|U) \;=\; E^{X=x}(Y|Z), \quad x = 0,1. \tag{9.60}$$

According to Theorem 9.14, this implies $Y \vDash U|(X,Z)$.

Note that the true outcome variables τ_0 and τ_1 are uniquely defined, because, in this example, $D_X = U$ and $P(X=0\,|\,U)$, $P(X=1\,|\,U) > 0$. This implies that τ_0 and τ_1 are also P-unique. Furthermore, according to RS-Equation (5.36) and Equation (9.60),

$$E(Y|X,U) \underset{P}{=} \sum_{x \in X(\Omega)} E^{X=x}(Y|U) \cdot 1_{X=x} \underset{P}{=} \sum_{x \in X(\Omega)} E^{X=x}(Y|Z) \cdot 1_{X=x} \underset{P}{=} E(Y|X,Z).$$

Hence, in this example, in which $\sigma(Z) \subset \sigma(D_X)$,

$$E(Y|X,Z,D_X) \underset{P}{=} E(Y|X,D_X) \underset{P}{=} E(Y|X,U) \underset{P}{=} E(Y|X,Z).$$

This is exactly what defines $Y \vDash U|(X,Z)$ [see Box 9.1 (viii) for $D_X = U$]. Furthermore, because τ_0 and τ_1 are P-unique, according to Remark 9.39, $Y \vDash D_X|(X,Z)$ implies that $E(Y|X,Z)$ is unbiased.

Because, $P(X=x, Z=z) > 0$ for all pairs (x,z) of values of X and Z, according to Theorem 9.42, $Y \vDash D_X|(X,Z)$ also implies that the conditional prima facie effects

$$PFE_{Z;10}(m) \;=\; E(Y|X=1, Z=m) - E(Y|X=0, Z=m) \;=\; 95 - 90 \;=\; 5$$

and

$$PFE_{Z;10}(f) \;=\; E(Y\,|\,X{=}1, Z{=}f) - E(Y\,|\,X{=}0, Z{=}f) \;=\; 110 - 100 \;=\; 10$$

are unbiased as well, that is, $PFE_{Z;10}(m) = CTE_{Z;10}(m)$ and $PFE_{Z;10}(f) = CTE_{Z;10}(f)$. In other words, the $(Z{=}z)$-conditional prima facie effects are equal to the causal $(Z{=}z)$-conditional total effects.

Another conclusion is that the expectation of the conditional total effects is equal to the causal average total effect [see Eq. (6.47)], that is,

$$\begin{aligned} ATE_{10} \;=\; E\big(PFE_{Z;10}(Z)\big) \;&=\; PFE_{Z;10}(m)\cdot 4/6 + PFE_{Z;10}(f)\cdot 2/6 \\ &=\; 5\cdot 4/6 + 10\cdot 2/6 \;\approx\; 6.667. \end{aligned}$$

Finally, note that, if $Y \vDash D_X|(X,Z)$ holds, then the individual treatment probabilities $P(X{=}1\,|\,U{=}u)$, which played a crucial role in our examples for independence and Z-conditional independence of X and D_X (see ch. 8) do not matter any more in the following sense: Even though, in this example, the individual treatment probabilities $P(X{=}1\,|\,U{=}u)$ differ between males and also between females, the $(Z{=}z)$-conditional prima facie effects $PFE_{Z;10}(m)$ and $PFE_{Z;10}(f)$ are unbiased.

9.4 Methodological Consequences

Now we discuss the implications of the causality conditions $Y \perp\!\!\!\perp D_X|X$, $Y \perp\!\!\!\perp D_X|(X,Z)$, $Y \vDash D_X|X$, and $Y \vDash D_X|(X,Z)$ for the design and analysis of experiments and quasi-experiments.

9.4.1 Methodological Consequences of $Y \perp\!\!\!\perp D_X|X$ *and* $Y \vDash D_X|X$

First of all, we have to realize that $Y \vDash D_X|X$ has no implications on any design technique. Either the empirical phenomenon we study (i. e., the random experiment that we consider in a specific empirical application) is such that $Y \perp\!\!\!\perp D_X|X$ or $Y \vDash D_X|X$ holds, or it is such that none of them holds. In contrast to the Fisher condition $D_X \perp\!\!\!\perp X$ treated in chapter 8, which can be created by randomized assignment of the observational unit to one of the treatment conditions, there is nothing we can do that implies $Y \perp\!\!\!\perp D_X|X$ or $Y \vDash D_X|X$.

In data analysis, however, we can test the hypothesis $Y \vDash D_X|X$, which implies

$$E(Y\,|\,X,W) \underset{P}{=} E(Y\,|\,X)\,. \tag{9.61}$$

provided that W is a potential confounder of X. Hence, if this equation is rejected, then we also reject the hypothesis that $Y \vDash D_X|X$ holds. Furthermore, because $Y \perp\!\!\!\perp D_X|X$ implies $Y \vDash D_X|X$, if Equation (9.61) does not hold, then $Y \perp\!\!\!\perp D_X|X$ cannot hold as well. For a test of the hypothesis that Equation (9.61) holds, standard procedures of regression analysis can be used.

9.4.2 Methodological Consequences of $Y \perp\!\!\!\perp D_X|(X,Z)$ *and* $Y \vDash D_X|(X,Z)$

In contrast to $Y \vDash D_X|X$, we can use $Y \vDash D_X|(X,Z)$ for *covariate selection* in quasi-experiments aiming at the analysis of causal conditional and average total effects. By definition, there is no randomization in a quasi-experiment. However, even under initial randomization, systematic attrition of subjects may invalidate the Fisher condition $X \perp\!\!\!\perp D_X$ (see,

e. g., Abraham & Russell, 2004; Fichman & Cummings, 2003; Graham & Donaldson, 1993; Shadish et al., 2002). In this case, we will say that randomization failed and that the initially randomized experiment turned into a quasi-experiment.

In quasi-experiments, selecting the random variables Z_i in the m-variate covariate $Z := (Z_1, \ldots, Z_m)$ for which we can hope that $Y \vDash D_X|(X, Z)$ holds, is a useful strategy in the analysis of causal conditional and average total treatment effects. As compared to the causality condition $X \perp\!\!\!\perp D_X | Z$, it might be easier to find covariates such that $Y \vDash D_X|(X, Z)$ holds. If, for example, X is a treatment variable, in many cases a pretest (i. e., a pre-treatment measure of the criterion variable) may already suffice.

Remember that $X \perp\!\!\!\perp D_X | Z$ is equivalent to

$$\forall\, x \in X(\Omega)\colon\; P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x \,|\, Z).$$

It might be difficult to find a (possibly multivariate) covariate of X such that this equation holds because there usually are many covariates determining the treatment probabilities. For instance, the *severity of the disorder*, *knowing about the treatment*, and *availability of the treatment* are candidates for such covariates. In contrast, often only $Z :=$ *pre-treatment severity of the disorder* (the pretest) may be important as a covariate to satisfy $Y \vDash D_X|(X, Z)$ if the outcome variable Y is *post-treatment severity of the disorder*. However, there is no guarantee that the pretest is sufficient as a covariate for $Y \vDash D_X|(X, Z)$ to hold. Whether or not it holds depends on the application considered.

As already mentioned before, in contrast to some of the other causality conditions treated in chapters 6 and 7, the Suppes-Reichenbach conditions can be tested in empirical applications, at least in the sense that some consequences of these assumptions can be checked. Falsifiability is important, because otherwise we would not have any criterion for covariate selection, that is, for deciding whether or not a specific covariate should be included in the m-variate covariate $Z := (Z_1, \ldots, Z_m)$ for which we hope that $Y \vDash D_X|(X, Z)$ holds.

It is easy to test $Y \vDash D_X|(X, Z)$. For example, if W is a potential confounder of X, then we can test the hypothesis

$$E(Y \,|\, X, Z, W) \underset{P}{=} E(Y \,|\, X, Z), \tag{9.62}$$

which follows from $Y \vDash D_X|(X, Z)$ [see Th. 9.21 for details]. If we reject this hypothesis, then we also reject the causality condition $Y \vDash D_X|(X, Z)$. And, because $Y \perp\!\!\!\perp D_X|(X, Z)$ implies $Y \vDash D_X|(X, Z)$, if Equation (9.62) does not hold, then $Y \perp\!\!\!\perp D_X|(X, Z)$ cannot hold as well. Again, we can use the well-known techniques of regression analysis for such a test (see, e. g., Aiken & West, 1996; Allen, 1997; Cohen et al., 2003; Draper & Smith, 1998; Gelman & Hill, 2007; Györfi, Kohler, Krzyzak, & Walk, 2002; von Eye & Schuster, 1998; West & Aiken, 2005).

9.5 Summary and Conclusions

In chapter 6 we introduced unbiasedness of the conditional expectation values $E(Y \,|\, X{=}x)$ and $E(Y \,|\, X{=}x, Z{=}z)$, as well as unbiasedness of the conditional expectations $E(Y \,|\, X)$, $E^{X=x}(Y \,|\, Z)$ and $E(Y \,|\, X, Z)$. There we showed that unbiasedness of these terms is crucial for computing causal conditional and average total effects. In chapter 7, we treated the Rosenbaum-Rubin conditions that imply unbiasedness of those conditional expectation values and conditional expectations. However, the Rosenbaum-Rubin conditions as well as the unbiasedness conditions themselves cannot be tested empirically.

In chapter 8, we introduced the Fisher conditions, a first class of empirically falsifiable causality conditions that focus on independence and Z-conditional independence of the putative cause variable X and a global potential confounder D_X of X. There we showed, for example, that $X \perp\!\!\!\perp D_X \,|\, Z$ implies Rosenbaum and Rubin's *strong ignorability*, that is, it implies $\tau \perp\!\!\!\perp X \,|\, Z$, where $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$ consists of the $J+1$ true outcome variables τ_x and Z is a covariate of X. The Fisher conditions share the focus on (conditional) independence of X and all potential confounders of X.

In contrast, the causality conditions introduced in this chapter, share the focus on conditional independence or conditional mean-independence of the outcome variable Y and all potential confounders of X. These conditions have been referred to as the *Suppes-Reichenbach conditions*. Box 9.1 displays their definitions and notation. We also studied the implication relations among the Suppes-Reichenbach conditions and showed that they imply the Rosenbaum-Rubin conditions. Table 9.2 displays the details. Combining these results with those already displayed in Tables 7.1 and 7.2 reveals the far reaching implications of the Suppes-Reichenbach conditions. Unlike unbiasedness and the Rosenbaum-Rubin conditions, and just like the Fisher conditions, the Suppes-Reichenbach conditions are empirically falsifiable. Therefore, their most important methodological implication is that they provide a viable alternative to the Fisher conditions for covariate selection aiming at creating unbiasedness. In contrast to $D_X \perp\!\!\!\perp X \,|\, Z$, which can be used to select the covariate $Z = (Z_1, \ldots, Z_m)$ such that X and the global potential confounder D_X are Z-conditionally independent, we can use $Y \vDash D_X \,|\, (X, Z)$ to select Z such that Y is (X, Z)-conditionally mean-independent from D_X.

Last but not least it should be emphasized that, just like the Fisher conditions, the Suppes-Reichenbach conditions are *generalizable*. Hence, if W is a potential confounder of X and $Y \perp\!\!\!\perp D_X \,|\, X$ holds, then $Y \perp\!\!\!\perp D_X \,|\, (X, W)$ holds as well. Furthermore, if $Y \perp\!\!\!\perp D_X \,|\, (X, Z)$ holds, then $Y \perp\!\!\!\perp D_X \,|\, (X, Z, W)$ holds as well. The corresponding mean-independence conditions $Y \vDash D_X \,|\, X$ and $Y \vDash D_X \,|\, (X, Z)$ have the same properties (see sect. 9.2 for the relevant theorems and their methodological implications).

9.6 Proofs

Proof of Theorem 9.7

$$
\begin{aligned}
& Y \vDash D_X \,|\, X \\
\Leftrightarrow \quad & E(Y \,|\, X, D_X) \underset{P}{=} E(Y \,|\, X) && \text{[Box 9.1 (iv)]} \\
\Rightarrow \quad & \forall\, x \in X(\Omega)\colon E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y). && \text{[SN-(14.82)]}
\end{aligned}
$$

Vice versa,

$$
\begin{aligned}
& \forall\, x \in X(\Omega)\colon\; E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) \\
\Rightarrow \quad & E(Y \,|\, X, D_X) \underset{P}{=} \sum_{x \in X(\Omega)} E^{X=x}(Y) \cdot 1_{X=x} && \text{[RS-(5.36)]}
\end{aligned}
$$

$$
\begin{aligned}
&\Rightarrow \quad E(Y|X, D_X) \underset{P}{=} E(Y|X) && \text{[RS-(4.1), RS-(3.24)]}\\
&\Leftrightarrow \quad Y \vDash D_X|X. && \text{[Box 9.1 (iv)]}
\end{aligned}
$$

Proof of Theorem 9.14

$$
\begin{aligned}
&\quad Y \vDash D_X|(X,Z)\\
&\Leftrightarrow \quad E(Y|X, D_X, Z) \underset{P}{=} E(Y|X,Z) && \text{[Box 9.1 (viii)]}\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X, Z) \underset{P^{X=x}}{=} E^{X=x}(Y|Z). && \text{[SN-(14.81)]}
\end{aligned}
$$

Vice versa,

$$
\begin{aligned}
&\quad \forall\, x \in X(\Omega)\colon\; E^{X=x}(Y|D_X, Z) \underset{P^{X=x}}{=} E^{X=x}(Y|Z)\\
&\Rightarrow \quad E(Y|X, D_X, Z) \underset{P}{=} \sum_{x\in X(\Omega)} E^{X=x}(Y|Z)\cdot 1_{X=x} && \text{[RS-(5.36)]}\\
&\Rightarrow \quad E(Y|X, D_X, Z) \underset{P}{=} E(Y|X,Z) && \text{[RS-(5.36)]}\\
&\Leftrightarrow \quad Y \vDash D_X|(X,Z). && \text{[Box 9.1 (viii)]}
\end{aligned}
$$

Proof of Theorem 9.17

Proposition (9.23).

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X|(X{=}x) \quad &\Leftrightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X && \text{[(9.1)]}\\
&\Rightarrow \quad E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && \text{[RS-Box 4.1 (v)]}\\
&\Leftrightarrow \quad Y \vDash D_X|(X{=}x). && \text{[Box 9.1 (ii)]}
\end{aligned}
$$

Proposition (9.24).

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X|(X{=}x, Z) \quad &\Leftrightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X|Z && \text{[(9.9)]}\\
&\Rightarrow \quad E^{X=x}(Y|D_X) \underset{P^{X=x}}{\vDash} E^{X=x}(Y|Z) && \text{[RS-Box 4.1 (v)]}\\
&\Leftrightarrow \quad Y \vDash D_X|(X{=}x, Z). && \text{[Box 9.1 (vi)]}
\end{aligned}
$$

Proof of Theorem 9.18

Proposition (9.25). This is a special case of Proposition (9.26) for $\sigma(Z) = \{\Omega, \emptyset\}$.
Proposition (9.26).

$$\begin{aligned} Y \vDash D_X | (X, Z) \quad &\Leftrightarrow \quad E(Y | X, D_X, Z) \underset{P}{=} E(Y | X, Z) && \text{[Box 9.1 (viii)]} \\ &\Rightarrow \quad E^{X=x}(Y | D_X, Z) \underset{P^{X=x}}{=} E^{X=x}(Y | Z) && \text{[SN-(14.81)]} \\ &\Leftrightarrow \quad Y \vDash D_X | (X{=}x, Z). && \text{[Box 9.1 (vi)]} \end{aligned}$$

Proof of Theorem 9.19

Proposition (9.27) . This proposition is a special case of (9.28) for $\sigma(Z) = \{\Omega, \emptyset\}$.

Proposition (9.28). By definition of a potential confounder of X, $\sigma(W) \subset \sigma(D_X)$, and this implies $\sigma(D_X, W) = \sigma(D_X)$ [see RS-Prop. (2.19)]. Therefore,

$$\begin{aligned} Y \perp\!\!\!\perp D_X | (X, Z) \quad &\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon Y \perp\!\!\!\perp (D_X, W) \,|\, (X, Z) && [\sigma(D_X, W) = \sigma(D_X)] \\ &\Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon Y \perp\!\!\!\perp D_X | (X, Z, W). && \text{[RS-Box 6.1 (viii)]} \end{aligned}$$

Proof of Theorem 9.21

Proposition (9.29). This is a special case of Proposition (9.30) for $\sigma(Z) = \{\Omega, \emptyset\}$.

Proposition (9.30).

$$\begin{aligned} & Y \vDash D_X | (X, Z) \\ \Leftrightarrow\quad & E(Y | X, Z, D_X) \underset{P}{=} E(Y | X, Z) && \text{[Box 9.1 (viii)]} \\ \Leftrightarrow\quad & \forall\, W \in \mathscr{W}_X\colon E(Y | X, Z, D_X) \underset{P}{=} E\big(E(Y | X, Z) \,\big|\, X, Z, W\big) \\ & && \big[\sigma\big(E(Y | X, Z)\big) \subset \sigma(X, Z, W), \text{ RS-Box 4.1 (xi)}\big] \\ \Leftrightarrow\quad & \forall\, W \in \mathscr{W}_X\colon E(Y | X, Z, D_X) \underset{P}{=} E\big(E(Y | X, Z, D_X) \,\big|\, X, Z, W\big) \\ & && [Y \vDash D_X | (X, Z), \text{ Box 9.1 (viii)}] \\ \Leftrightarrow\quad & \forall\, W \in \mathscr{W}_X\colon E(Y | X, Z, D_X) \underset{P}{=} E(Y | X, Z, W) \\ & && [\sigma(X, Z, W) \subset \sigma(X, Z, D_X), \text{ RS-Box 4.1 (xiii)}] \\ \Leftrightarrow\quad & \forall\, W \in \mathscr{W}_X\colon Y \vDash D_X | (X, Z, W). && \text{[Box 9.1 (viii)]} \end{aligned}$$

Proof of Theorem 9.22

Proposition (9.31). This is a special case of Proposition (9.32) for $\sigma(Z) = \{\Omega, \emptyset\}$.

Proposition (9.32).

$$\begin{aligned} & Y \vDash D_X | (X{=}x, Z) \\ \Leftrightarrow\quad & E^{X=x}(Y | Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y | Z) && \text{[Box 9.1 (vi)]} \\ \Leftrightarrow\quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y | Z, D_X) \underset{P^{X=x}}{=} E^{X=x}\big(E^{X=x}(Y | Z) \,\big|\, Z, W\big) \\ & && \big[\sigma\big(E^{X=x}(Y | Z)\big) \subset \sigma(Z, W), \text{ RS-Box 4.1 (xi)}\big] \end{aligned}$$

$$
\begin{aligned}
&\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y \,|\, Z, D_X) \underset{P^{X=x}}{=} E^{X=x}\big(E^{X=x}(Y \,|\, Z, D_X) \,\big|\, Z, W\big) \\
&\hspace{16em} [Y \vDash D_X \,|\, (X{=}x, Z), \text{ Box 9.1 (vi)}] \\
&\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y \,|\, Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y \,|\, Z, W) \\
&\hspace{16em} [\sigma(Z, W) \subset \sigma(Z, D_X), \text{ RS-Box 4.1 (xiii)}] \\
&\Leftrightarrow \quad \forall\, W \in \mathscr{W}_X\colon Y \vDash D_X \,|\, (X{=}x, Z, W). \hspace{6em} [\text{Box 9.1 (vi)}]
\end{aligned}
$$

Proof of Theorem 9.25

$$
\begin{aligned}
Y \vDash D_X \,|\, (X{=}x) \quad &\Leftrightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [\text{Box 9.1 (ii)}] \\
&\Rightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P}{=} E^{X=x}(Y) && [\tau_x = E^{X=x}(Y \,|\, D_X) \text{ is } P\text{-unique}] \\
&\Leftrightarrow \quad \tau_x \underset{P}{=} E^{X=x}(Y) && [\tau_x \underset{P}{=} E^{X=x}(Y \,|\, D_X)] \\
&\Rightarrow \quad \tau_x \perp\!\!\!\perp X. && [\text{RS-Box 2.1 (iii)}]
\end{aligned}
$$

Proof of Theorem 9.28

First of all note that $\sigma\big(E^{X=x}(Y \,|\, Z)\big) \subset \sigma(Z)$. According to RS-Box 6.1 (iv), this implies

$$
E^{X=x}(Y \,|\, Z) \perp\!\!\!\perp X \,|\, Z.
$$

Hence,

$$
\begin{aligned}
Y \vDash D_X \,|\, (X{=}x) \quad &\Leftrightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y \,|\, Z) && [\text{Box 9.1 (ii)}] \\
&\Leftrightarrow \quad \tau_x \underset{P}{=} E^{X=x}(Y \,|\, Z) && [\tau_x = E^{X=x}(Y \,|\, D_X),\ \tau_x \text{ is } P\text{-unique}] \\
&\Rightarrow \quad \tau_x \perp\!\!\!\perp X \,|\, Z. && [\text{RS-Box 6.1 (x)},\ E^{X=x}(Y \,|\, Z) \perp\!\!\!\perp X \,|\, Z]
\end{aligned}
$$

Proof of Theorem 9.30

Proposition (9.37).

$$
\begin{aligned}
Y \vDash D_X \,|\, X \quad &\Rightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [\text{Th. 9.7}] \\
&\Leftrightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P}{=} E^{X=x}(Y) && [\tau_x = E^{X=x}(Y \,|\, D_X) \text{ is } P\text{-unique}] \\
&\Leftrightarrow \quad \tau_x \underset{P}{=} E^{X=x}(Y) && [\tau_x = E^{X=x}(Y \,|\, D_X)] \\
&\Rightarrow \quad \tau_x \perp\!\!\!\perp X. && [\text{RS-Box 2.1 (iii)}]
\end{aligned}
$$

Proposition (9.38).

$$\begin{aligned} Y \vDash D_X|X \quad &\Rightarrow \quad \tau_x \underset{P}{=} E^{X=x}(Y) && \text{[see proof of (9.37)]} \\ &\Rightarrow \quad \tau_x \perp\!\!\!\perp X\,|\,Z. && \text{[RS-Box 6.1 (v)]} \end{aligned}$$

Proof of Theorem 9.32

Proposition (9.41).

$$\begin{aligned} & Y \vDash D_X|X \\ \Leftrightarrow\quad & \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && \text{[Th. 9.7]} \\ \Leftrightarrow\quad & \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X) \underset{P}{=} E^{X=x}(Y) && \text{[all } \tau_x = E^{X=x}(Y|D_X) \text{ are } P\text{-unique]} \\ \Leftrightarrow\quad & \forall\, x \in X(\Omega)\colon \tau_x \underset{P}{=} E^{X=x}(Y) && [\tau_x = E^{X=x}(Y|D_X)] \\ \Leftrightarrow\quad & \tau \underset{P}{=} \big(E^{X=0}(Y), E^{X=1}(Y), \dots, E^{X=J}(Y)\big) && [\tau = (\tau_0, \tau_1, \dots, \tau_J)] \\ \Rightarrow\quad & \tau \perp\!\!\!\perp X. && \text{[RS-Box 2.1 (iii)]} \end{aligned}$$

Proposition (9.42).

$$\begin{aligned} Y \vDash D_X|X \quad &\Leftrightarrow \quad \tau \underset{P}{=} \big(E^{X=0}(Y), E^{X=1}(Y), \dots, E^{X=J}(Y)\big) && \text{[see proof of (9.41)]} \\ &\Rightarrow \quad \tau \perp\!\!\!\perp X\,|\,Z. && \text{[RS-Box 6.1 (v)]} \end{aligned}$$

Proof of Theorem 9.35

Note that $\sigma\big(E^{X=x}(Y|Z)\big) \subset \sigma(Z)$. According to RS-Box 6.1 (iv), this implies

$$E^{X=x}(Y|Z) \perp\!\!\!\perp X\,|\,Z.$$

Hence,

$$\begin{aligned} Y \vDash D_X|(X{=}x, Z) \quad &\Leftrightarrow \quad E^{X=x}(Y|Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) && \text{[Box 9.1 (vi)]} \\ &\Leftrightarrow \quad E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) && [\sigma(Z) \subset \sigma(D_X)] \\ &\Leftrightarrow \quad \tau_x \underset{P}{=} E^{X=x}(Y|Z) && [\tau_x = E^{X=x}(Y|D_X),\ \tau_x \text{ is } P\text{-unique}] \\ &\Rightarrow \quad \tau_x \perp\!\!\!\perp X\,|\,Z. && [E^{X=x}(Y|Z) \perp\!\!\!\perp X\,|\,Z] \end{aligned}$$

Proof of Theorem 9.37

$$\begin{aligned} & Y \vDash D_X|(X{=}x, Z) \\ \Leftrightarrow\quad & E^{X=x}(Y|Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) && \text{[Box 9.1 (vi)]} \end{aligned}$$

$$
\begin{aligned}
&\Leftrightarrow \quad E^{X=x}(Y\,|\,D_X) \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) && [\sigma(Z) \subset \sigma(D_X)]\\
&\Leftrightarrow \quad \tau_x \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) && [\tau_x = E^{X=x}(Y\,|\,D_X)]\\
&\Leftrightarrow \quad \tau_x \underset{P^{Z=z}}{=} E^{X=x}(Y\,|\,Z) && [E^{X=x}(Y\,|\,Z) \in \mathscr{E}^{X=x}(Y\,|\,D_X),\ \tau_x \text{ is } P^{Z=z}\text{-unique}]\\
&\Rightarrow \quad E^{Z=z}(\tau_x) = E^{Z=z}\big(E^{X=x}(Y\,|\,Z)\big) && [P(Z=z) > 0, \text{ RS-Box 3.1 (v)}]\\
&\Leftrightarrow \quad E(\tau_x|Z=z) = E^{X=x}(Y|Z=z)\,. && [\text{RS-(3.24), RS-(5.24), SN-(14.22)}]
\end{aligned}
$$

Proof of Theorem 9.38

Proposition (9.49).

$$
\begin{aligned}
& Y \vDash D_X | (X, Z)\\
\Leftrightarrow \quad & E(Y\,|\,X, Z, D_X) \underset{P}{=} E(Y\,|\,X, Z) && [\text{Box 9.1 (viii)}]\\
\Leftrightarrow \quad & \forall\, x \in X(\Omega){:}\ \ E^{X=x}(Y\,|\,Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) && [\text{Th. 9.14}]\\
\Leftrightarrow \quad & \forall\, x \in X(\Omega){:}\ \ E^{X=x}(Y\,|\,D_X) \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) && [\sigma(Z) \subset \sigma(D_X)]\\
\Leftrightarrow \quad & \forall\, x \in X(\Omega){:}\ \ \tau_x \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) && [\forall\, x \in X(\Omega){:}\ \tau_x = E^{X=x}(Y\,|\,D_X)]\\
\Leftrightarrow \quad & \forall\, x \in X(\Omega){:}\ \ \tau_x \underset{P}{=} E^{X=x}(Y\,|\,Z) && [\text{all } \tau_x \text{ are } P\text{-unique}]\\
\Leftrightarrow \quad & \tau \underset{P}{=} \big(E^{X=0}(Y\,|\,Z), E^{X=1}(Y\,|\,Z), \dots, E^{X=J}(Y\,|\,Z)\big) && [\tau = (\tau_0, \tau_1, \dots, \tau_J)]\\
\Rightarrow \quad & \tau \perp\!\!\!\perp X\,|\,Z. && [\text{RS-Box 6.1 (iv), (x)}]
\end{aligned}
$$

Proposition (9.50).

$$
\begin{aligned}
Y \vDash D_X | (X, Z) \quad &\Rightarrow \quad \tau \perp\!\!\!\perp X\,|\,Z && [9.49]\\
&\Rightarrow \quad \tau \perp\!\!\!\perp 1_{X=x}|Z. && [\sigma(1_{X=x}) \subset \sigma(X), \text{ RS-6.1 (vi)}]
\end{aligned}
$$

Proof of Theorem 9.40

$$
\begin{aligned}
Y \vDash D_X | (X, Z) \quad &\Leftrightarrow \quad E(Y\,|\,X, Z, D_X) \underset{P}{=} E(Y\,|\,X, Z) && [\text{Box 9.1 (viii)}]\\
&\Rightarrow \quad E^{X=x}(Y\,|\,Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) && [\text{SN-(14.81)}]\\
&\Leftrightarrow \quad E^{X=x}(Y\,|\,D_X) \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) && [\sigma(Z, D_X) = \sigma(D_X)]\\
&\Leftrightarrow \quad Y \vDash D_X | (X{=}x, Z) && [\text{Box 9.1 (vi)}]\\
&\Rightarrow \quad E^{X=x}(Y|Z{=}z) \vdash D_X\,. && [\text{Th. 9.37}]
\end{aligned}
$$

Proof of Theorem 9.42

$$
\begin{array}{lll}
& Y \vDash D_X | (X,Z) & \\
\Rightarrow & \tau \perp\!\!\!\perp X \,|\, Z & \text{[Th. 9.38]} \\
\Rightarrow & \forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} | Z & \text{[(7.57), (7.59)]} \\
\Rightarrow & \forall\, x \in X(\Omega)\colon E^{X=x}(Y\,|\,Z) \vdash D_X & \text{[(7.62)]} \\
\Rightarrow & \forall\, x \in X(\Omega)\colon E^{X=x}(Y\,|\,Z) \underset{P}{=} E(\tau_x | Z) & \text{[Def. 6.13 (i)]} \\
\Rightarrow & PFE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x|Z) - E(\tau_{x'}|Z) & \text{[(9.54)]} \\
\Rightarrow & PFE_{Z;xx'}(Z) \underset{P}{=} CTE_{Z;xx'}(Z). & \text{[(9.55)]}
\end{array}
$$

Proof of Theorem 9.43

$$
\begin{array}{lll}
& Y \vDash D_X | (X,Z) & \\
\Rightarrow & E^{X=x}(Y\,|\,Z{=}z) \vdash D_X \;\wedge\; E^{X=x'}(Y\,|\,Z{=}z) \vdash D_X & \text{[(9.53)]} \\
\Rightarrow & E^{X=x}(Y\,|\,Z{=}z) = E(\tau_x|Z{=}z) \;\wedge\; E^{X=x'}(Y\,|\,Z{=}z) = E(\tau_{x'}|Z{=}z) & \text{[Def. 6.13 (ii)]} \\
\Rightarrow & PFE_{Z;xx'}(z) = CTE_{Z;xx'}(z). & \text{[(9.57), (9.58)]}
\end{array}
$$

9.7 Exercises

▷ **Exercise 9-1** Assuming $P(X{=}x) > 0$, show that $E(Y) < \infty$ implies $E^{X=x}(Y) < \infty$.

▷ **Exercise 9-2** Show that $Y \vDash D_X | (X,Z)$ implies $E(Y\,|\,X,Z,W) \underset{P}{=} E(Y\,|\,X,Z)$ if Z and W are measurable with respect to the global potential confounder D_X.

▷ **Exercise 9-3** Let the Assumptions 9.1 (a) to (e) and (j) hold. Which unbiasedness conditions follow from $Y \vDash D_X | X$?

▷ **Exercise 9-4** Let the Assumptions 9.1 (a) to (e), (g), and (j) hold. Which unbiasedness conditions follow from $Y \vDash D_X | (X,Z)$?

▷ **Exercise 9-5** Show that Proposition (9.11) holds if for all values $z \in Z(\Omega)$: $P^{X=x}(Z{=}z) > 0$.

▷ **Exercise 9-6** Check if and where the implications listed in Table 9.2 have been proven in this chapter and prove those that have not. Use and specify the appropriate choice of assumptions listed in the Assumptions 9.1.

▷ **Exercise 9-7** Check if and where the implications listed in Table 9.3 have been proven in this chapter and prove those that have not. Use and specify the appropriate choice of assumptions listed in the Assumptions 9.1.

Solutions

▷ **Solution 9-1** According to SN-Equation (9.11),

$$E^{X=x}(Y) \;=\; \frac{1}{P(X=x)} \cdot E(1_{X=x} \cdot Y).$$

According to SN-Remark 3.34 and RS-Definition 3.1, $E(Y) < \infty$ implies $E(1_{X=x} \cdot Y) < \infty$. Hence, the assumptions $P(X=x) > 0$ and $E(Y) < \infty$ imply $E^{X=x}(Y) < \infty$.

▷ **Solution 9-2**

$$\begin{aligned}
E(Y|X,Z,W) &\underset{P}{=} E\big(E(Y|X,Z,D_X)\,\big|\,X,Z,W\big) && [\sigma(Z,W) \subset \sigma(D_X), \text{RS-Box 6.1 (vi)}]\\
&\underset{P}{=} E\big(E(Y|X,Z)\,\big|\,X,Z,W\big) && [Y \vDash D_X|(X,Z)]\\
&\underset{P}{=} E(Y|X,Z). && [\text{RS-Box 4.1 (xi)}]
\end{aligned}$$

▷ **Solution 9-3** Under the assumptions mentioned in the exercise, $Y \vDash D_X|X$ implies that $E(Y|X)$, all its values $E(Y|X=x)$, and all prima facie effects $PFE_{xx'}$, $x, x' \in X(\Omega)$, are unbiased. If we additionally consider a covariate Z of X, then $Y \vDash D_X|X$ also implies that $E(Y|X,Z)$ and all $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, are unbiased (see Rem. 9.34).

▷ **Solution 9-4** Under the assumptions of this exercise, $Y \vDash D_X|(X,Z)$ implies that $E(Y|X,Z)$ and all conditional expectations $E^{X=x}(Y|Z)$, $x = 0, 1, \dots, J$, are unbiased. Furthermore, if z is a value of Z satisfying $P(X=x, Z=z) > 0$, then $Y \vDash D_X|(X,Z)$ also implies that $E(Y|X=x, Z=z)$ is unbiased.

▷ **Solution 9-5**

$$\begin{aligned}
& Y \vDash D_X|(X=x, Z)\\
\Leftrightarrow\quad & E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z)\\
\Leftrightarrow\quad & E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} \sum_{z \in Z(\Omega)} E^{X=x}(Y|Z=z) \cdot 1_{Z=z} && [\text{RS-Def. 4.1, RS-Rem. 4.8}]\\
\Leftrightarrow\quad & \exists A \in \mathscr{A}\colon \Big(\forall \omega \in \Omega \setminus A\colon\; E^{X=x}(Y|D_X)(\omega) = \sum_{z \in Z(\Omega)} E^{X=x}(Y|Z=z) \cdot 1_{Z=z}(\omega)\\
& \qquad \wedge\; P^{X=x}(A) = 0\Big) && [\text{RS-Def. 2.46}]\\
\Leftrightarrow\quad & \forall\, \omega \in \Omega\colon\; E^{X=x}(Y|D_X)(\omega) = E^{X=x}(Y|Z=z), \quad \text{if } \omega \in \{Z=z\}. && [P^{X=x}(Z=z) > 0]
\end{aligned}$$

▷ **Solution 9-6** We check the implications listed in Table 9.2 row-wise.

(1) Under Assumptions 9.1 (a) to (c) and (g): $Y \vDash D_X|(X,Z) \;\Rightarrow\; Y \vDash D_X|(X=x, Z)$.
This is Proposition (9.26).

(2) Under Assumptions 9.1 (a) to (c) and (h): $Y \vDash D_X|(X=x) \;\Rightarrow\; Y \vDash D_X|(X=x, Z)$.
This immediately follows from Proposition (9.31) because $\sigma(Z) \subset \sigma(D_X)$.

(3) Under Assumptions 9.1 (a) to (c) and (h): $Y \vDash D_X|X \;\Rightarrow\; Y \vDash D_X|(X=x, Z)$.

$$\begin{aligned}
Y \vDash D_X|X \quad &\Rightarrow\quad Y \vDash D_X|(X,Z) && [\text{Prop. (9.29)}]\\
&\Rightarrow\quad Y \vDash D_X|(X=x, Z). && [(1)]
\end{aligned}$$

(4) Under Assumptions 9.1 (a) to (c) and (h): $Y \vDash D_X|X \;\Rightarrow\; Y \vDash D_X|(X,Z)$.
This immediately follows from Proposition (9.29) because $\sigma(Z) \subset \sigma(D_X)$.

(5) Under Assumptions 9.1 (a) to (c): $Y \vDash D_X|X \;\Rightarrow\; Y \vDash D_X|(X=x)$.
This is Proposition (9.25).

(6) Under Assumptions 9.1 (a) to (c) and (g): $Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z) \;\Rightarrow\; Y \vDash D_X \,|\, (X{=}x, Z)$.
This is Proposition (9.24).

(7) Under Assumptions 9.1 (a) to (c) and (g): $Y \perp\!\!\!\perp D_X \,|\, (X, Z) \;\Rightarrow\; Y \vDash D_X \,|\, (X{=}x, Z)$.

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X \,|\, (X, Z) \quad &\Rightarrow \quad Y \vDash D_X \,|\, (X, Z) && [(9.22)] \\
&\Rightarrow \quad Y \vDash D_X \,|\, (X{=}x, Z). && [(1)]
\end{aligned}
$$

(8) Under Assumptions 9.1 (a), (c), and (g): $Y \perp\!\!\!\perp D_X \,|\, (X, Z) \;\Rightarrow\; Y \vDash D_X \,|\, (X, Z)$.
This is Proposition (9.22).

(9) Under Assumptions 9.1 (a) to (c) and (g): $Y \perp\!\!\!\perp D_X \,|\, (X, Z) \;\Rightarrow\; Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z)$.

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X \,|\, (X, Z) \quad &\Rightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X \,|\, Z && [\text{RS-Th. } 6.24] \\
&\Leftrightarrow \quad Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z). && [(9.8)]
\end{aligned}
$$

(10) Under Assumptions 9.1 (a) to (c) and (h): $Y \perp\!\!\!\perp D_X \,|\, (X{=}x) \;\Rightarrow\; Y \vDash D_X \,|\, (X{=}x, Z)$.
This immediately follows from Proposition (9.33) because $\sigma(Z) \subset \sigma(D_X)$.

(11) Under Assumptions 9.1 (a) to (c): $Y \perp\!\!\!\perp D_X \,|\, (X{=}x) \;\Rightarrow\; Y \vDash D_X \,|\, (X{=}x)$.
This is Proposition (9.23).

(12) Under Assumptions 9.1 (a), (b), and (h): $Y \perp\!\!\!\perp D_X \,|\, (X{=}x) \;\Rightarrow\; Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z)$.

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X \,|\, (X{=}x) \quad &\Leftrightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X && [(9.1)] \\
&\Rightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X \,|\, Z && [\sigma(Z) \subset \sigma(D_X), \text{ RS-Prop. } (6.5)] \\
&\Leftrightarrow \quad Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z). && [(9.8)]
\end{aligned}
$$

(13) Under Assumptions 9.1 (a) to (c) and (h): $Y \perp\!\!\!\perp D_X \,|\, X \;\Rightarrow\; Y \vDash D_X \,|\, (X{=}x, Z)$.
This immmediately follows from Proposition (9.33) because $\sigma(Z) \subset \sigma(D_X)$.

(14) Under Assumptions 9.1 (a), (c), and (h): $Y \perp\!\!\!\perp D_X \,|\, X \;\Rightarrow\; Y \vDash D_X \,|\, (X, Z)$.

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X \,|\, X \quad &\Rightarrow \quad Y \vDash D_X \,|\, X && [(9.21)] \\
&\Rightarrow \quad Y \vDash D_X \,|\, (X, Z). && [\sigma(Z) \subset \sigma(D_X), (9.29)]
\end{aligned}
$$

(15) Under Assumptions 9.1 (a) to (c): $Y \perp\!\!\!\perp D_X \,|\, X \;\Rightarrow\; Y \vDash D_X \,|\, (X{=}x)$.

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X \,|\, X \quad &\Rightarrow \quad Y \vDash D_X \,|\, X && [(9.21)] \\
&\Rightarrow \quad Y \vDash D_X \,|\, (X{=}x). && [(5)]
\end{aligned}
$$

(16) Under Assumptions 9.1 (a) and (c): $Y \perp\!\!\!\perp D_X \,|\, X \;\Rightarrow\; Y \vDash D_X \,|\, X$.
This is Proposition (9.21).

(17) Under Assumptions 9.1 (a), (b), and (h): $Y \perp\!\!\!\perp D_X \,|\, X \;\Rightarrow\; Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z)$.

$$
\begin{aligned}
Y \perp\!\!\!\perp D_X \,|\, X \quad &\Leftrightarrow \quad Y \perp\!\!\!\perp (D_X, Z) \,|\, X && [\sigma(Z) \subset \sigma(D_X)] \\
&\Rightarrow \quad Y \perp\!\!\!\perp (D_X, Z) \,|\, (X, Z) && [\text{RS-Box } 6.1 \text{ (viii)}] \\
&\Leftrightarrow \quad Y \perp\!\!\!\perp D_X \,|\, (X, Z) && [\sigma(Z) \subset \sigma(D_X)] \\
&\Rightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X \,|\, Z && [P(X{=}x) > 0, \text{ RS-Th. } 6.24] \\
&\Leftrightarrow \quad Y \perp\!\!\!\perp D_X \,|\, (X{=}x, Z). && [(9.8)]
\end{aligned}
$$

(18) Under Assumptions 9.1 (a) and (h): $Y \perp\!\!\!\perp D_X | X \;\Rightarrow\; Y \perp\!\!\!\perp D_X | (X, Z)$.
This immediately follows from Proposition (9.27) because $\sigma(Z) \subset \sigma(D_X)$.

(19) Under Assumptions 9.1 (a) and (b): $Y \perp\!\!\!\perp D_X | X \;\Rightarrow\; Y \perp\!\!\!\perp D_X | (X{=}x)$.

$$\begin{aligned} Y \perp\!\!\!\perp D_X | X \quad &\Rightarrow \quad Y \underset{P^{X=x}}{\perp\!\!\!\perp} D_X && \text{[RS-Prop. (6.37)]} \\ &\Leftrightarrow \quad Y \perp\!\!\!\perp D_X | (X{=}x)\,. && \text{[(9.1)]} \end{aligned}$$

▷ **Solution 9-7** Again, we check the implications listed in Table 9.3 row-wise.

(1) Under the Assumptions 9.1 (a) to (d), (h), and (i): $Y \vDash D_X | (X{=}x, Z) \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} | Z$.
This is Proposition (9.47).

(2) Under the Assumptions 9.1 (a) to (d), (h), and (i): $Y \vDash D_X | (X{=}x, Z) \;\Rightarrow\; \tau_x \perp\!\!\!\perp X | Z$.
This is Proposition (9.45).

(3) Under the Assumptions 9.1 (a) to (d), (h), and (i): $Y \perp\!\!\!\perp D_X | (X, Z) \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} | Z$.
This is Proposition (9.52).

(4) Under the Assumptions 9.1 (a) to (d), (h), and (i): $Y \perp\!\!\!\perp D_X | (X, Z) \;\Rightarrow\; \tau_x \perp\!\!\!\perp X | Z$.
This is Proposition (9.51).

(5) Under the Assumptions 9.1 (a) to (e), (h), and (j): $Y \vDash D_X | (X, Z) \;\Rightarrow\; \tau \perp\!\!\!\perp X | Z$.
This is Proposition (9.49).

(6) Under the Assumptions 9.1 (a) to (e), (h), and (j): $Y \vDash D_X | (X, Z) \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} | Z$.
This is Proposition (9.50).

(7) Under the Assumptions 9.1 (a) to (d), (g), and (i): $Y \vDash D_X | (X{=}x) \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} | Z$.
This is Remark 9.29 (i).

(8) Under the Assumptions 9.1 (a) to (d), (g), and (i): $Y \vDash D_X | (X{=}x) \;\Rightarrow\; \tau_x \perp\!\!\!\perp X | Z$.
This is Proposition (9.36).

(9) Under the Assumptions 9.1 (a) to (d) and (i): $Y \vDash D_X | (X{=}x) \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x}$.

$$\begin{aligned} Y \vDash D_X | (X{=}x) \quad &\Rightarrow \quad \tau_x \perp\!\!\!\perp X && \text{[Th. 9.25]} \\ &\Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x}\,. && [\sigma(1_{X=x}) \subset \sigma(X),\ \text{RS-Box 2.1 (iv)}] \end{aligned}$$

(10) Under the Assumptions 9.1 (a) to (d) and (i): $Y \vDash D_X | (X{=}x) \;\Rightarrow\; \tau_x \perp\!\!\!\perp X$.
This is Theorem 9.25.

(11) Under the Assumptions 9.1 (a) to (d), (g), and (i): $Y \vDash D_X | X \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} | Z$.
This is Proposition (9.40).

(12) Under the Assumptions 9.1 (a) to (d), (g), and (i): $Y \vDash D_X | X \;\Rightarrow\; \tau_x \perp\!\!\!\perp X | Z$.
This is Proposition (9.38).

(13) Under the Assumptions 9.1 (a) to (e), (g), and (j): $Y \vDash D_X | X \;\Rightarrow\; \tau \perp\!\!\!\perp X | Z$.
This is Proposition (9.42).

(14) Under the Assumptions 9.1 (a) to (e), (g), and (j): $Y \vDash D_X | X \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x} | Z$.
This is Proposition (9.44).

(15) Under the Assumptions 9.1 (a) to (d) and (i): $Y \vDash D_X | X \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x}$.
This is Proposition (9.39).

(16) Under the Assumptions 9.1 (a) to (d) and (i): $Y \vDash D_X | X \;\Rightarrow\; \tau_x \perp\!\!\!\perp X$.
This is Proposition (9.37)

(17) Under the Assumptions 9.1 (a) to (e) and (j): $Y \vDash D_X | X \;\Rightarrow\; \tau \perp\!\!\!\perp X$.
This is Proposition (9.41).

(18) Under the Assumptions 9.1 (a) to (e) and (j): $Y \vDash D_X | X \;\Rightarrow\; \tau \perp\!\!\!\perp 1_{X=x}$.
This is Proposition (9.43).

Chapter 10
Unconfoundedness

In this chapter we study another kind of causality conditions called *unconfoundedness*. To my knowledge Cartwright (1979) was the first to mention these conditions, calling them "effective strategies". In Steyer (1984, 1992) I called these properties "weak causality", and Steyer, Gabler, von Davier, and Nachtigall (2000) "unconfoundedness".

These conditions deserve special attention, because they are, to my knowledge, the weakest causality conditions that are still empirically testable, at least in the sense of falsifiability. 'Empirically testable' means that, in empirical applications, the hypothesis that an unconfoundedness condition holds for a conditional expectation value $E^{X=x}(Y)$ or the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$, or $E(Y|X,Z)$ can be tested and rejected if necessary, using statistical procedures. Hence, just like the Fisher conditions and the Suppes-Reichenbach conditions, the unconfoundedness conditions can be used as criteria for selecting relevant covariates, that is, for selecting those covariates that would induce bias if not included in the m-variate covariate $Z := (Z_1, \ldots, Z_m)$ of X when considering the conditional expectations $E^{X=x}(Y|Z)$ or $E(Y|X,Z)$ in order to describe the causal dependence of the outcome variable Y on the putative cause variable X, conditioning on the covariate Z.

In this and some other respects, the unconfoundedness conditions are similar to the Fisher conditions and the Suppes-Reichenbach conditions. Just like the F-condition $D_X \perp\!\!\!\perp X$, the corresponding unconfoundedness condition holds in a perfect randomized experiment in which X is the treatment variable and Y the outcome variable. Similarly, just like the F-condition $D_X \perp\!\!\!\perp X|Z$, the corresponding unconfoundedness condition holds in a perfect randomized experiment and in a perfect Z-conditionally randomized experiment.

We start with the basic idea and the definitions of the unconfoundedness conditions and then treat their implication structure. We continue showing that they are implied by the F-conditions and by the SR-conditions. Next we illustrate them by an example and show that the unconfoundedness conditions imply the corresponding unbiasedness conditions. Finally, we treat an important property, namely *expectation stability of causal total effects and conditional effect functions*. This property holds under unconfoundedness, but also under the F-conditions and under the SR-conditions.

In this chapter, we will often refer to the following assumptions and notation.

Notation and Assumptions 10.1

(a) *Let* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *be a regular probabilistic causality setup, let* D_X *denote a global potential confounder of* X*, and let* $\mathscr{W}_X$ *denote the set of all*

potential confounders of X, that is, the set of all random variables on $(\Omega, \mathscr{A}, P)$ that are D_X-measurable.

(b) *Let $(\Omega'_X, \mathscr{A}'_X)$ denote the value space of X, let $x \in \Omega'_X$, let $\{x\} \in \mathscr{A}'_X$, and let $1_{X=x}$ denote the indicator of the event $\{X{=}x\} = \{\omega \in \Omega\colon X(\omega) = x\}$.*

(c) *Let Y be real-valued with positive variance.*

(d) *Let $X(\Omega) = \{0, 1, \dots, J\}$ be the image of Ω under X, let $x \in X(\Omega)$, $0 < P(X{=}x) < 1$, define the probability measure $P^{X=x}\colon \mathscr{A} \to [0,1]$ by $P^{X=x}(A) = P(A \mid X{=}x)$ for all $A \in \mathscr{A}$, let $\tau_x = E^{X=x}(Y \mid D_X)$ denote a (version of the) true outcome variable of Y given x, and $\mathscr{E}^{X=x}(Y \mid D_X)$ the set of all such versions.*

(e) *Let Z be a covariate of X, that is, let $\sigma(Z) \subset \sigma(D_X)$, and let $(\Omega'_Z, \mathscr{A}'_Z)$ denote its value space.*

(f) *Let $z \in \Omega'_Z$ be a value of Z, let $\{z\} \in \mathscr{A}'_Z$, assume $0 < P(X{=}x, Z{=}z) < 1$, and define the probability measure $P^{Z=z}\colon \mathscr{A} \to [0,1]$ by $P^{Z=z}(A) = P(A \mid Z{=}z)$, for all $A \in \mathscr{A}$.*

(g) *For all $x \in X(\Omega) = \{0, 1, \dots, J\}$, let $\{x\} \in \mathscr{A}'_X$ and $0 < P(X{=}x) < 1$. Furthermore, let $\tau := (\tau_0, \tau_1, \dots, \tau_J)$ denote the $(J+1)$-variate random variable consisting of the true outcome variables $\tau_0, \tau_1, \dots, \tau_J$ of Y.*

(h) *Assume that τ_x is P-unique.*

(i) *Assume that all τ_x, $x \in X(\Omega)$, are P-unique.*

(j) *Assume that τ_x is $P^{Z=z}$-unique.*

(k) *Assume that all τ_x, $x \in X(\Omega)$, are $P^{Z=z}$-unique.*

10.1 Unconfoundedness Conditions

Table 10.1 summarizes the definitions of unconfoundedness of various conditional expectation values and conditional expectations. In this section, these definitions are treated in more detail. We start motivating the definition of the unconfoundedness condition of $E(Y \mid X)$ by comparing it to unbiasedness of the conditional expectation $E(Y \mid X)$, which is briefly reviewed in Remarks 10.2 to 10.5.

Remark 10.2 [Unbiasedness of $E(Y \mid X)$] In the definition of *unbiasedness of $E(Y \mid X)$*, denoted $E(Y \mid X) \vdash D_X$ [see Def. 6.3 (i)], we require that the Assumptions 10.1 (a) to (d) hold, including that the outcome variable Y is real-valued with positive variance. This implies that the expectation $E(Y)$ of Y is finite [see SN-Rem. 6.25 (iii) and SN-Def. 6.28]. Additionally, we require that, for all $x \in X(\Omega)$, the true outcome variable $\tau_x = E^{X=x}(Y \mid D_X)$ is P-unique [see Ass. 10.1 (i)]. Under these assumptions, the crucial condition defining unbiasedness of $E(Y \mid X)$ is

$$\forall\, x \in X(\Omega)\colon\; E^{X=x}(Y) \;=\; E(\tau_x). \tag{10.1}$$

Remember that $E(Y \mid X{=}x)$ is an alternative notation for the expectation $E^{X=x}(Y)$ of Y with respect to the probability measure $P^{X=x}$ defined in Assumption 10.1 (d). That is, if $P(X{=}x) > 0$, then

$$E(Y \mid X{=}x) \;=\; E^{X=x}(Y) \tag{10.2}$$

[see RS-Eq. (3.24)]. ◁

Remark 10.3 [P-Uniqueness of τ_x] According to RS-Box 5.1 (vii), P-uniqueness of all true outcome variables τ_x, $x \in X(\Omega)$, implies that the expectations of all τ_x are uniquely defined and, according to SN-Theorem 9.4, finiteness of $E(Y)$ and Equation (10.1) imply that all expectations $E(\tau_x)$, $x \in X(\Omega)$, are finite as well. Hence, according to RS-Theorem 5.27, P-uniqueness of all τ_x, $x \in X(\Omega)$, is equivalent to the expectations of all $\tau_x = E^{X=x}(Y \mid D_X)$ being uniquely defined. ◁

Remark 10.4 [An Implication of P-Uniqueness of τ_x] According to Rule (iv) of RS-Box 5.1, P-uniqueness of $\tau_x = E^{X=x}(Y \mid D_X)$ implies P-uniqueness of $E^{X=x}(Y \mid W)$, provided that W is D_X-measurable. Because, according to Definition 4.11 (iv), a potential confounder W of X is D_X-measurable and we assume P-uniqueness of τ_x, this implies: If W is a potential confounder of X, then the expectation $E\big(E^{X=x}(Y \mid W)\big)$ is uniquely defined. Note, however, that the expectations $E\big(E^{X=x}(Y \mid W_1)\big)$ and $E\big(E^{X=x}(Y \mid W_2)\big)$ can differ if W_1 and W_2 are different potential confounders of X. ◁

Remark 10.5 [Meaning of Unbiasedness] With Equation (10.1) we postulate that each conditional expectation value $E^{X=x}(Y)$, $x \in X(\Omega)$, is identical to the expectation of the true outcome variable $\tau_x = E^{X=x}(Y \mid D_X)$. That is, we postulate

$$\forall\, x \in X(\Omega)\colon\; E^{X=x}(Y) \;=\; E\big(E^{X=x}(Y \mid D_X)\big). \tag{10.3}$$

Of course, we may loose information and precision in approximating Y by $E^{X=x}(Y)$ given the value x of the putative cause variable X. However, if $E^{X=x}(Y)$ is unbiased, then it is identical to the expectation $E(\tau_x)$ of the true outcome variable $\tau_x = E^{X=x}(Y \mid D_X)$ in which we control for the global potential confounder D_X and with it, for *all* potential confounders of X.

If we consider a second value x' of X and $E^{X=x'}(Y)$ is unbiased as well, then $E^{X=x'}(Y)$ is identical to the expectation $E(\tau_{x'})$ of the true outcome variable $\tau_{x'} = E^{X=x'}(Y \mid D_X)$. This implies

$$E^{X=x}(Y) - E^{X=x'}(Y) \;=\; E(Y \mid X{=}x) - E(Y \mid X{=}x') \;=\; E(\tau_x - \tau_{x'}) \;=\; ATE_{xx'},$$

that is, the difference $E^{X=x}(Y) - E^{X=x'}(Y)$ is identical to the causal average total effect on Y comparing x to x' (see Def. 5.26). ◁

Remark 10.6 [Basic Idea of Unconfoundedness of $E(Y \mid X)$] The basic idea of *unconfoundedness* of $E(Y \mid X)$ is to postulate

$$\forall\, x \in X(\Omega)\colon\; E^{X=x}(Y) \;=\; E\big(E^{X=x}(Y \mid W)\big) \tag{10.4}$$

for all potential confounders W of X, that is, for *all* $W \in \mathscr{W}_X$, and not only for the *global* potential confounder D_X, as in the definition of unbiasedness [see Eq. (10.3)]. In other words, we postulate that each conditional expectation value $E^{X=x}(Y)$ is identical to the expectation (with respect to the probability measure P) of the conditional expectation $E^{X=x}(Y \mid W)$ *for each* potential confounder W of X, and not only for the *global* potential confounder D_X of X. In contrast to

$$E^{X=x}(Y) \;=\; E^{X=x}\big(E^{X=x}(Y \mid W)\big), \tag{10.5}$$

Equation (10.4) does not always hold even if $P(X{=}x) > 0$. This is illustrated by Simpson's paradox for $W = U$ (see sect. 1.1 and Exercise 10-1). ◁

Box 10.1 Unconfoundedness conditions

Simple unconfoundedness conditions

$E^{X=x}(Y) \Vdash D_X$ — *Unconfoundedness of* $E^{X=x}(Y)$. Let Assumptions 10.1 (a) to (d) hold. Then $E^{X=x}(Y)$ is called *unconfounded* if τ_x is P-unique and

$$\forall W \in \mathscr{W}_X\colon\; E^{X=x}(Y) = E\big(E^{X=x}(Y|W)\big). \tag{i}$$

$E^{X=x}(Y) \Vdash D_X$ implies that $E^{X=x}(Y) = E(Y|X{=}x)$ is unbiased.

$E(Y|X) \Vdash D_X$ — *Unconfoundedness of* $E(Y|X)$. Let Assumptions 10.1 (a) to (d) and (g) hold. Then it is defined by

$$\forall x \in X(\Omega)\colon\; E^{X=x}(Y) \Vdash D_X. \tag{ii}$$

$E(Y|X) \Vdash D_X$ implies that $E(Y|X)$ is unbiased.

Conditional unconfoundedness conditions

$E^{X=x}(Y|Z) \Vdash D_X$ — *Unconfoundedness of* $E^{X=x}(Y|Z)$. Let Assumptions 10.1 (a) to (e) hold. Then $E^{X=x}(Y|Z)$ is called *unconfounded* if τ_x is P-unique and

$$\forall W \in \mathscr{W}_X\colon\; E^{X=x}(Y|Z) \underset{P}{=} E\big(E^{X=x}(Y|Z,W)\,\big|\,Z\big). \tag{iii}$$

$E^{X=x}(Y|Z) \Vdash D_X$ implies that $E^{X=x}(Y|Z)$ is unbiased.

$E^{X=x}(Y|Z{=}z) \Vdash D_X$ — *Unconfoundedness of* $E^{X=x}(Y|Z{=}z)$. Let Assumptions 10.1 (a) to (f) hold. Then $E^{X=x}(Y|Z{=}z)$ is called *unconfounded* if τ_x is $P^{Z=z}$-unique and

$$\forall W \in \mathscr{W}_X\colon\; E^{X=x}(Y|Z{=}z) = E\big(E^{X=x}(Y|Z,W)\,\big|\,Z{=}z\big). \tag{iv}$$

$E^{X=x}(Y|Z{=}z) \Vdash D_X$ implies that $E^{X=x}(Y|Z{=}z) = E(Y|X{=}x, Z{=}z)$ is unbiased.

$E(Y|X,Z) \Vdash D_X$ — *Unconfoundedness of* $E(Y|X,Z)$. Let Assumptions 10.1 (a) to (e) and (g) hold. Then it is defined by

$$\forall x \in X(\Omega)\colon\; E^{X=x}(Y|Z) \Vdash D_X. \tag{v}$$

$E(Y|X,Z) \Vdash D_X$ implies that $E(Y|X,Z)$ is unbiased.

Note: The proofs that the unconfoundedness conditions imply unbiasedness of the specified conditional expectations or expectation values are found in section 10.5.

Definition 10.7 [Unconfoundedness of $E^{X=x}(Y)$ and $E(Y|X)$]
Let the Assumptions 10.1 (a) to (d) hold.

(i) Then the conditional expectation value $E^{X=x}(Y)$ is called unconfounded, denoted $E^{X=x}(Y) \Vdash D_X$, if τ_x is P-unique and

$$\forall W \in \mathscr{W}_X\colon\; E^{X=x}(Y) \;=\; E\big(E^{X=x}(Y|W)\big). \tag{10.6}$$

Let the Assumptions 10.1 (a) to (d) and (g) hold.

(ii) *Then the conditional expectation* $E(Y|X)$ *is called* *unconfounded*, *denoted* $E(Y|X) \Vdash D_X$, *if for all* $x \in X(\Omega)$: $E^{X=x}(Y) \Vdash D_X$.

Remark 10.8 [Unconfoundedness of $E(Y|X)$] Note that $E^{X=x}(Y) \Vdash D_X$ implies that τ_x is P-unique. Hence, under the assumptions of Definition 10.7 (ii), the conditional expectation $E(Y|X)$ is unconfounded if

(a) For all $x \in X(\Omega)$: τ_x is P-unique
(b) For all $x \in X(\Omega)$, $\forall\, W \in \mathscr{W}_X$: $E^{X=x}(Y) = E\big(E^{X=x}(Y|W)\big)$.

$\triangleleft$

Remark 10.9 [Expectation Stability of $(X{=}x)$-Conditional Expectation Values] If the Assumptions 10.1 (a) to (d) and (h) hold, and W is a potential confounder of X, then $E^{X=x}(Y) \Vdash D_X$ implies Equation (10.6). This equation is also called *expectation stability of the* $(X{=}x)$*-conditional expectation value* $E^{X=x}(Y)$. It means that the expectation $E\big(E^{X=x}(Y|W)\big)$ is identical to $E^{X=x}(Y)$ and identical for all potential confounders $W \in \mathscr{W}_X$. In section 10.6 we will study the consequences of this property for the prima facie effect $PFE_{xx'}$. $\triangleleft$

Remark 10.10 [Equivalent Equations for $E(E^{X=x}(Y|W))$] If P_W denotes the distribution of W (see RS-Def. 2.38), then

$$
\begin{aligned}
E\big(E^{X=x}(Y|W)\big) &= \int E^{X=x}(Y|W)\; dP && \text{[RS-Def. 3.1]}\\
&= \int E^{X=x}(Y|W{=}w)\; P_W(dw) && \text{[RS-(3.12)]} \qquad (10.7)\\
&= \int E(Y|X{=}x, W{=}w)\; P_W(dw). && \text{[RS-(5.26)]}
\end{aligned}
$$

If the image $W(\Omega)$ of Ω under W is finite, then, according to RS-Equation (3.13), then the second of these equations can also be written

$$E\big(E^{X=x}(Y|W)\big) = \sum_{w \in W(\Omega)} E^{X=x}(Y|W{=}w) \cdot P(W{=}w). \tag{10.8}$$

Hence, if $W(\Omega)$ is finite and we assume $P(X{=}x, W{=}w) > 0$ for all $w \in W(\Omega)$, then $E^{X=x}(Y) = E\big(E^{X=x}(Y|W)\big)$ [see Eq. (10.6)] can be written as

$$E^{X=x}(Y) = \sum_{w \in W(\Omega)} E^{X=x}(Y|W{=}w) \cdot P(W{=}w). \tag{10.9}$$

To emphasize again, Equation (10.9) is not always true even if we assume that $W(\Omega)$ is finite. In contrast, if $P(X{=}x, W{=}w) > 0$ for all $w \in W(\Omega)$, then

$$E^{X=x}(Y) = \sum_{w \in W(\Omega)} E^{X=x}(Y|W{=}w) \cdot P^{X=x}(W{=}w) \tag{10.10}$$

is always true [see RS-Box 3.2 (ii)]. Comparing the highlighted terms in these two equations is useful in order to more easily comprehend the concept of unconfoundedness of a conditional expectation value $E^{X=x}(Y) = E(Y|X{=}x)$. $\triangleleft$

Remark 10.11 [Two Sufficient Conditions] Let Assumptions 10.1 (a) to (d) hold and let $0 < P(X{=}x, W{=}w) < 1$ for all $w \in W(\Omega)$. Then, comparing Equations (10.9) and (10.10) to each other reveals two sufficient conditions for Equation (10.9):

(i) $P^{X=x}(W{=}w) = P(W{=}w), \quad \forall\, w \in W(\Omega).$
(ii) $E^{X=x}(Y \mid W{=}w) = E^{X=x}(Y), \quad \forall\, w \in W(\Omega).$

If condition (i) is true, then it is obvious that Equation (10.9) holds because, under the assumption $P(X{=}x, W{=}w) > 0$ for all $w \in W(\Omega)$, Equation (10.10) is always true.

If condition (ii) holds, then

$$\begin{aligned} E^{X=x}(Y) &= E^{X=x}(Y) \cdot \sum_{w \in W(\Omega)} P(W{=}w) && \Big[\sum_{w \in W(\Omega)} P(W{=}w) = 1\Big] \\ &= \sum_{w \in W(\Omega)} E^{X=x}(Y) \cdot P(W{=}w) && [E^{X=x}(Y) \text{ is a constant}] \\ &= \sum_{w \in W(\Omega)} E^{X=x}(Y \mid W{=}w) \cdot P(W{=}w). && [\text{(ii)}] \end{aligned}$$

Hence, (ii) implies Equation (10.9) as well. These sufficient conditions will be treated in more detail and generality in section 10.2. ◁

Now we additionally consider a covariate Z of X and define unconfoundedness of the conditional expectation value $E^{X=x}(Y \mid Z{=}z)$ and the conditional expectations $E^{X=x}(Y \mid Z)$ and $E(Y \mid X, Z)$.

Definition 10.12 [Unconfoundedness of $E^{X=x}(Y \mid Z)$, $E^{X=x}(Y \mid Z{=}z)$, and $E(Y \mid X, Z)$]
Let the Assumptions 10.1 (a) to (e) hold.

(i) Then $E^{X=x}(Y \mid Z)$ is called unconfounded, denoted $E^{X=x}(Y \mid Z) \Vdash D_X$, if τ_x is P-unique and

$$\forall\, W \in \mathscr{W}_X\colon\; E^{X=x}(Y \mid Z) \underset{P}{=} E\big(E^{X=x}(Y \mid Z, W) \,\big|\, Z\big). \tag{10.11}$$

Let the Assumptions 10.1 (a) to (f) hold.

(ii) Then $E^{X=x}(Y \mid Z{=}z)$ is called unconfounded, denoted $E^{X=x}(Y \mid Z{=}z) \Vdash D_X$, if τ_x is $P^{Z=z}$-unique and

$$\forall\, W \in \mathscr{W}_X\colon\; E^{X=x}(Y \mid Z{=}z) = E\big(E^{X=x}(Y \mid Z, W) \,\big|\, Z{=}z\big). \tag{10.12}$$

Let the Assumptions 10.1 (a) to (e) and (g) hold.

(iii) Then $E(Y \mid X, Z)$ is called unconfounded, denoted $E(Y \mid X, Z) \Vdash D_X$, if for all $x \in X(\Omega)$: $E^{X=x}(Y \mid Z) \Vdash D_X$.

Remark 10.13 [Expectation Stability of the Conditional Expectation $E^{X=x}(Y \mid Z)$] If the Assumptions 10.1 (a) to (e) and (h) hold, and W is a potential confounder of X, then the condition $E^{X=x}(Y \mid Z) \Vdash D_X$ implies Equation (10.11). This equation is also called *expectation stability of the conditional expectation* $E^{X=x}(Y \mid Z)$. It means that the Z-conditional expectation $E\big(E^{X=x}(Y|W)\,\big|\,Z\big)$ is identical to $E^{X=x}(Y|Z)$ and identical for all potential confounders $W \in \mathscr{W}_X$ of X. In section 10.6 we will study the consequences of this property for the Z-conditional effect variable $PFE_{Z;xx'}$. $\triangleleft$

Remark 10.14 [Equivalent Conditions] Let the Assumptions 10.1 (a) to (e) hold, assume that τ_x is P-unique, and that the conditional distribution $P_{W|Z}$ exists (see SN-section 17.3). Furthermore, let P_Z denote the distribution of Z (see RS-Def. 2.38) and $P_{W|Z=z}$ a $(Z{=}z)$-conditional distribution of W (see SN-Def. 17.7). Then Equation (10.11) is equivalent to

$$\forall\, W \in \mathscr{W}_X\!:\; E^{X=x}(Y|Z{=}z) \;=\; \int E^{X=x}(Y|Z{=}z, W{=}w)\; P_{W|Z=z}\,(dw), \quad \text{for } P_Z\text{-almost all } z \in \Omega'_Z \tag{10.13}$$

(see Exercise 10-2). If we assume $P(X{=}x, Z{=}z, W{=}w) > 0$ for all values $w \in W(\Omega)$, then the equation in Proposition (10.13) simplifies to

$$E^{X=x}(Y|Z{=}z) \;=\; \sum_{w \in W(\Omega)} E^{X=x}(Y|Z{=}z, W{=}w) \cdot P(W{=}w\,|\,Z{=}z). \tag{10.14}$$

Note again that Equation (10.14) is not always true even if $P(X{=}x, Z{=}z, W{=}w) > 0$ for all values $w \in W(\Omega)$. In contrast, if this probability is positive, then

$$E^{X=x}(Y|Z{=}z) \;=\; \sum_{w \in W(\Omega)} E^{X=x}(Y|Z{=}z, W{=}w) \cdot P^{X=x}(W{=}w\,|\,Z{=}z) \tag{10.15}$$

is always true [see RS-Box 3.2 (ii)]. And again, comparing the highlighted terms in the last two equations is useful in order to more easily understand the concept of unconfoundedness of a conditional expectation value $E^{X=x}(Y|Z{=}z) = E(Y|X{=}x, Z{=}z)$. $\triangleleft$

Remark 10.15 [Two Sufficient Conditions] Let the Assumptions 10.1 (a) to (e) hold and, additionally, let $P(X{=}x, Z{=}z, W{=}w) > 0$ for all $w \in W(\Omega)$. Then, comparing Equations (10.14) and (10.15) to each other reveals two sufficient conditions of Equation (10.14):

(i) $P^{X=x}(W{=}w\,|\,Z{=}z) = P(W{=}w\,|\,Z{=}z), \quad \forall\, w \in W(\Omega)$
(ii) $E^{X=x}(Y|Z{=}z, W{=}w) = E^{X=x}(Y|Z{=}z), \quad \forall\, w \in W(\Omega)$.

If condition (i) is true, then it is obvious that Equation (10.14) holds because, under the assumption $P(X{=}x, Z{=}z, W{=}w) > 0$ for all $w \in W(\Omega)$, Equation (10.15) is always true.

If condition (ii) holds, then

$$\begin{aligned}
&E^{X=x}(Y|Z{=}z) \\
&= E^{X=x}(Y|Z{=}z) \cdot \sum_{w \in W(\Omega)} P(W{=}w\,|\,Z{=}z) && \Big[\sum_{w \in W(\Omega)} P(W{=}w\,|\,Z{=}z) = 1\Big] \\
&= \sum_{w \in W(\Omega)} E^{X=x}(Y|Z{=}z) \cdot P(W{=}w\,|\,Z{=}z) && [E^{X=x}(Y|Z{=}z) \text{ is a constant}] \\
&= \sum_{w \in W(\Omega)} E^{X=x}(Y|Z{=}z, W{=}w) \cdot P(W{=}w\,|\,Z{=}z). && [\text{(ii)}]
\end{aligned}$$

Hence, (ii) implies Equation (10.14) as well. Again, these two sufficient conditions will be treated in more detail and generality in section 10.2. $\triangleleft$

Table 10.1 summarizes the implication relations of the unconfoundedness conditions that have been defined in Box 10.1. It also displays the assumptions under which these implications hold.

Remark 10.16 [Two Immediate Implications] One of the implications in Table 10.1 is

$$E(Y\,|\,X) \Vdash D_X \quad \Rightarrow \quad E^{X=x}(Y) \Vdash D_X, \tag{10.16}$$

which, under the Assumptions 10.1 (a) to (d), (g), and (i), immediately follows from the definition of unconfoundedness of the conditional expectation $E(Y\,|\,X)$ [see Def. 10.7 (ii)]. Similarly, under the Assumptions 10.1 (a) to (e), (g), and (i),

$$E(Y\,|\,X,Z) \Vdash D_X \quad \Rightarrow \quad E^{X=x}(Y\,|\,Z) \Vdash D_X \tag{10.17}$$

immediately follows from the definition of unconfoundedness of the conditional expectation $E(Y\,|\,X,Z)$ [see Def. 10.12 (iii)]. $\lhd$

Two other implications listed in Table 10.1 are specified in the following theorem.

Theorem 10.17 [Consequences of $E^{X=x}(Y\,|\,Z) \Vdash D_X$ and $E(Y\,|\,X,Z) \Vdash D_X$]
Let the Assumptions 10.1 (a) to (f) and (h) hold. Then

$$E^{X=x}(Y\,|\,Z) \Vdash D_X \quad \Rightarrow \quad E^{X=x}(Y\,|\,Z{=}z) \Vdash D_X. \tag{10.18}$$

If we additionally assume 10.1 (g) and (i), then

$$E(Y\,|\,X,Z) \Vdash D_X \quad \Rightarrow \quad E^{X=x}(Y\,|\,Z{=}z) \Vdash D_X. \tag{10.19}$$

(Proof p. 329)

Finally, we turn to a proposition that is not listed in Table 10.1. According to the following theorem, if Z is discrete and $0 < P(X{=}x, Z{=}z) < 1$ for all $z \in Z(\Omega)$, then unconfoundedness of $E^{X=x}(Y\,|\,Z{=}z)$ for all values of Z is equivalent to unconfoundedness of the conditional expectation $E^{X=x}(Y\,|\,Z)$.

Theorem 10.18 [A Condition Equivalent to Unconfoundedness of $E^{X=x}(Y\,|\,Z)$]
Let Assumptions 10.1 (a) to (f) hold, and, for all $z \in Z(\Omega)$, let $0 < P(X{=}x, Z{=}z) < 1$ and τ_x be $P^{Z=z}$-unique. Then

$$\big(\forall\, z \in Z(\Omega)\colon E^{X=x}(Y\,|\,Z{=}z) \Vdash D_X\big) \quad \Leftrightarrow \quad E^{X=x}(Y\,|\,Z) \Vdash D_X. \tag{10.20}$$

(Proof p. 330)

10.2 Sufficient Conditions of Unconfoundedness

In Remarks 10.11 and 10.15 we already treated special cases in which Fisher conditions and Suppes-Reichenbach conditions imply an unconfoundedness condition. In this section, we revisit these sufficient conditions in more detail and generality.

Table 10.1. Implications among unconfoundedness conditions

	$E^{X=x}(Y \mid Z=z) \Vdash D_X$	$E^{X=x}(Y \mid Z) \Vdash D_X$	$E^{X=x}(Y) \Vdash D_X$
$E(Y \mid X) \Vdash D_X$			(a)-(d), (g), (i)
$E^{X=x}(Y \mid Z) \Vdash D_X$	(a)-(f), (h)		
$E(Y \mid X, Z) \Vdash D_X$	(a)-(g), (i)	(a)-(e), (g), (i)	

Note: An entry such as (a)-(f), (h) means that the condition in the row implies the condition in the column, provided that Assumptions 10.1 (a) to (f) and (h) hold. Trivial implications such as $E(Y \mid X) \Vdash D_X \Rightarrow E(Y \mid X) \Vdash D_X$ are omitted.

10.2.1 Fisher Conditions

We study the implications on unconfoundedness of the following conditions:

(a) $D_X \perp\!\!\!\perp 1_{X=x}$, that is, independence of the global potential confounder D_X and the indicator $1_{X=x}$ of the event that the putative cause variable X takes on the value x.
(b) $D_X \perp\!\!\!\perp X$, that is, independence of D_X and X.
(c) $D_X \perp\!\!\!\perp 1_{X=x} \mid (Z=z)$, that is, $(Z=z)$-conditional independence of D_X and $1_{X=x}$.
(d) $D_X \perp\!\!\!\perp 1_{X=x} \mid Z$, that is, Z-conditional independence of D_X and $1_{X=x}$.
(e) $D_X \perp\!\!\!\perp X \mid (Z=z)$, that is, $(Z=z)$-conditional independence of D_X and X.
(f) $D_X \perp\!\!\!\perp X \mid Z$, that is, Z-conditional independence of D_X and X.

These conditions have been introduced in Box 8.1. They are summarily called the *Fisher conditions* (F-conditions). The implications of these F-conditions on various unconfoundedness conditions are summarized in Table 10.2 and proved in the theorems presented this subsection.

The first theorem presents the implications of independence of the global potential confounder D_X of X and the indicator $1_{X=x}$ on some unconfoundedness conditions. Remember that $D_X \perp\!\!\!\perp 1_{X=x}$ implies that τ_x is P-unique [see Th. 8.25 (ii)]. Therefore, it is not needed as an extra assumption.

Theorem 10.19 [Implications of $D_X \perp\!\!\!\perp 1_{X=x}$ on Unconfoundedness]
Let the Assumptions 10.1 (a) to (d) hold. Then

(i) $D_X \perp\!\!\!\perp 1_{X=x} \;\Rightarrow\; E^{X=x}(Y) \Vdash D_X$.

If we additionally assume that Z is a covariate of X [i. e., Ass. 10.1 (e)], then

(ii) $D_X \perp\!\!\!\perp 1_{X=x} \;\Rightarrow\; E^{X=x}(Y \mid Z) \Vdash D_X$.

If we additionally assume $0 < P(X=x, Z=z) < 1$ *[i. e., Ass. 10.1 (f)], then*

Table 10.2. Implications of the Fisher conditions on unconfoundedness

	$E^{X=x}(Y) \Vdash D_X$	$E(Y \mid X) \Vdash D_X$	$E^{X=x}(Y \mid Z) \Vdash D_X$	$E^{X=x}(Y \mid Z=z) \Vdash D_X$	$E(Y \mid X, Z) \Vdash D_X$
$D_X \perp\!\!\!\perp 1_{X=x}$	(a)-(d)		(a)-(e)	(a)-(f)	
$D_X \perp\!\!\!\perp X$	(a)-(d)	(a)-(d), (g)	(a)-(e)	(a)-(f)	(a)-(e), (g)
$D_X \perp\!\!\!\perp 1_{X=x} \mid (Z=z)$				(a)-(f)	
$D_X \perp\!\!\!\perp 1_{X=x} \mid Z$			(a)-(e), (h)	(a)-(f)	
$D_X \perp\!\!\!\perp X \mid (Z=z)$				(a)-(f)	
$D_X \perp\!\!\!\perp X \mid Z$			(a)-(e), (h)	(a)-(f)	(a)-(e), (g), (i)

Note: An entry such as (a)-(d) means that the condition in the row implies the condition in the column, provided that the Assumptions 10.1 (a) to (d) hold.

(iii) $D_X \perp\!\!\!\perp 1_{X=x} \;\Rightarrow\; E^{X=x}(Y \mid Z=z) \Vdash D_X$.

(Proof p. 330)

According to Theorem 10.19 (i), independence of a global potential confounder D_X and the indicator variable $1_{X=x}$ implies that the conditional expectation value $E^{X=x}(Y) = E(Y \mid X=x)$ is unconfounded. Furthermore, according to Proposition (ii) of this theorem, independence of D_X and $1_{X=x}$ also implies that the conditional expectation $E^{X=x}(Y \mid Z)$ is unconfounded, provided that Z is a covariate of X. Finally, according to Proposition (iii), independence of D_X and $1_{X=x}$ also implies that the conditional expectation value $E^{X=x}(Y \mid Z=z) = E(Y \mid X=x, Z=z)$ is unconfounded, provided that Z is a covariate of X and $0 < P(X=x, Z=z) < 1$.

In Theorem 10.19 we only considered a single value x of (the putative cause variable) X and the indicator variable of the event $\{X=x\}$. Now we turn to the implications of independence of a global potential confounder D_X and X itself. This includes the case in which X has more than just two values. We start with the following corollary, which immediatly follows from Theorem 10.19 because $D_X \perp\!\!\!\perp X$ implies $D_X \perp\!\!\!\perp 1_{X=x}$.

Corollary 10.20 [First Implications of $D_X \perp\!\!\!\perp X$ on Unconfoundedness]
Let the Assumptions 10.1 (a) to (d) hold. Then

(i) $D_X \perp\!\!\!\perp X \;\Rightarrow\; E^{X=x}(Y) \Vdash D_X$.

If we additionally assume that Z is a covariate of X [i. e., Ass. 10.1 (e)], then

(ii) $D_X \perp\!\!\!\perp X \;\Rightarrow\; E^{X=x}(Y \mid Z) \Vdash D_X$.

If we additionally assume $0 < P(X=x, Z=z) < 1$ *[i. e., Ass. 10.1 (f)], then* *(iii)* $D_X \perp\!\!\!\perp X \;\Rightarrow\; E^{X=x}(Y \mid Z{=}z) \Vdash D_X$.

Hence, under the assumptions of Corollary 10.20, independence of the putative cause variable X and the global potential confounder D_X of X implies that the conditional expectation value $E^{X=x}(Y)$ is unconfounded [see Prop. (i)]. Furthermore, according to Proposition (ii), independence of D_X and X also implies that the conditional expectation $E^{X=x}(Y \mid Z)$ is unconfounded, provided that we additionally assume that Z is a covariate of X. Finally, according to Proposition (iii), independence of D_X and X implies that the conditional expectation value $E^{X=x}(Y \mid Z{=}z) = E(Y \mid X{=}x, Z{=}z)$ is unconfounded, provided that Z is a covariate of X and we additionally assume $0 < P(X{=}x, Z{=}z) < 1$.

Theorem 10.21 [More Implications of $D_X \perp\!\!\!\perp X$ on Unconfoundedness]
Let the Assumptions 10.1 (a) to (d) and (g) hold. Then

(i) $D_X \perp\!\!\!\perp X \;\Rightarrow\; E(Y \mid X) \Vdash D_X$.

If we additionally assume that Z is a covariate of X [i. e., Ass. 10.1 (e)], then

(ii) $D_X \perp\!\!\!\perp X \;\Rightarrow\; \forall\, x \in X(\Omega)\colon E^{X=x}(Y \mid Z) \Vdash D_X$
(iii) $D_X \perp\!\!\!\perp X \;\Rightarrow\; E(Y \mid X, Z) \Vdash D_X$.

(Proof p. 331)

Hence, under the assumptions of Theorem 10.21, independence of the putative cause variable X and the global potential confounder D_X of X implies that the conditional expectation $E(Y \mid X)$ is unconfounded [see Prop. (i)]. Furthermore, according to Propositions (ii) and (iii), independence of D_X and X also implies that all conditional expectations $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$, and $E(Y \mid X, Z)$ are unconfounded, provided that we additionally assume that Z is a covariate of X.

Now we turn to implications of Z-conditional independence of a global potential confounder D_X of X and an indicator $1_{X=x}$. Note, if Z is a covariate of X, then $D_X \perp\!\!\!\perp 1_{X=x} \mid Z$ follows from $D_X \perp\!\!\!\perp 1_{X=x}$ [see RS-Box 6.1 (ix)], but not vice versa. That is, $D_X \perp\!\!\!\perp 1_{X=x}$ does not follow from $D_X \perp\!\!\!\perp 1_{X=x} \mid Z$. Therefore, it is meaningful to study the implications of $D_X \perp\!\!\!\perp 1_{X=x} \mid Z$.

In Theorem 10.22 we assume that τ_x is P-unique. This assumption has not been necessary in Theorems 10.19 and 10.21 because it already follows from $D_X \perp\!\!\!\perp 1_{X=x}$ and the assumption that Z is a covariate of X. Under this assumption, according to RS-Box 5.1 (iv), P-uniqueness of $E^{X=x}(Y \mid Z)$ follows from P-uniqueness of τ_x because then $\sigma(Z) \subset \sigma(D_X)$ [see Def. 4.11 (iv) and Rem. 4.16]. Also remember that P-uniqueness of $\tau_x = E^{X=x}(Y \mid D_X)$ is equivalent to $P(X{=}x \mid D_X) \underset{P}{>} 0$ whereas P-uniqueness of $E^{X=x}(Y \mid Z)$ is equivalent to $P(X{=}x \mid Z) \underset{P}{>} 0$ (see RS-Th. 5.27).

Theorem 10.22 [$D_X \perp\!\!\!\perp 1_{X=x} \mid Z$ Implies Unconfoundedness of $E^{X=x}(Y \mid Z)$]
Let the Assumptions 10.1 (a) to (e) and (h) hold. Then

$$D_X \perp\!\!\!\perp 1_{X=x} \mid Z \quad\Rightarrow\quad E^{X=x}(Y \mid Z) \Vdash D_X. \tag{10.21}$$

(Proof p. 331)

Hence, under the assumptions of Theorem 10.22, Z-conditional independence of a global potential confounder D_X of X and the indicator $1_{X=x}$ of the event $\{X=x\}$ that the putative cause variable X takes on the value x implies that $E^{X=x}(Y|Z)$ is unconfounded.

Now we study the conditions under which Z-conditional independence as well as $(Z=z)$-conditional independence of D_X and $1_{X=x}$ imply that the conditional expectation value $E^{X=x}(Y|Z=z) = E(Y|X=x, Z=z)$ is unconfounded.

Theorem 10.23 [$D_X \perp\!\!\!\perp 1_{X=x} | Z$ **Implies That** $E^{X=x}(Y|Z=z)$ **is Unconfounded**]
Let the Assumptions 10.1 (a) to (f) hold. Then

$$D_X \perp\!\!\!\perp 1_{X=x} | Z \quad \Rightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} | (Z=z) \tag{10.22}$$

$$\Rightarrow \quad E^{X=x}(Y|Z=z) \Vdash D_X. \tag{10.23}$$

(Proof p. 331)

Now we turn to implications of Z-conditional independence of D_X and X. Again note that, if Z is a covariate of X, then $D_X \perp\!\!\!\perp X | Z$ follows from $D_X \perp\!\!\!\perp X$ [see RS-Box 6.1 (ix)], but not vice versa. Therefore, it is meaningful to study the implications of $D_X \perp\!\!\!\perp X | Z$. In this theorem, we need the assumption that τ_x is P-unique. Remember that P-uniqueness of τ_x follows from $P^{Z=z}$-uniqueness for all values z of Z, which presumes $P(Z=z) > 0$ for all $z \in Z(\Omega)$ (see Lem. 5.37 for details). Also note that τ_x can be P-unique if Z is continuous and $P(Z=z) = 0$ for all $z \in Z(\Omega)$.

Theorem 10.24 [$D_X \perp\!\!\!\perp X | Z$ **Implies That** $E(Y|X, Z)$ **is Unconfounded**]
Let the Assumptions 10.1 (a) to (e), (g), and (i) hold. Then

$$D_X \perp\!\!\!\perp X | Z \quad \Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|Z) \Vdash D_X \tag{10.24}$$

$$\Leftrightarrow \quad E(Y|X, Z) \Vdash D_X. \tag{10.25}$$

(Proof p. 332)

Hence, under the assumptions of Theorem 10.24, Z-conditional independence of the putative cause variable X and a global potential confounder D_X of X implies that all conditional expectations $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, are unconfounded, and this is equivalent to unconfoundedness of the conditional expectation $E(Y|X, Z)$.

According to the last theorem in this subsection, Z-conditional independence of the putative cause variable X and the global potential confounder D_X of X implies that the conditional expectation value $E^{X=x}(Y|Z=z) = E(Y|X=x, Z=z)$ is unconfounded.

Theorem 10.25 [$D_X \perp\!\!\!\perp X | Z$ **Implies That** $E^{X=x}(Y|Z=z)$ **is Unconfounded**]
Let the Assumptions 10.1 (a) to (f) hold. Then

$$D_X \perp\!\!\!\perp X | Z \quad \Rightarrow \quad D_X \perp\!\!\!\perp X | (Z=z) \tag{10.26}$$

$$\Rightarrow \quad E^{X=x}(Y|Z=z) \Vdash D_X. \tag{10.27}$$

(Proof p. 332)

Table 10.2 summarizes the propositions and their assumptions that have been detailed and proved in the theorems of this subsection.

10.2.2 Suppes-Reichenbach Conditions

In this subsection, we show that the unconfoundedness conditions also follow from the Suppes-Reichenbach conditions (SR-conditions). Remember, while the F-conditions focus on (Z-conditional) independence of *the putative cause variable* X and the global potential confounder D_X of X, the SR-conditions deal with X-conditional [or (X,Z)-conditional independence] of the *outcome variable* Y and D_X.

Note that the conditional independence SR-conditions imply their corresponding conditional mean-independence SR-conditions (see Box 9.1). Therefore, it suffices to confine ourselves to the conditional mean-independence SR-conditions. The implications of the conditional independence SR-conditions on unconfoundedness follow immediately. Therefore, in Table 10.3, which summarizes all implications of the SR-mean-independence conditions on unconfoundedness that are proved in this subsection, we may simply replace the symbol $\vDash$ for *mean-independence* by the symbol $\perp\!\!\!\perp$ for *independence*.

Reading the first theorem, remember

$$Y \vDash D_X \,|\, (X{=}x) \quad :\Leftrightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) \tag{10.28}$$

[see Box 9.1 (ii)]. Furthermore, under the Assumptions 10.1 (a) to (f) and (j),

$$E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) \quad \Rightarrow \quad E\big(E^{X=x}(Y \,|\, Z,W) \,\big|\, Z{=}z\big) = E^{X=x}(Y) \tag{10.29}$$

(see Exercise 10-4).

Finally, assuming that the true outcome variable $\tau_x = E^{X=x}(Y \,|\, D_X)$ is P-unique [see Ass. 10.1 (h)] yields

$$Y \vDash D_X \,|\, (X{=}x) \quad \Leftrightarrow \quad E^{X=x}(Y \,|\, D_X) \underset{P}{=} E^{X=x}(Y). \tag{10.30}$$

Under the Assumptions 10.1 (a) to (e), (h), and that W is a potential confounder of X,

$$E^{X=x}(Y \,|\, D_X) \underset{P}{=} E^{X=x}(Y) \quad \Rightarrow \quad E\big(E^{X=x}(Y \,|\, Z,W) \,\big|\, Z\big) \underset{P}{=} E^{X=x}(Y) \tag{10.31}$$

(see Exercise 10-5).

Theorem 10.26 [Implications of $Y \vDash D_X \,|\, (X{=}x)$ on Unconfoundedness]
Let Assumptions 10.1 (a) to (d) hold and let τ_x be P-unique [i. e., Ass. 10.1 (h)]. Then

(i) $\quad Y \vDash D_X \,|\, (X{=}x) \quad \Rightarrow \quad E^{X=x}(Y) \Vdash D_X.$

If we additionally assume that Z is a covariate of X [i. e., Ass. 10.1 (e)], then

(ii) $\quad Y \vDash D_X \,|\, (X{=}x) \quad \Rightarrow \quad E^{X=x}(Y \,|\, Z) \Vdash D_X.$

Let the Assumptions 10.1 (a) to (f) hold and let τ_x be $P^{Z=z}$-unique [i. e., Ass. 10.1 (j)]. Then

(iii) $\quad Y \vDash D_X \,|\, (X{=}x) \quad \Rightarrow \quad E^{X=x}(Y \,|\, Z{=}z) \Vdash D_X.$

(Proof p. 332)

Table 10.3. Implications of Suppes-Reichenbach conditions on unconfoundedness

	$E^{X=x}(Y) \Vdash D_X$	$E(Y \mid X) \Vdash D_X$	$E^{X=x}(Y \mid Z) \Vdash D_X$	$E^{X=x}(Y \mid Z=z) \Vdash D_X$	$E(Y \mid X, Z) \Vdash D_X$
$Y \vDash D_X \mid (X=x)$	(a)-(d), (h)		(a)-(e), (h)	(a)-(f), (j)	
$Y \vDash D_X \mid X$	(a)-(d), (h)	(a)-(d), (g), (i)	(a)-(e), (h)	(a)-(f), (j)	(a)-(e), (g), (i)
$Y \vDash D_X \mid (X=x, Z)$			(a)-(e), (h)	(a)-(f), (j)	
$Y \vDash D_X \mid (X, Z)$			(a)-(e), (h)	(a)-(f), (j)	(a)-(e), (g), (i)

Note: An entry such as (a)-(d), (h) means that the condition in the row implies the condition in the column, provided that the Assumptions 10.1 (a) to (d) and (h) hold.

Hence, under the assumptions of Theorem 10.26 (i), $(X=x)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X [see Eq. (10.28)] implies unconfoundedness of $E^{X=x}(Y)$, provided that τ_x is P-unique. More generally, according to Theorem 10.26 (ii), $(X=x)$-conditional mean-independence of Y from D_X also implies unconfoundedness of $E^{X=x}(Y \mid Z)$, provided that τ_x is P-unique and Z is a covariate of X. Finally, if we consider a value z of Z, then under the assumptions of Theorem 10.26 (iii), $Y \vDash D_X \mid (X=x)$ also implies that the conditional expectation value $E^{X=x}(Y \mid Z=z) = E(Y \mid X=x, Z=z)$ is unconfounded. For this implication we do not have to assume that τ_x is P-unique. Instead, $P^{Z=z}$-uniqueness of τ_x suffices.

Now we turn to the implications of X-conditional mean-independence of Y from the global potential confounder D_X on various unconfoundedness conditions. In Box 9.1 (iv), this condition has been defined by

$$Y \vDash D_X \mid X \quad :\Leftrightarrow \quad E(Y \mid X, D_X) \underset{P}{=} E(Y \mid X)\,. \tag{10.32}$$

Corollary 10.27 [First Implications of $Y \vDash D_X \mid X$ on Unconfoundedness]
Let the Assumptions 10.1 (a) to (d) hold and let τ_x be P-unique [i. e., Ass. 10.1 (h)]. Then

(i) $Y \vDash D_X \mid X \quad \Rightarrow \quad E^{X=x}(Y) \Vdash D_X$.

If we additionally assume that Z is a covariate of X [i. e., Ass. 10.1 (e)], then

(ii) $Y \vDash D_X \mid X \quad \Rightarrow \quad E^{X=x}(Y \mid Z) \Vdash D_X$.

Let the Assumptions 10.1 (a) to (f) and (j) hold. Then

(iii) $Y \vDash D_X \mid X \quad \Rightarrow \quad E^{X=x}(Y \mid Z=z) \Vdash D_X$.

(Proof p. 333)

Hence, under the assumptions of Corollary 10.27 (i), X-conditional mean independence of the outcome variable Y from the global potential confounder D_X of X implies that the conditional expectation value $E^{X=x}(Y)$ is unconfounded. Furthermore, if Z is a covariate of X, then it also implies that the conditional expectation $E^{X=x}(Y|Z)$ is unconfounded. Finally, if we consider a value z of Z, then under the assumptions of Corollary 10.27 (iii), $Y \vDash D_X|X$ also implies that the conditional expectation value $E^{X=x}(Y|Z=z) = E(Y|X=x, Z=z)$ is unconfounded. Again, for this implication we do not have to assume that τ_x is P-unique. Instead, $P^{Z=z}$-unique of τ_x suffices.

Corollary 10.28 [More Implications of $Y \vDash D_X|X$ on Unconfoundedness]
Let the Assumptions 10.1 (a) to (d) and (g) hold. Furthermore, let that all τ_x, $x \in X(\Omega)$, be P-unique [i. e., Ass. 10.1 (i)]. Then

(i) $Y \vDash D_X|X \quad \Rightarrow \quad E(Y|X) \Vdash D_X.$

If we additionally assume that Z is a covariate of X [i. e., Ass. 10.1 (e)], then

(ii) $Y \vDash D_X|X \quad \Rightarrow \quad \forall\, x \in X(\Omega)\colon\ E^{X=x}(Y|Z) \Vdash D_X$

(iii) $Y \vDash D_X|X \quad \Rightarrow \quad E(Y|X,Z) \Vdash D_X.$

(Proof p. 334)

Hence, under the assumptions of Corollary 10.28 (i), X-conditional mean independence of the outcome variable Y from the global potential confounder D_X of X implies that $E(Y|X)$ is unconfounded. Furthermore, if we additionally assume that Z is a covariate of X, then X-conditional mean independence of Y from D_X also implies that each $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, is unconfounded [see Cor. 10.28 (ii)]. Finally, under the assumptions of Corollary 10.28 (iii), X-conditional mean independence of Y from D_X also implies that $E(Y|X,Z)$ is unconfounded, provided that Z is a covariate of X.

Now we turn to the implications of $(X=x, Z)$-conditional mean-independence of Y from D_X, which has been defined by

$$Y \vDash D_X|(X=x, Z) \quad :\Leftrightarrow \quad E^{X=x}(Y|Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) \tag{10.33}$$

[see Box 9.1 (vi)]. Assuming that Z is a covariate of X implies $\sigma(Z) \subset \sigma(D_X)$ and

$$Y \vDash D_X|(X=x, Z) \quad \Leftrightarrow \quad E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z)\,. \tag{10.34}$$

Theorem 10.29 [$Y \vDash D_X|(X=x, Z)$ Implies Unconfoundedness of $E^{X=x}(Y|Z)$]
Let Assumptions 10.1 (a) to (e) hold and assume that τ_x is P-unique [i. e., Ass. 10.1 (h)]. Then

$$Y \vDash D_X|(X=x, Z) \quad \Rightarrow \quad E^{X=x}(Y|Z) \Vdash D_X. \tag{10.35}$$

(Proof p. 334)

Hence, under the assumptions of Theorem 10.29, $(X=x, Z)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X implies unconfoundedness of $E^{X=x}(Y|Z)$, provided that Z is a covariate of X and τ_x is P-unique.

Now we show that $(X=x, Z)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X implies unconfoundedness of the conditional expectation value $E^{X=x}(Y|Z=z)$. For this implication it suffices to assume that τ_x is $P^{Z=z}$-unique.

Theorem 10.30 [$Y \vDash D_X | (X{=}x, Z)$ **Implies Unconfoundedness of** $E^{X=x}(Y|Z{=}z)$]
Let Assumptions 10.1 (a) to (f) hold. Furthermore, assume that τ_x is $P^{Z=z}$-unique [i. e., Ass. 10.1 (j)]. Then

$$Y \vDash D_X | (X{=}x, Z) \quad \Rightarrow \quad E^{X=x}(Y|Z{=}z) \Vdash D_X. \tag{10.36}$$

(Proof p. 334)

Hence, under the assumptions of Theorem 10.30, $(X{=}x, Z)$-conditional mean-independence of Y from D_X also implies unconfoundedness of the conditional expectation value $E^{X=x}(Y|Z{=}z) = E(Y|X{=}x, Z{=}z)$.

Now we turn to an implication of (X, Z)-conditional mean-independence of Y from the global potential confounder D_X on unconfoundedness. Remember,

$$Y \vDash D_X | (X, Z) \quad :\Leftrightarrow \quad E(Y|X, Z, D_X) \underset{P}{=} E(Y|X, Z), \tag{10.37}$$

and if Z is a covariate of X, that is, if Z is D_X-measurable, then

$$Y \vDash D_X | (X, Z) \quad \Leftrightarrow \quad E(Y|X, D_X) \underset{P}{=} E(Y|X, Z). \tag{10.38}$$

Theorem 10.31 [Implications of $Y \vDash D_X | (X, Z)$ on Unconfoundedness]
Let Assumptions 10.1 (a) to (e) hold and assume that τ_x is P-unique [i. e., Ass. 10.1 (h)]. Then

(i) $Y \vDash D_X | (X, Z) \quad \Rightarrow \quad E^{X=x}(Y|Z) \Vdash D_X.$

If we additionally assume 10.1 (g) and that for all $x \in X(\Omega)$, τ_x is P-unique [i. e., Ass. 10.1 (i)], then

(ii) $Y \vDash D_X | (X, Z) \quad \Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|Z) \Vdash D_X$

(iii) $Y \vDash D_X | (X, Z) \quad \Rightarrow \quad E(Y|X, Z) \Vdash D_X.$

(Proof p. 335)

Hence, under the assumptions of Theorem 10.31 (i), (X, Z)-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X implies that $E^{X=x}(Y|Z)$ is unconfounded. If we additionally assume 10.1 (g) and that each τ_x, $x \in X(\Omega)$, is P-unique, then $Y \vDash D_X | (X, Z)$ also implies that each $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, and $E(Y|X, Z)$ are unconfounded [see Prop. (iii)].

In the next theorem we show that $(X{=}x, Z{=}z)$-conditional mean-independence of the outcome variable Y from the global potential confounder D_X of X implies unconfoundedness of the conditional expectation value $E^{X=x}(Y|Z{=}z) = E(Y|X{=}x, Z{=}z)$.

Theorem 10.32 [An Implication of $Y \vDash D_X | (X{=}x, Z{=}z)$ on Unconfoundedness]
Let Assumptions 10.1 (a) to (f) hold and let τ_x be $P^{Z=z}$-unique [i. e., Ass. 10.1 (j)]. Then

$$Y \vDash D_X | (X{=}x, Z{=}z) \quad \Rightarrow \quad E^{X=x}(Y|Z{=}z) \Vdash D_X. \tag{10.39}$$

(Proof p. 335)

The proposition of Theorem 10.32 is now used to show that unconfoundedness of the conditional expectation value $E^{X=x}(Y|Z=z)$ also follows from (X,Z)-conditional mean-independence of Y from the global potential confounder D_X of X, provided that we assume $P(X=x, Z=z) > 0$ and $P^{Z=z}$-uniqueness of the true outcome variable τ_x.

Corollary 10.33 [$Y \vDash D_X | (X,Z)$ **Implies Unconfoundedness of** $E^{X=x}(Y|Z=z)$]
Let Assumptions 10.1 (a) to (f) hold. Furthermore, assume that τ_x is $P^{Z=z}$-unique [i. e., Ass. 10.1 (j)]. Then

$$Y \vDash D_X | (X,Z) \quad \Rightarrow \quad E^{X=x}(Y|Z=z) \Vdash D_X. \tag{10.40}$$

(Proof p. 335)

Table 10.3 summarizes the propositions and their assumptions that have been proved in the theorems and corollaries of this subsection.

Remark 10.34 [Methodological Conclusions] Note that the unconfoundedness conditions are less restrictive than the F-conditions and less restrictive than the SR-conditions. However, according to Theorem 10.21, we can conclude, for example, that unconfoundedness of $E(Y|X)$ and $E(Y|X,Z)$ hold in a perfect randomized experiment because, in this case, $D_X \perp\!\!\!\perp X$ and $D_X \perp\!\!\!\perp X|Z$ (see section 9.4). Similarly, we can conclude, that $E(Y|X,Z)$ is unconfounded in a perfect Z-conditionally randomized experiment because, in this case, $D_X \perp\!\!\!\perp X|Z$ holds (see Th. 10.24). $\triangleleft$

10.3 Hybrid Sufficient Conditions of Unconfoundedness

The propositions of the following corollary immediately follow from Theorems 10.19 and 10.26. Reading this corollary, remember

$$D_X \perp\!\!\!\perp 1_{X=x} \quad \Leftrightarrow \quad P(X=x|D_X) \underset{P}{=} P(X=x) \tag{10.41}$$

[see Eq. (8.1)] and

$$Y \vDash D_X | (X=x) \quad \Leftrightarrow \quad E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) \tag{10.42}$$

[see Eqs. (9.3) and (9.4)].

Corollary 10.35 [More Sufficient Conditions of Unconfoundedness]
Let Assumptions 10.1 (a) to (d) hold. Then

(i) $\Big(D_X \perp\!\!\!\perp 1_{X=x} \vee \big(Y \vDash D_X | (X=x) \wedge \tau_x \text{ is } P\text{-unique}\big)\Big) \quad \Rightarrow \quad E^{X=x}(Y) \Vdash D_X.$

If we additionally assume that Z is a covariate of X i. e., Ass. Assumptions 10.1 (a) [i. e., Ass. 10.1 (e)], then

(ii) $\Big(D_X \perp\!\!\!\perp 1_{X=x} \vee \big(Y \vDash D_X | (X=x) \wedge \tau_x \text{ is } P\text{-unique}\big)\Big) \quad \Rightarrow \quad E^{X=x}(Y|Z) \Vdash D_X.$

If, additionally, $z \in Z(\Omega)$ and $0 < P(X{=}x, Z{=}z) < 1$ [i. e., Ass. 10.1 (f)], then

(iii) $\Big(D_X \perp\!\!\!\perp 1_{X=x} \vee \big(Y \vDash D_X \,|\, (X{=}x) \wedge \tau_x \textit{ is } P\textit{-unique}\big)\Big) \;\Rightarrow\; E^{X=x}(Y \,|\, Z{=}z) \Vdash D_X.$

Hence, under the assumptions of Corollary 10.35, the conditional expectation value $E^{X=x}(Y)$ is unconfounded, if the global potential confounder D_X of X and the indicator $1_{X=x}$ are independent or, if the outcome variable Y is $(X{=}x)$-conditionally independent from D_X [see Eq. (10.28)] and τ_x is P-unique. If we additionally assume that Z is a covariate of X, then the same premises also imply that the conditional expectation $E^{X=x}(Y\,|\,Z)$ is unconfounded. Furthermore, if $0 < P(X{=}x, Z{=}z) < 1$ holds as well, then the conditional expectation value $E^{X=x}(Y\,|\,Z{=}z) = E(Y\,|\,X{=}x, Z{=}z)$ is unconfounded if D_X and the indicator $1_{X=x}$ are independent or if $E^{X=x}(Y\,|\,D_X) \underset{P^{X=x}}{=} E^{X=x}(Y)$ and τ_x is P-unique.

Remark 10.36 [Immediate Implications] According to Corollary 10.35 and Definitions 10.7 (ii) and 10.12 (iii), we can conclude:

(i) Under the Assumptions 10.1 (a) to (d) and (g)
$$\Big(\forall\, x \in X(\Omega)\colon \Big(D_X \perp\!\!\!\perp 1_{X=x} \vee \big(Y \vDash D_X \,|\, (X{=}x) \wedge \tau_x \text{ is } P\text{-unique}\big)\Big)\Big) \;\Rightarrow\; E(Y\,|\,X) \Vdash D_X.$$
(ii) Under the Assumptions 10.1 (a) to (e), and (g),
$$\Big(\forall\, x \in X(\Omega)\colon \Big(D_X \perp\!\!\!\perp 1_{X=x} \vee \big(Y \vDash D_X \,|\, (X{=}x) \wedge \tau_x \text{ is } P\text{-unique}\big)\Big)\Big) \;\Rightarrow\; E(Y\,|\,X, Z) \Vdash D_X$$
[see Def. 10.12 (iii)].

$\triangleleft$

Now we turn to other conditions implying unconfoundedness of the conditional expectations $E^{X=x}(Y\,|\,Z)$ and $E(Y\,|\,X, Z)$. The proposition of the following corollary immediately follows from Theorems 10.22 and 10.29. Reading this corollary, remember

$$D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \quad\Leftrightarrow\quad P(X{=}x \,|\, Z, D_X) \underset{P}{=} P(X{=}x \,|\, Z) \tag{10.43}$$

[see Prop. (8.7)] and

$$Y \vDash D_X \,|\, (X{=}x, Z) \quad\Leftrightarrow\quad E^{X=x}(Y\,|\,Z, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y\,|\,Z) \tag{10.44}$$

[see Eqs. (9.8) and (9.9)].

Corollary 10.37 [Another Sufficient Condition of Unconfoundedness of $E^{X=x}(Y\,|\,Z)$] *Let Assumptions 10.1 (a) to (e) hold and assume that τ_x is P-unique [i. e., Ass. 10.1 (h)]. Then*

$$\big(D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z \;\vee\; Y \vDash D_X \,|\, (X{=}x, Z)\big) \quad\Rightarrow\quad E^{X=x}(Y\,|\,Z) \Vdash D_X. \tag{10.45}$$

Hence, under the assumptions of Corollary 10.37, the disjunction of Z-conditional independence of the indicator $1_{X=x}$ for a value x of X and the global potential confounder

D_X of X and $(X{=}x, Z)$-conditional mean-independence of the outcome variable Y from D_X implies that $E^{X=x}(Y|Z)$ is unconfounded.

The proposition of the next corollary immediately follows from Theorems 10.23 and 10.30. While the premise is the same as in Corollary 10.37, now we only assume that the true outcome variable τ_x is $P^{Z=z}$-unique and $0 < P(X{=}x, Z{=}z) < 1$.

Corollary 10.38 [Another Condition Implying Unconfoundedness of $E^{X=x}(Y|Z{=}z)$]
Let Assumptions 10.1 (a) to (f) hold and assume that τ_x is $P^{Z=z}$-unique [i. e., Ass. 10.1 (j)]. Then

$$\big(D_X \perp\!\!\!\perp 1_{X=x} | Z \,\vee\, Y \vDash D_X | (X{=}x, Z)\big) \quad \Rightarrow \quad E^{X=x}(Y|Z{=}z) \Vdash D_X. \tag{10.46}$$

According to the following corollary, which immediately follows from Theorems 10.23 and 10.32, the disjunction of $D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z)$ and $Y \vDash D_X | (X{=}x, Z{=}z)$ also implies unconfoundedness of the conditional expectation value $E^{X=x}(Y|Z{=}z) = E(Y|X{=}x, Z{=}z)$.

Corollary 10.39 [A Last Condition Implying Unconfoundedness of $E^{X=x}(Y|Z{=}z)$]
Let Assumptions 10.1 (a) to (f) hold and let τ_x be $P^{Z=z}$-unique [i. e., Ass. 10.1 (j)]. Then

$$\big(D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z) \,\vee\, Y \vDash D_X | (X{=}x, Z{=}z)\big) \quad \Rightarrow \quad E^{X=x}(Y|Z{=}z) \Vdash D_X. \tag{10.47}$$

Remark 10.40 [Methodological Consequences] If $0 < P(X{=}x, Z{=}z) < 1$ holds for all pairs (x, z) of values of X and Z, then, according to Corollary 10.39, we can check for each pair (x, z) if the disjunction in Proposition (10.47) holds. Obviously, $D_X \perp\!\!\!\perp 1_{X=x} | (Z{=}z)$ may hold for some pairs (x, z) and $Y \vDash D_X | (X{=}x, Z{=}z)$ for all other pairs of values of X and Z. And still this implies that all conditional expectation values $E^{X=x}(Y|Z{=}z)$, $(x, z) \in X(\Omega) \times Z(\Omega)$, are unconfounded. Remember, according to Proposition (10.20),

$$\big(\forall z \in Z(\Omega)\colon E^{X=x}(Y|Z{=}z) \Vdash D_X\big) \quad \Rightarrow \quad E^{X=x}(Y|Z) \Vdash D_X$$

and according to Definition 10.12 (iii),

$$\big(\forall x \in X(\Omega)\colon E^{X=x}(Y|Z) \Vdash D_X\big) \quad \Leftrightarrow \quad E(Y|X, Z) \Vdash D_X.$$

Hence, checking the premise of Proposition (10.47) for a fixed value x of X and all values z of Z is a way to check unconfoundedness of $E^{X=x}(Y|Z)$. Similarly, checking this premise for all pairs (x, z) of values of X and Z we can check unconfoundedness of $E(Y|X, Z)$, provided, of course, that $0 < P(X{=}x, Z{=}z) < 1$ holds for all pairs of values of X and Z involved (see Example 10.42). $\triangleleft$

10.4 Example

Before we turn to the implications of unconfoundedness on unbiasedness, we illustrate this causality condition by a numerical example. The most important parameters are presented in Table 10.4.

Table 10.4. Unconfoundedness of $E(Y|X,Z)$

		Fundamental parameters							Derived parameters				
Person u	Sex z	$P(U=u)$	$P(X=0\|U=u)$	$P(X=1\|U=u)$	$P(X=2\|U=u)$	$E^{X=0}(Y\|U=u)$	$E^{X=1}(Y\|U=u)$	$E^{X=2}(Y\|U=u)$	$CTE_{U;10}(u)$	$CTE_{U;20}(u)$	$P^{X=0}(U=u)$	$P^{X=1}(U=u)$	$P^{X=2}(U=u)$
Tom	m	1/6	2/6	3/6	1/6	82	100	110	18	28	4/18	3/20	2/14
Tim	m	1/6	2/6	2/6	2/6	89	100	110	11	21	4/18	2/20	4/14
Joe	m	1/6	2/6	3/6	1/6	95	100	110	5	15	4/18	3/20	2/14
Ann	f	1/6	2/12	8/12	2/12	102	113	120	11	18	2/18	4/20	2/14
Sue	f	1/6	1/12	8/12	3/12	102	117	120	15	18	1/18	4/20	3/14
Eva	f	1/6	3/12	8/12	1/12	102	125	120	23	18	3/18	4/20	1/14

	$x=0$	$x=1$	$x=2$		
$E(\tau_x)$:	95.33	109.17	115.00	$ATE_{10} = 13.83$	$ATE_{20} = 19.67$
$E^{X=x}(Y)$:	93.11	111.00	114.29	$PFE_{10} = 17.89$	$PFE_{20} = 21.18$
$E(\tau_x\|Z=m)$:	88.67	100.00	110.00	$CTE_{Z;10}(m) = 11.33$	$CTE_{Z;20}(m) = 21.33$
$E^{X=x}(Y\|Z=m)$:	88.67	100.00	110.00	$PFE_{Z;10}(m) = 11.33$	$PFE_{Z;20}(m) = 21.33$
$E(\tau_x\|Z=f)$:	102.00	118.33	120.00	$CTE_{Z;10}(f) = 16.33$	$CTE_{Z;20}(f) = 18.00$
$E^{X=x}(Y\|Z=f)$:	102.00	118.33	120.00	$PFE_{Z;10}(f) = 16.33$	$PFE_{Z;20}(f) = 18.00$

Again we assume that this random experiment has the same structure as the random experiments treated in section 6.5. That is,

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big),$$

as specified in section 6.5, is the regular probabilistic causality setup. However, now the treatment variable X has three values. In this example, the person variable U again is a global potential confounder of X, that is, $D_X = U$. This implies

$$E(Y|X, D_X) \underset{P}{=} E(Y|X, U) \tag{10.48}$$

and

$$P(X=1|D_X) \underset{P}{=} P(X=1|U). \tag{10.49}$$

If the values u of U represent the observational units *at the onset of treatment*, then these assumptions are realistic also in empirical applications if a simple experiment as described in section 2.1 is considered.

Example 10.41 [$E(Y|X)$ **is Confounded**] In this example, $E(Y|X)$, the X-conditional expectation of Y, *is confounded* because, in this example, it is *not true* that Equation (10.9) is satisfied for each of the three values x of X and each potential confounder of X. Let us consider the potential confounder $Z = sex$. In this case, Z takes the role of W in Equation (10.9). In order to see that Equation (10.9) does not hold we compute both sides of this equation and see if they are identical for all three values $x \in \{0, 1, 2\}$ of the treatment variable X.

The conditional expectation value $E^{X=x}(Y) = E(Y|X=x)$ of the outcome variable Y in treatment x can be computed applying

$$E^{X=x}(Y) \;=\; \sum_z E^{X=x}(Y|Z=z) \cdot P^{X=x}(Z=z). \tag{10.50}$$

This equation is always true if $P(X=x, Z=z) > 0$ for all values z of Z [see RS-Box 3.2 (ii)]. For $x=0$, this equation yields

$$E^{X=0}(Y) \;=\; \sum_z E^{X=0}(Y|Z=z) \cdot P^{X=0}(Z=z) \;=\; 88.\overline{6} \cdot \frac{2}{3} + 102 \cdot \frac{1}{3} \;\approx\; 93.11$$

(see Exercises 10-6, 10-7, and 10-8). In contrast, computing the right-hand side of Equation (10.9) for $x=0$ yields

$$\sum_z E^{X=0}(Y|Z=z) \cdot P(Z=z) \;=\; 88.\overline{6} \cdot \frac{1}{2} + 102 \cdot \frac{1}{2} \;\approx\; 95.33.$$

Because $93.11 \neq 95.33$, unconfoundedness of $E^{X=0}(Y)$ and with it unconfoundedness of $E(Y|X)$ [see Box 10.1 (i), (ii)] does not hold in this example. ◁

Example 10.42 [All $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$ are Unconfounded] Now we check if the conditional expectations $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, and $E(Y|X,Z)$ are unconfounded for the covariate $Z = sex$. For this purpose, we apply Theorem 10.18 and Corollary 10.39 and show that, in this example, the disjunction

$$U \perp\!\!\!\perp 1_{X=x} \,|\, (Z=z) \quad \vee \quad Y \vDash U \,|\, (X=x, Z=z) \tag{10.51}$$

holds for all $3 \cdot 2 = 6$ pairs (x, z) of values of X and Z. Remember, Proposition (10.51) is equivalent to

$$P^{Z=z}(X=x|U) \underset{P^{Z=z}}{=} P^{Z=z}(X=x) \quad \vee \quad E^{X=x,Z=z}(Y|U) \underset{P^{X=x,Z=z}}{=} E^{X=x,Z=z}(Y) \tag{10.52}$$

[see (8.20), (9.3), and (9.4)].

Inspecting Table 10.4 for $(x, z) = (0, m)$ shows that

$$P^{Z=m}(X=0|U)(\omega) \;=\; P^{Z=m}(X=0) \;=\; 2/6, \quad \text{if } \omega \in \{Z=m\}$$

because $P^{Z=m}(X=0|U=u_i) = P(X=0|U=u_i, Z=m) = P^{Z=m}(X=0) = 2/6$ for the three male persons *Tom*, *Tim*, and *Joe*. Furthermore, inspecting the table for $(x, z) = (0, f)$ shows that

$$E^{X=0,Z=f}(Y|U)(\omega) \;=\; E^{X=0,Z=f}(Y) \;=\; 102, \quad \text{if } \omega \in \{X=x, Z=f\}$$

because $E^{X=0,Z=f}(Y|U)(\omega) = E^{X=0}(Y|U=u_i, Z=f) = 102$ for all three female persons *Ann*, *Sue*, and *Eva*. Hence, according to Theorem 10.18, we can conclude that $E^{X=0}(Y|Z)$ is unconfounded.

Checking the pairs $(1, m)$ and $(1, f)$ of values of X and Z in the same way shows that Proposition (10.52) also holds for these two pairs, which proves that the conditional expectation $E^{X=1}(Y|Z)$ is unconfounded as well. Finally, checking Proposition (10.52) for the pairs $(x,z) = (2, m)$ and $(x,z) = (2, f)$ yields the corresponding positive result. This means that $E^{X=2}(Y|Z)$ is unconfounded, too.

Now, according to Definition 10.12 (iii), unconfoundedness of all three conditional expectations $E^{X=x}(Y|Z)$, $x \in X(\Omega) = \{0,1,2\}$, is equivalent to unconfoundedness of $E(Y|X,Z)$. Hence, we can conclude that $E(Y|X,Z)$ is unconfounded as well. $\triangleleft$

10.5 Implications of Unconfoundedness on Unbiasedness

Now we study the implications of unconfoundedness on unbiasedness of various conditional expectations, prima facie effects, and prima facie effect functions. Remember, unbiasedness is the prerequisite for the identification of causal total effects and effect functions (see sect. 6.4).

10.5.1 Unbiasedness of $E(Y|X)$

We start considering the implications of unconfoundedness of $E(Y|X)$ on unbiasedness of conditional expectations and prima facie effects. Reading the following theorem, remember that $\tau_x = E^{X=x}(Y|D_X)$ denotes the true outcome variable (see Def. 5.4), and unbiasedness of $E^{X=x}(Y) = E(Y|X{=}x)$ is defined by τ_x being P-unique and

$$E^{X=x}(Y) \;=\; E(\tau_x) \tag{10.53}$$

(see Def. 6.3).

Theorem 10.43 [Unconfoundedness of $E(Y|X)$ Implies That It Is Unbiased]
Let the Assumptions 10.1 (a) to (d) and (h) hold.

(i) If $E^{X=x}(Y)$ is unconfounded, then it is also unbiased, that is, $E^{X=x}(Y) \Vdash D_X \;\Rightarrow\; E^{X=x}(Y) \vdash D_X$.
(ii) If we additionally assume 10.1 (g) and (i) and $E(Y|X)$ is unconfounded, then it is unbiased. That is, $E(Y|X) \Vdash D_X \;\Rightarrow\; E(Y|X) \vdash D_X$.

(Proof p. 336)

Hence, the conditional expectation value $E^{X=x}(Y) = E(Y|X{=}x)$ and the conditional expectation $E(Y|X)$ are unbiased if they are unconfounded.

Theorem 10.43 has immediate implications on unbiasedness of the prima facie effect

$$PFE_{xx'} \;=\; E^{X=x}(Y) - E^{X=x'}(Y) \;=\; E(Y|X{=}x) - E(Y|X{=}x')\,. \tag{10.54}$$

This is obvious if we consider Equation (10.53) and remember that

$$ATE_{xx'} \;=\; E(\tau_x) - E(\tau_{x'}) \tag{10.55}$$

is the causal average total effect on Y comparing x to x' (see Def. 5.26). Using Equations (10.53) to (10.55) yields

$$PFE_{xx'} \;=\; ATE_{xx'} \;=\; E(\tau_x) - E(\tau_{x'}), \tag{10.56}$$

provided that we assume that $E^{X=x}(Y)$ and $E^{X=x'}(Y)$ are unbiased.

Corollary 10.44 [Implications on Unbiasedness of Prima Facie Effects]
Let the Assumptions 10.1 (a) to (d) hold, let $x' \in X(\Omega)$, $P(X=x') > 0$, and assume that the true outcome variables τ_x and $\tau_{x'}$ are P-unique. Then unconfoundedness of $E^{X=x}(Y)$ and $E^{X=x'}(Y)$ implies that $PFE_{xx'}$ is unbiased.

Hence, if $E^{X=x}(Y)$ and $E^{X=x'}(Y)$ are unconfounded, then we can conclude that the prima facie effect $PFE_{xx'} = E^{X=x}(Y) - E^{X=x'}(Y) = E(Y|X=x) - E(Y|X=x')$ is unbiased.

10.5.2 Unbiasedness of $E(Y|X,Z)$, $E^{X=x}(Y|Z)$, and $E^{X=x}(Y|Z=z)$

Now we turn to the implications of unconfoundedness on unbiasedness of the conditional expectation values $E^{X=x}(Y|Z=z)$ and the conditional expectations $E^{X=x}(Y|Z)$ and $E(Y|X,Z)$.

Theorem 10.45 [Implications on Unbiasedness of Conditional Expectations]
Let the Assumptions 10.1 (a) to (e) and (h) hold.

(i) If $E^{X=x}(Y|Z)$ is unconfounded, then it is unbiased. That is,

$$E^{X=x}(Y|Z) \Vdash D_X \quad \Rightarrow \quad E^{X=x}(Y|Z) \vdash D_X. \tag{10.57}$$

(ii) If we additionally assume 10.1 (i) and $E(Y|X,Z)$ is unconfounded, then it is unbiased. That is,

$$E(Y|X,Z) \Vdash D_X \quad \Rightarrow \quad E(Y|X,Z) \vdash D_X. \tag{10.58}$$

(Proof p. 336)

Theorem 10.45 (i) has immediate implications on unbiasedness of the prima facie effect function

$$PFE_{Z;xx'}(Z) \underset{P}{=} E^{X=x}(Y|Z) - E^{X=x'}(Y|Z). \tag{10.59}$$

This is obvious if we consider

$$E^{X=x}(Y|Z) \underset{P}{=} E(\tau_x|Z), \tag{10.60}$$

which, in conjunction with P-uniqueness of τ_x, defines unbiasedness of $E^{X=x}(Y|Z)$. Remember that

$$E(\tau_x|Z) - E(\tau_{x'}|Z) \underset{P}{=} CTE_{Z;xx'}(Z) \tag{10.61}$$

is a causal Z-conditional total effect variable comparing x to x' [see Def. 5.38 (i)].

Corollary 10.46 [Implications on Unbiasedness of a Prima Facie Effect Function]
Let the Assumptions 10.1 (a) to (e) hold, let $x' \in X(\Omega)$, $P(X{=}x') > 0$, and assume that the true outcome variables τ_x and $\tau_{x'}$ are P-unique. If $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ are unconfounded, then $PFE_{Z;xx'}(Z)$ is unbiased.

Hence, if $E(Y|X,Z)$ or at least the conditional expectations $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ are unconfounded, then we can conclude that the prima facie effect variable $PFE_{xx'}(Z) = E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$ is unbiased.

In our last theorem of this section, we show that unconfoundedness of the conditional expectation value $E^{X=x}(Y|Z{=}z)$ implies that it is also unbiased.

Theorem 10.47 [Implications on Unbiasedness of Conditional Expectations]
Let the Assumptions 10.1 (a) to (f) and (j) hold. If $E^{X=x}(Y|Z{=}z)$ is unconfounded, then it is unbiased, that is,

$$E^{X=x}(Y|Z{=}z) \Vdash D_X \quad \Rightarrow \quad E^{X=x}(Y|Z{=}z) \vdash D_X. \tag{10.62}$$

(Proof p. 337)

Now we consider the implications of Theorem 10.47 on unbiasedness of the $(Z{=}z)$-conditional prima facie effect

$$PFE_{Z;xx'}(z) \ = \ E^{X=x}(Y|Z{=}z) - E^{X=x'}(Y|Z{=}z). \tag{10.63}$$

This is obvious if we consider the equation

$$E^{X=x}(Y|Z{=}z) \ = \ E(\tau_x|Z{=}z), \tag{10.64}$$

which, in conjunction with $P^{Z=z}$-uniqueness of τ_x, defines unbiasedness of the conditional expectation value $E^{X=x}(Y|Z{=}z)$. Also remember that

$$E(\tau_x|Z{=}z) - E(\tau_{x'}|Z{=}z) \ = \ CTE_{Z;xx'}(z) \tag{10.65}$$

is a causal $(Z{=}z)$-conditional total effect comparing x to x' [see Def. 5.38 (ii)]. Hence, if $E(Y|X{=}x, Z{=}z) = E^{X=x}(Y|Z{=}z)$ is unconfounded, then we can conclude that the $(Z{=}z)$-conditional prima facie effect $PFE_{Z;xx'}(z) = E^{X=x}(Y|Z{=}z) - E^{X=x'}(Y|Z{=}z)$ is unbiased.

10.6 Expectation Stability of Prima Facie Effects and Effect Functions

The unconfoundedness conditions have some further consequences that are important for the interpretation of the conditional expectation values $E^{X=x}(Y)$, the conditional expectations $E(Y|X)$, $E^{X=x}(Y|Z)$, $E(Y|X,Z)$, the prima effects $E^{X=x}(Y) - E^{X=x'}(Y)$, and the conditional prima effect variables $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$. In Remark 10.9 we already mentioned expectation stability of the conditional expectation values $E^{X=x}(Y)$. Because

$$E(Y|X) \underset{P}{=} \sum_{x \in X(\Omega)} E^{X=x}(Y) \cdot 1_{X=x}, \tag{10.66}$$

this property also concerns the interpretation of the conditional expectation $E(Y|X)$. Furthermore, in Remark 10.13 we hinted at expectation stability of the conditional expectations $E^{X=x}(Y|Z)$, and because

$$E(Y|X,Z) \underset{P}{=} \sum_{x \in X(\Omega)} E^{X=x}(Y|Z) \cdot 1_{X=x}, \tag{10.67}$$

expectation stability of the conditional expectations $E^{X=x}(Y|Z)$ also concerns the interpretation of $E(Y|X,Z)$.

In this section, we turn to expectation stability of prima facie effects and prima facie effect variables. Remember, in the example described in section 1.1, we had a *negative* prima facie effect PFE_{10} and two *positive* prima facie effects $PFE_{U;10}(Joe)$ and $PFE_{U;10}(Ann)$ for Joe and Ann, respectively. Hence, in that example, PFE_{10} *is not* the expectation of the person-specific effects. However, according to the following theorem, unconfoundedness of $E^{X=x}(Y) = E(Y|X=x)$ and $E^{X=x'}(Y) = E(Y|X=x')$ implies that PFE_{10} *is*, for each potential confounder W of X, the expectation of the W-conditional prima facie effect variable

$$PFE_{W;xx'}(W) \underset{P}{=} E^{X=x}(Y|W) - E^{X=x'}(Y|W) \tag{10.68}$$

over the distribution of W. In the example mentioned above, this includes the potential confounder $W = U$. Correspondingly, unconfoundedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ implies that $PFE_{Z;xx'}(Z)$ is, for each potential confounder W of X, the Z-conditional expectation of the (Z,W)-conditional prima facie effect variable

$$PFE_{Z,W;xx'}(Z,W) \underset{P}{=} E^{X=x}(Y|Z,W) - E^{X=x'}(Y|Z,W). \tag{10.69}$$

Theorem 10.48 [Expectation Stability of Prima Facie Effects and Effect Variables]
Let the Assumptions 10.1 (a) to (d) hold, let $x' \in X(\Omega)$, $P(X=x') > 0$, and assume that the true outcome variables τ_x and $\tau_{x'}$ are P-unique. Finally, let W be a potential confounder of X.

(i) Then $E^{X=x}(Y) \Vdash D_X \wedge E^{X=x'}(Y) \Vdash D_X$ implies

$$PFE_{xx'} = E\big(PFE_{W;xx'}(W)\big). \tag{10.70}$$

(ii) If we additionally assume 10.1 (e), then $E^{X=x}(Y|Z) \Vdash D_X \wedge E^{X=x'}(Y|Z) \Vdash D_X$ implies

$$PFE_{Z;xx'}(Z) \underset{P}{=} E\big(PFE_{Z,W;xx'}(Z,W) \,\big|\, Z\big). \tag{10.71}$$

(Proof p. 337)

Remark 10.49 [Expectation Stability of the Prima Facie Effect $PFE_{xx'}$] According to this theorem, unconfoundedness of $E^{X=x}(Y)$ and $E^{X=x'}(Y)$ implies that each prima facie effect is the expectation of the W-conditional prima facie effect variable over the distribution of the potential confounder W of X. Equation (10.70) is what we intuitively and erroneously expect to hold *in general* and this seems to be the reason why we tend to be astonished about Simpson's paradox where this equation does *not* hold. Note that unconfoundedness of $E^{X=x}(Y)$ and $E^{X=x'}(Y)$ follows from unconfoundedness of $E(Y|X)$. ◁

Remark 10.50 [Expectation Stability of the Prima Facie Effect Variable $PFE_{Z;xx'}(Z)$ **]** According to Theorem 10.48 (ii), unconfoundedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ implies that the Z-conditional prima facie effect variable is the Z-conditional expectation of the (Z, W)-conditional prima facie effect variable, provided that W is a potential confounder of X. Again, note that unconfoundedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ follows from unconfoundedness of $E(Y|X,Z)$. ◁

10.7 Summary and Conclusions

In this chapter we studied another class of sufficient conditions of unbiasedness, called *unconfoundedness* conditions. Together with the additional structural components of a probabilistic causality setup, each of these causality conditions make the difference between an *ordinary mean-dependency* (i. e., a dependency that can be described by a conditional expectation) and a mean-dependency that has a causal meaning. Box 10.1 summarizes the definitions of the unconfoundedness conditions.

Unconfoundedness and the Randomized Experiment

The unconfoundedness conditions are implied by their corresponding Fisher conditions (see Table 10.2) and Suppes-Reichenbach conditions (see Table 10.3), which have been treated in chapters 8 and 9, respectively. Hence, unconfoundedness of $E(Y|X)$ and of $E(Y|X,Z)$ hold in randomized experiments, presuming that X is the treatment variable, Z a covariate of X, and Y the outcome variable. Similarly, unconfoundedness of $E(Y|X,Z)$ holds in randomized and Z-conditionally randomized experiments (see Ths. 10.21, 10.24 and Rems. 8.58, 8.59).

Covariate Selection in the Analysis of Total Effects

To my knowledge, unconfoundedness is the *weakest condition implying unbiasedness that is still empirically falsifiable*. If unconfoundedness does not hold in a particular empirical application, then there is serious doubt that unbiasedness — our weakest causality condition — holds in this application.

In order to falsify unconfoundedness of $E^{X=x}(Y)$ and with it unconfoundedness of $E(Y|X)$, it suffices to find a potential confounder W of X for which the equation

$$E^{X=x}(Y) \;=\; E\big(E^{X=x}(Y|W)\big) \tag{10.72}$$

does not hold (see Def. 10.7). If such a potential confounder of X is identified, then one might rename it $Z_1 = W$ and try to find another potential confounder W of X for which

$$E^{X=x}(Y|Z_1) \underset{P}{=} E\big(E^{X=x}(Y|Z_1,W)\big|Z_1\big) \tag{10.73}$$

does not hold (see Def. 10.12). Again, if such a potential confounder of X is identified, then one may rename it $Z_2 = W$ and try out other potential confounders of X, checking whether or not

$$E^{X=x}(Y|Z_1,Z_2) \underset{P}{=} E\big(E^{X=x}(Y|Z_1,Z_2,W)\big|Z_1,Z_2\big) \tag{10.74}$$

holds. Hence, in quasi-experiments and observational studies, unconfoundedness may serve as a criterion for selecting those covariates in the vector $Z := (Z_1, \ldots, Z_m)$ for which the conditional expectations $E^{X=x}(Y|Z)$, $x \in X(\Omega)$, can hoped for being unconfounded. In these kinds of studies we cannot be sure that unconfoundedness holds. However, we can try to falsify the hypothesis of unconfoundedness of conditional expectation values and conditional expectations.

Expectation Stability

According to Theorem 10.48 and Remark 10.49, unconfoundedness of $E(Y|X)$ implies that each prima facie effect $E^{X=x}(Y) - E^{X=x'}(Y)$ is the expectation of the corresponding W-conditional prima facie effect variable $E^{X=x}(Y|W) - E^{X=x'}(Y|W)$ over the distribution of the potential confounder W of X. Hence, the causal average total effect $E^{X=x}(Y) - E^{X=x'}(Y)$ might be less informative than the causal W-conditional total effect variable $E^{X=x}(Y|W) - E^{X=x'}(Y|W)$, but $E^{X=x}(Y) - E^{X=x'}(Y)$ is not biased, provided that it is unconfounded.

Similarly, unconfoundedness of $E(Y|X,Z)$ implies that the Z-conditional prima facie effect variable $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$ is the Z-conditional expectation of the (Z,W)-conditional prima facie effect variable $E^{X=x}(Y|Z,W) - E^{X=x'}(Y|Z,W)$, provided that W is a potential confounder of X. Again, it might contain less information than the causal (Z,W)-conditional total effect variable $E^{X=x}(Y|Z,W) - E^{X=x'}(Y|Z,W)$, but it is not biased if $E(Y|X,Z)$ is unconfounded.

Finally, note that expectation stability does not only hold for unconfoundedness of $E(Y|X)$ and $E(Y|X,Z)$, but also for all causality conditions that imply unconfoundedness of $E(Y|X)$ and $E(Y|X,Z)$, respectively. In particular, this applies to $D_X \perp\!\!\!\perp X$ and $Y \vDash D_X | X$, each of which implies unconfoundedness of $E(Y|X)$ [see Th. 10.21 (i) and Cor. 10.28 (i)], but also to $D_X \perp\!\!\!\perp X | Z$ and $Y \vDash D_X | (X,Z)$, each of which implies unconfoundedness of $E(Y|X,Z)$ [see Ths. 10.24 and 10.31 (iii)].

10.8 Proofs

Proof of Theorem 10.17

Proposition (10.18). We assume that τ_x is P-unique, which in turn implies that it is also $P^{Z=z}$-unique [see RS-Box 5.1 (v)]. Hence,

$$\begin{aligned}
& E^{X=x}(Y|Z) \Vdash D_X \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\; E^{X=x}(Y|Z) \underset{P}{=} E\big(E^{X=x}(Y|Z,W)\,|\,Z\big) && \text{[Def. 10.12 (i)]} \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\; E^{X=x}(Y|Z{=}z) = E\big(E^{X=x}(Y|Z,W)\,|\,Z{=}z\big) && \text{[RS-(4.19)]} \\
\Rightarrow\quad & E^{X=x}(Y|Z{=}z) \Vdash D_X\,. && \text{[Def. 10.12 (ii)]}
\end{aligned}$$

Proposition (10.19). We assume that all τ_x, $x \in X(\Omega)$, are P-unique and this implies that they are also $P^{Z=z}$-unique [see RS-Box 5.1 (v)]. Hence,

$$\begin{aligned}
E(Y|X,Z) \Vdash D_X \quad & \Leftrightarrow \quad \forall x \in X(\Omega)\colon\; E^{X=x}(Y|Z) \Vdash D_X && \text{[Def. 10.12 (iii)]} \\
& \Rightarrow \quad E^{X=x}(Y|Z{=}z) \Vdash D_X\,. && \text{[(10.18)]}
\end{aligned}$$

Proof of Theorem 10.18

$$
\begin{aligned}
& \forall z \in Z(\Omega)\colon\ E^{X=x}(Y \mid Z{=}z) \Vdash D_X \\
\Leftrightarrow\quad & \forall z \in Z(\Omega), \forall W \in \mathscr{W}_X\colon\ E^{X=x}(Y \mid Z{=}z) = E\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z{=}z\big) \\
& \quad \wedge\ \forall z \in Z(\Omega)\colon \tau_x \text{ is } P^{Z=z}\text{-unique} && [10.12\ \text{(ii)}] \\
\Leftrightarrow\quad & \forall W \in \mathscr{W}_X\colon\ E^{X=x}(Y \mid Z) \underset{P}{=} E\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z\big) \\
& \quad \wedge\ \tau_x \text{ is } P\text{-unique} \\
& && [\forall z \in Z(\Omega)\colon P(Z{=}z) > 0,\ \text{RS-(4.19), Lem. 5.37}] \\
\Leftrightarrow\quad & E^{X=x}(Y \mid Z) \Vdash D_X. && [10.12\ \text{(i)}]
\end{aligned}
$$

Proof of Theorem 10.19

Proposition (i). This is a special case of Proposition (ii) in which $\sigma(Z) = \{\Omega, \emptyset\}$.

Proposition (ii). Note that $D_X \perp\!\!\!\perp 1_{X=x}$ implies that τ_x is P-unique [see Th. 8.25 (ii)]. Furthermore, remember that $\sigma(Z) \subset \sigma(D_X)$ if Z is a covariate of X. Hence,

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \\
\Leftrightarrow\quad & (D_X, Z) \perp\!\!\!\perp 1_{X=x} && [\sigma(Z, D_X) = \sigma(D_X),\ \text{Box 8.1 (i)}] \\
\Rightarrow\quad & D_X \perp\!\!\!\perp 1_{X=x} \mid Z && [\text{RS-Box 6.1 (ix)}] \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\ E^{X=x}(Y \mid Z, W) \perp\!\!\!\perp 1_{X=x} \mid Z \\
& && \big[\sigma\big(E^{X=x}(Y \mid Z, W)\big) \subset \sigma(D_X),\ \text{RS-Box 6.1 (vi)}\big] \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\ E^{X=x}\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z\big) \underset{P^{X=x}}{=} E\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z\big) \\
& && [\{X{=}x\} = \{1_{X=x}{=}1\},\ \text{RS-Th. 6.26}] \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\ E^{X=x}(Y \mid Z) \underset{P^{X=x}}{=} E\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z\big) && [\text{RS-Box 4.1 (xiii)}] \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\ E^{X=x}(Y \mid Z) \underset{P}{=} E\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z\big) \\
& && [\tau_x \text{ is } P\text{-unique} \Rightarrow E^{X=x}(Y \mid Z) \text{ is } P\text{-unique, RS-Box 5.1 (iv)}] \\
\Rightarrow\quad & E^{X=x}(Y \mid Z) \Vdash D_X. && [\text{Def. 10.12 (i)}]
\end{aligned}
$$

Proposition (iii). P-uniqueness of τ_x implies that τ_x is also $P^{Z=z}$-unique [see RS-Box 5.1 (v)]. Furthermore,

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \\
\Rightarrow\quad & E^{X=x}(Y \mid Z) \Vdash D_X && [\text{(ii)}] \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\ E^{X=x}(Y \mid Z) \underset{P}{=} E\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z\big) && [\text{Def. 10.12 (i)}] \\
\Rightarrow\quad & \forall W \in \mathscr{W}_X\colon\ E^{X=x}(Y \mid Z{=}z) = E\big(E^{X=x}(Y \mid Z, W)\,\big|\,Z{=}z\big) \\
& && [\text{RS-Th. 4.27},\ P(X{=}x, Z{=}z) > 0 \Rightarrow P(Z{=}z) > 0] \\
\Rightarrow\quad & E^{X=x}(Y \mid Z{=}z) \Vdash D_X. && [\text{Def. 10.12 (ii)}]
\end{aligned}
$$

Proof of Theorem 10.21

Proposition (i).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; D_X \perp\!\!\!\perp 1_{X=x} && [(8.6)]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon\; E^{X=x}(Y) \Vdash D_X && [\text{Th. } 10.19\ (\text{i})]\\
&\Leftrightarrow \quad E(Y \mid X) \Vdash D_X. && [\text{Def. } 10.7\ (\text{ii})]
\end{aligned}
$$

Proposition (ii).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; D_X \perp\!\!\!\perp 1_{X=x} && [(8.6)]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon\; E^{X=x}(Y \mid Z) \Vdash D_X. && [\text{Th. } 10.19\ (\text{ii})]
\end{aligned}
$$

Proposition (iii).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \quad &\Rightarrow \quad \forall\, x \in X(\Omega)\colon\; E^{X=x}(Y \mid Z) \Vdash D_X && [(\text{ii})]\\
&\Leftrightarrow \quad E(Y \mid X, Z) \Vdash D_X. && [\text{Def. } 10.12\ (\text{iii})]
\end{aligned}
$$

Proof of Theorem 10.22

The assumptions include that τ_x is P-unique [see Ass. 10.1 (h)] and this implies that $E^{X=x}(Y \mid Z)$ is P-unique as well [see RS-Box 5.1 (iv)]. Hence,

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \mid Z\\
\Rightarrow\quad & \forall\, W \in \mathscr{W}_X\colon\; E^{X=x}(Y \mid Z, W) \perp\!\!\!\perp 1_{X=x} \mid Z\\
& \hspace{12em} [\sigma\big(E^{X=x}(Y \mid Z, W)\big) \subset \sigma(D_X),\ \text{RS-Box } 6.1\ (\text{vi})]\\
\Rightarrow\quad & \forall\, W \in \mathscr{W}_X\colon\; E^{X=x}\big(E^{X=x}(Y \mid Z, W) \,\big|\, Z\big) \underset{P^{X=x}}{=} E\big(E^{X=x}(Y \mid Z, W) \,\big|\, Z\big)\\
& \hspace{12em} [\{X{=}x\} = \{1_{X=x}{=}1\},\ \text{RS-Th. } 6.26]\\
\Rightarrow\quad & \forall\, W \in \mathscr{W}_X\colon\; E^{X=x}(Y \mid Z) \underset{P^{X=x}}{=} E\big(E^{X=x}(Y \mid Z, W) \,\big|\, Z\big) \qquad [\text{RS-Box } 4.1\ (\text{xiii})]\\
\Rightarrow\quad & \forall\, W \in \mathscr{W}_X\colon\; E^{X=x}(Y \mid Z) \underset{P}{=} E\big(E^{X=x}(Y \mid Z, W) \,\big|\, Z\big)\\
& \hspace{12em} [E^{X=x}(Y \mid Z) \text{ is } P\text{-unique}]\\
\Rightarrow\quad & E^{X=x}(Y \mid Z) \Vdash D_X. \qquad [\tau_x \text{ is } P\text{-unique, Def. } 10.12\ (\text{i})]
\end{aligned}
$$

Proof of Theorem 10.23

Proposition (10.22). This implication immediately follows from RS-Propositions (6.47) and (6.48).

Proposition (10.23). According to RS-Proposition (6.47),

$$
D_X \perp\!\!\!\perp 1_{X=x} \mid (Z{=}z) \quad :\Leftrightarrow \quad D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x}\,.
$$

Hence, we can use all properties and implications of independence of two random variables (with respect to a measure P), simply replacing the measure P by the measure $P^{Z=z}$. Therefore, according to Theorem 8.25 (ii), $D_X \perp\!\!\!\perp 1_{X=x} \mid (Z{=}z)$ implies that $\tau_x = E^{X=x}(Y \mid D_X)$ is $P^{Z=z}$-unique. Furthermore,

$$
\begin{aligned}
& D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z) \\
\Leftrightarrow\;& D_X \underset{P^{Z=z}}{\perp\!\!\!\perp} 1_{X=x} && [\text{RS-}(6.47)] \\
\Rightarrow\;& E^{Z=z}(Y \,|\, X{=}x) \Vdash D_X && [\tau_x \text{ is } P^{Z=z}\text{-unique, Th. } 10.19 \text{ (i)}] \\
\Leftrightarrow\;& E^{X=x}(Y \,|\, Z{=}z) \Vdash D_X\,. && [\text{RS-}(5.26)]
\end{aligned}
$$

Proof of Theorem 10.24

$$
\begin{aligned}
& D_X \perp\!\!\!\perp X \,|\, Z \\
\Leftrightarrow\;& \forall\, x \in X(\Omega)\colon D_X \perp\!\!\!\perp 1_{X=x} \,|\, Z && [(8.12)] \\
\Rightarrow\;& \forall\, x \in X(\Omega)\colon E^{X=x}(Y \,|\, Z) \Vdash D_X && [\tau_x \text{ is } P\text{-unique, Th. } 10.22] \\
\Leftrightarrow\;& E(Y \,|\, X, Z) \Vdash D_X\,. && [\text{Def. } 10.12 \text{ (iii)}]
\end{aligned}
$$

Proof of Theorem 10.25

Proposition (10.26). This is Proposition (8.19).

Proposition (10.27).

$$
\begin{aligned}
D_X \perp\!\!\!\perp X \,|\, (Z{=}z) \;&\Leftrightarrow\; \forall\, x \in X(\Omega)\colon D_X \perp\!\!\!\perp 1_{X=x} \,|\, (Z{=}z) && [(8.26)] \\
&\Rightarrow\; E^{X=x}(Y \,|\, Z{=}z) \Vdash D_X\,. && [\text{Th. } 10.23]
\end{aligned}
$$

Proof of Theorem 10.26

Proposition (i).

$$
\begin{aligned}
& Y \vDash D_X \,|\, (X{=}x) \\
\Leftrightarrow\;& E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(10.28)] \\
\Rightarrow\;& \forall\, W \in \mathscr{W}_X\colon\; E^{X=x}(Y \,|\, W) \underset{P^{X=x}}{=} E^{X=x}(Y) && [\forall W\colon \sigma(W) \subset \sigma(D_X), \text{RS-Th. } 4.39] \\
\Rightarrow\;& \forall\, W \in \mathscr{W}_X\colon\; E^{X=x}(Y \,|\, W) \underset{P}{=} E^{X=x}(Y) \\
& && [\forall W\colon \sigma(W) \subset \sigma(D_X), E^{X=x}(Y \,|\, W) \text{ is } P\text{-unique, RS-Box } 5.1 \text{ (iv)}] \\
\Rightarrow\;& \forall\, W \in \mathscr{W}_X\colon\; E\big(E^{X=x}(Y \,|\, W)\big) = E^{X=x}(Y) && [\text{RS-Box } 3.1 \text{ (i), (v)}] \\
\Leftrightarrow\;& E^{X=x}(Y) \Vdash D_X\,. && [\tau_x \text{ is } P\text{-unique, Def. } 10.7 \text{ (i)}]
\end{aligned}
$$

Proposition (ii).

$$
\begin{aligned}
& Y \vDash D_X \,|\, (X{=}x) \\
\Leftrightarrow\;& E^{X=x}(Y \,|\, D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(10.28)]
\end{aligned}
$$

$$\begin{aligned}
&\Rightarrow \quad E^{X=x}(Y|Z) \underset{P^{X=x}}{=} E^{X=x}(Y) && [\sigma(Z) \subset \sigma(D_X), \text{RS-Th. } 4.39]\\
&\Rightarrow \quad E^{X=x}(Y|Z) \underset{P}{=} E^{X=x}(Y) \\
& && [\sigma(Z) \subset \sigma(D_X),\ E^{X=x}(Y|Z) \text{ is } P\text{-unique, RS-Box } 5.1\ (iv)]\\
&\Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon\ E^{X=x}(Y|Z) \underset{P}{=} E\big(E^{X=x}(Y|Z,W)\,\big|\,Z\big) \\
& && [\forall\, W \in \mathscr{W}_X\colon \sigma(Z,W) \subset \sigma(D_X),\ (10.31)]\\
&\Leftrightarrow \quad E^{X=x}(Y|Z) \Vdash D_X. && [\tau_x \text{ is } P\text{-unique, Def. } 10.12\ (i)]
\end{aligned}$$

Proposition (iii). Assuming $0 < P(X{=}x, Z{=}z) < 1$ implies $P^{X=x}(Z{=}z), P(Z{=}z) > 0$. Furthermore, we assume that $E^{X=x}(Y|D_X) = \tau_x$ is $P^{Z=z}$-unique. According to RS-Box 5.1 (v), this implies that τ_x is also $P^{X=x,Z=z}$-unique, which in turn implies that the conditional expectation value $E^{X=x,Z=z}(\tau_x) = E^{X=x}(\tau_x\,|\,Z{=}z)$ is uniquely defined. Hence,

$$\begin{aligned}
& \quad Y \vDash D_X|(X{=}x) \\
&\Leftrightarrow \quad E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(10.28)]\\
&\Rightarrow \quad E^{X=x}(Y|Z) \underset{P^{X=x}}{=} E^{X=x}(Y) && [\sigma(Z) \subset \sigma(D_X), \text{RS-Th. } 4.39]\\
&\Rightarrow \quad E^{X=x}(Y|Z{=}z) = E^{X=x}(Y) && [\text{RS-Th. } 4.37\ (ii)]\\
&\Rightarrow \quad \forall\, W \in \mathscr{W}_X\colon\ E^{X=x}(Y|Z{=}z) = E\big(E^{X=x}(Y|Z,W)\,\big|\,Z{=}z\big) && [(10.29)]\\
&\Leftrightarrow \quad E^{X=x}(Y|Z{=}z) \Vdash D_X. && [\tau_x \text{ is } P\text{-unique, Def. } 10.12\ (ii)]
\end{aligned}$$

Proof of Corollary 10.27

Proposition (i).

$$\begin{aligned}
Y \vDash D_X|X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(9.5)]\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon Y \vDash D_X|(X{=}x) && [(10.28)]\\
&\Rightarrow \quad E^{X=x}(Y) \Vdash D_X. && [\text{Th. } 10.26\ (i)]
\end{aligned}$$

Proposition (ii).

$$\begin{aligned}
Y \vDash D_X|X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(9.5)]\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon Y \vDash D_X|(X{=}x) && [(10.28)]\\
&\Rightarrow \quad E^{X=x}(Y|Z) \Vdash D_X. && [\text{Th. } 10.26\ (ii)]
\end{aligned}$$

Proposition (iii).

$$\begin{aligned}
Y \vDash D_X|X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(9.5)]\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon Y \vDash D_X|(X{=}x) && [(10.28)]\\
&\Rightarrow \quad E^{X=x}(Y|Z{=}z) \Vdash D_X. && [\text{Th. } 10.26\ (iii)]
\end{aligned}$$

Proof of Corollary 10.28

Proposition (i).

$$\begin{aligned}
Y \vDash D_X | X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(9.5)]\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon Y \vDash D_X | (X=x) && [(10.28)]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y) \Vdash D_X && [\text{Th. } 10.26 \text{ (i)}]\\
&\Leftrightarrow \quad E(Y|X) \Vdash D_X. && [\text{Def. } 10.7 \text{ (ii)}]
\end{aligned}$$

Propositions (ii) and (iii).

$$\begin{aligned}
Y \vDash D_X | X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) && [(9.5)]\\
&\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon Y \vDash D_X | (X=x) && [(10.28)]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y|Z) \Vdash D_X && [\text{Th. } 10.26 \text{ (ii)}]\\
&\Leftrightarrow \quad E(Y|X,Z) \Vdash D_X. && [\text{Def. } 10.12 \text{ (iii)}]
\end{aligned}$$

Proof of Theorem 10.29

$$\begin{aligned}
& Y \vDash D_X | (X=x, Z)\\
\Leftrightarrow \quad & E^{X=x}(Y|Z) \underset{P^{X=x}}{=} E^{X=x}(Y|Z, D_X) && [(10.33)]\\
\Rightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y|Z) \underset{P^{X=x}}{=} E^{X=x}(Y|Z, W) && [\sigma(W) \subset \sigma(D_X), \text{ RS-Th. } 4.43]\\
\Rightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y|Z) \underset{P}{=} E^{X=x}(Y|Z, W) && [E^{X=x}(Y|Z) \text{ is } P\text{-unique}]\\
\Rightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E\big(E^{X=x}(Y|Z) \,\big|\, Z\big) \underset{P}{=} E\big(E^{X=x}(Y|Z, W) \,\big|\, Z\big) && [\text{RS-Box } 4.1 \text{ (xiv)}]\\
\Rightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y|Z) \underset{P}{=} E\big(E^{X=x}(Y|Z, W) \,\big|\, Z\big) && [\text{RS-Box } 4.1 \text{ (xi)}]\\
\Leftrightarrow \quad & E^{X=x}(Y|Z) \Vdash D_X. && [\tau_x \text{ is } P\text{-unique, Def. } 10.12 \text{ (i)}]
\end{aligned}$$

Proof of Theorem 10.30

$$\begin{aligned}
& Y \vDash D_X | (X=x, Z)\\
\Leftrightarrow \quad & E^{X=x}(Y|Z) \underset{P^{X=x}}{=} E^{X=x}(Y|Z, D_X) && [(10.33)]\\
\Rightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y|Z) \underset{P^{X=x}}{=} E^{X=x}(Y|Z, W) && [\sigma(W) \subset \sigma(D_X), \text{ RS-Th. } 4.43]\\
\Rightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E\big(E^{X=x}(Y|Z) \,\big|\, Z=z\big) = E\big(E^{X=x}(Y|Z, W) \,\big|\, Z=z\big)\\
& && [E^{X=x}(Y|Z, W) \text{ is } P^{Z=z}\text{-unique, RS-Box } 3.1 \text{ (v)}]\\
\Leftrightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y|Z=z) = E\big(E^{X=x}(Y|Z, W) \,\big|\, Z=z\big) && [\text{RS-Box } 3.2 \text{ (i)}]\\
\Leftrightarrow \quad & E^{X=x}(Y|Z=z) \Vdash D_X. && [\tau_x \text{ is } P^{Z=z}\text{-unique, Def. } 10.12 \text{ (ii)}]
\end{aligned}$$

Proof of Theorem 10.31

Proposition (i).

$$
\begin{aligned}
Y \vDash D_X | (X,Z) \quad &\Leftrightarrow \quad E(Y|X,Z,D_X) \underset{P}{=} E(Y|X,Z) && [(10.37)]\\
&\Leftrightarrow \quad \forall x \in X(\Omega)\colon E^{X=x}(Y|Z,D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) && [(9.15)]\\
&\Leftrightarrow \quad \forall x \in X(\Omega)\colon Y \vDash D_X | (X=x,Z) && [\text{Box 9.1 (vi)}]\\
&\Rightarrow \quad E^{X=x}(Y|Z) \Vdash D_X. && [\text{Th. 10.29}]
\end{aligned}
$$

Proposition (ii).

$$
\begin{aligned}
Y \vDash D_X | (X,Z) \quad &\Leftrightarrow \quad E(Y|X,Z,D_X) \underset{P}{=} E(Y|X,Z) && [(10.37)]\\
&\Leftrightarrow \quad \forall x \in X(\Omega)\colon E^{X=x}(Y|Z,D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z) && [(9.15)]\\
&\Leftrightarrow \quad \forall x \in X(\Omega)\colon Y \vDash D_X | (X=x,Z) && [\text{Box 9.1 (vi)}]\\
&\Rightarrow \quad \forall x \in X(\Omega)\colon E^{X=x}(Y|Z) \Vdash D_X. && [\text{Th. 10.29}]
\end{aligned}
$$

Proposition (iii). According to Definition 10.12 (iii), this proposition immediately follows from (ii).

Proof of Theorem 10.32

Note that, according to RS-Box 5.1 (iv), $P^{Z=z}$-uniqueness of τ_x implies that, for all $W \in \mathscr{W}_X$, $E^{X=x}(Y|Z,W)$ is $P^{Z=z}$-unique because all (Z,W), $W \in \mathscr{W}_X$, are (Z,D_X)-measuable. Now,

$$
\begin{aligned}
& Y \vDash D_X | (X=x, Z=z)\\
\Leftrightarrow \quad & E^{X=x,Z=z}(Y) \underset{P^{X=x,Z=z}}{=} E^{X=x,Z=z}(Y|D_X) && [\text{RS-(5.43)}]\\
\Rightarrow \quad & \forall W \in \mathscr{W}_X\colon E^{X=x,Z=z}(Y) \underset{P^{X=x,Z=z}}{=} E^{X=x,Z=z}(Y|Z,W)\\
& && [\sigma(Z,W) \subset \sigma(D_X), \text{RS-Th. 4.43}]\\
\Leftrightarrow \quad & \forall W \in \mathscr{W}_X\colon E^{X=x,Z=z}(Y) \underset{P^{X=x,Z=z}}{=} E^{X=x}(Y|Z,W) && [\text{SN-(14.72)}]\\
\Leftrightarrow \quad & \forall W \in \mathscr{W}_X\colon E^{X=x,Z=z}(Y) \underset{P^{Z=z}}{=} E^{X=x}(Y|Z,W)\\
& && [E^{X=x}(Y|Z,W) \text{ is } P^{Z=z}\text{-unique}]\\
\Rightarrow \quad & \forall W \in \mathscr{W}_X\colon E\big(E^{X=x,Z=z}(Y)\,\big|\,Z=z\big) = E\big(E^{X=x}(Y|Z,W)\,\big|\,Z=z\big)\\
& && [\text{RS-(3.24), RS-Box 3.1 (i)}, E^{X=x}(Y|Z,W) \text{ is } P^{Z=z}\text{-unique}]\\
\Rightarrow \quad & \forall W \in \mathscr{W}_X\colon E^{X=x}(Y|Z=z) = E\big(E^{X=x}(Y|Z,W)\,\big|\,Z=z\big)\\
& && [\text{RS-Box 3.1 (i), RS-(5.26)}]\\
\Leftrightarrow \quad & E^{X=x}(Y|Z=z) \Vdash D_X. && [\tau_x \text{ is } P^{Z=z}\text{-unique, Def. 10.12 (ii)}]
\end{aligned}
$$

Proof of Corollary 10.33

$$
\begin{aligned}
Y \vDash D_X \,|\, (X,Z) \quad &\Leftrightarrow \quad E(Y\,|\,X,Z) \underset{P}{=} E(Y\,|\,X,Z,D_X) && [(9.13), (9.14)]\\
&\Rightarrow \quad E^{X=x,\,Z=z}(Y) \underset{P^{X=x,Z=z}}{=} E^{X=x,\,Z=z}(Y\,|\,D_X) && [\text{SN-}(14.82)]\\
&\Leftrightarrow \quad Y \vDash D_X \,|\, (X{=}x, Z{=}z) && [\text{RS-}(5.43)]\\
&\Rightarrow \quad E^{X=x}(Y\,|\,Z{=}z) \Vdash D_X\,. && [(10.39)]
\end{aligned}
$$

Proof of Theorem 10.43

Proposition (i). If $E^{X=x}(Y)$ is unconfounded, then $P(X{=}x) > 0$ and $\tau_x = E^{X=x}(Y\,|\,D_X)$ is P-unique. Hence,

$$
\begin{aligned}
& E^{X=x}(Y) \Vdash D_X \\
\Leftrightarrow\quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y) = E\big(E^{X=x}(Y\,|\,W)\big) && [(10.6)]\\
\Rightarrow\quad & E^{X=x}(Y) = E\big(E^{X=x}(Y\,|\,D_X)\big) && [D_X \in \mathscr{W}_X]\\
\Leftrightarrow\quad & E^{X=x}(Y) = E(\tau_x) && [\text{Def. } 5.4]\\
\Leftrightarrow\quad & E(Y\,|\,X{=}x) \vdash D_X && [\text{Def. } 6.3\ (\text{i})]\\
\Leftrightarrow\quad & E^{X=x}(Y) \vdash D_X\,. && [\text{RS-}(3.24)]
\end{aligned}
$$

Proposition (ii).

$$
\begin{aligned}
E(Y\,|\,X) \Vdash D_X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y) \Vdash D_X && [\text{Def. } 10.7\ (\text{ii})]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y) \vdash D_X\,. && [(\text{i})]
\end{aligned}
$$

Proof of Theorem 10.45

Proposition (i). If $E^{X=x}(Y\,|\,Z)$ is unconfounded, then, according to Definition 10.12 (i), τ_x is P-unique. Hence,

$$
\begin{aligned}
& E^{X=x}(Y\,|\,Z) \Vdash D_X \\
\Leftrightarrow\quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y\,|\,Z) = E\big(E^{X=x}(Y\,|\,Z,W)\,\big|\,Z\big) && [(10.11)]\\
\Rightarrow\quad & E^{X=x}(Y\,|\,Z) = E\big(E^{X=x}(Y\,|\,Z,D_X)\,\big|\,Z\big) && [D_X \in \mathscr{W}_X]\\
\Leftrightarrow\quad & E^{X=x}(Y\,|\,Z) = E(\tau_x\,|\,Z) && [\sigma(Z) \subset \sigma(D_X), \text{Def. } 5.4]\\
\Leftrightarrow\quad & E^{X=x}(Y\,|\,Z) \vdash D_X\,. && [\text{Def. } 6.13\ (\text{i})]
\end{aligned}
$$

Proposition (ii).

$$
\begin{aligned}
E(Y\,|\,X,Z) \Vdash D_X \quad &\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y\,|\,Z) \Vdash D_X && [\text{Def. } 10.12\ (\text{iii})]\\
&\Rightarrow \quad \forall\, x \in X(\Omega)\colon E^{X=x}(Y\,|\,Z) \vdash D_X && [(\text{i})]\\
&\Leftrightarrow \quad E(Y\,|\,X,Z) \vdash D_X\,. && [\text{Def. } 6.18\ (\text{i})]
\end{aligned}
$$

Proof of Theorem 10.47

If the conditional expectation value $E^{X=x}(Y|Z=z)$ is unconfounded, then, according to Definition 10.12 (ii), $0 < P(X=x, Z=z) < 1$ and τ_x is $P^{Z=z}$-unique. Hence,

$$
\begin{aligned}
& E^{X=x}(Y|Z=z) \Vdash D_X \\
\Leftrightarrow \quad & \forall\, W \in \mathscr{W}_X\colon E^{X=x}(Y|Z=z) = E\big(E^{X=x}(Y|Z,W)\,\big|\,Z=z\big) && \text{[Def. 10.12 (ii)]} \\
\Rightarrow \quad & E^{X=x}(Y|Z=z) = E\big(E^{X=x}(Y|Z,D_X)\,\big|\,Z=z\big) && [D_X \in \mathscr{W}_X] \\
\Leftrightarrow \quad & E^{X=x}(Y|Z=z) = E(\tau_x\,|\,Z=z) && [\sigma(Z) \subset \sigma(D_X), \text{Def. 5.4}] \\
\Leftrightarrow \quad & E^{X=x}(Y|Z=z) \vdash D_X. && \text{[Def. 6.13 (ii)]}
\end{aligned}
$$

Proof of Theorem 10.48

Proposition (i) is a special case of (ii), Z being a constant. Hence we only prove (ii), using the definition of unconfoundedness of $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ [see Eq. (10.11)].

$$
\begin{aligned}
& PFE_{Z;xx'}(Z) \\
\underset{P}{=}\; & E^{X=x}(Y|Z) - E^{X=x'}(Y|Z) && [\text{def. of } PFE_{Z;xx'}(Z)] \\
\underset{P}{=}\; & E\big(E^{X=x}(Y|Z,W)\,\big|\,Z\big) - E\big(E^{X=x'}(Y|Z,W)\,\big|\,Z\big) && [(10.11)] \\
\underset{P}{=}\; & E\big(E^{X=x}(Y|Z,W) - E^{X=x'}(Y|Z,W)\,\big|\,Z\big) && \text{[RS-Box 4.1 (xviii)]} \\
\underset{P}{=}\; & E\big(PFE_{Z,W;xx'}(Z,W)\,\big|\,Z\big). && [\text{def. of } PFE_{Z,W;xx'}(Z,W)]
\end{aligned}
$$

10.9 Exercises

▷ **Exercise 10-1** Consider the example presented in section 1.1. Using Tables 1.3 and 1.4 show that $E^{X=0}(Y) = E\big(E^{X=0}(Y|W)\big)$ does not hold in that example for the person variable $W = U$.

▷ **Exercise 10-2** Prove that, under the assumptions of Remark 10.14, Equations (10.11) and (10.13) are equivalent to each other.

▷ **Exercise 10-3** Check if and where the implications listed in Table 10.1 have been proven in this chapter. Use and specify the appropriate choice of assumptions listed in the Assumptions 10.1.

▷ **Exercise 10-4** Let the Assumptions 10.1 (a) to (f) and (j) hold, and let W be potential confounders of X. Prove Proposition (10.29).

▷ **Exercise 10-5** Under the Assumptions 10.1 (a) to (e) and (h) and that W is a potential confounder of X, prove Proposition (10.31).

▷ **Exercise 10-6** Compute the probabilities $P(U=u_i\,|\,X=0, Z=m)$ and $P(U=u_i\,|\,X=0, Z=f)$ in the example presented in Table 10.4.

▷ **Exercise 10-7** Compute the conditional expectations $E(Y\,|\,X=0, Z=z)$ in the example presented in Table 10.4.

▷ **Exercise 10-8** Compute the conditional probabilities $P(Z{=}z\,|\,X{=}0)$ in the example presented in Table 10.4 for the covariate $Z := sex$.

▷ **Exercise 10-9** Define a map h such that the variable $Z := sex$ in the example of Table 10.4 is the composition $h(U)$ of h and U. [An alternative notation for $h(U)$ is $h \circ U$.]

▷ **Exercise 10-10** Using the example of Table 10.4, give four examples of a variable $W{=}f(U)$ and write down the associated four partitions of the total set of observational units $\Omega_U := \{u_1, u_2, \dots, u_6\}$. (A *partition* of the set Ω_U is a set of subsets of Ω_U such that the union of these subsets is Ω_U, the subsets are disjunct, that is, they do not intersect, and the empty set is not an element in the partition.)

▷ **Exercise 10-11** How many different partitions of the set of units $\Omega_U := \{u_1, u_2, \dots, u_6\}$ — excluding the partition $\{\Omega_U\}$ — exist in the example of Table 10.4?

▷ **Exercise 10-12** Show that Z-*conditional independence* of X and U does not hold in the example of Table 10.4.

▷ **Exercise 10-13** Show that $Y \vDash D_X|(X,Z)$ with $Z := sex$ does not hold in the example of Table 10.4.

Solutions

▷ **Solution 10-1** In the example of section 1.1, the response or outcome variable Y is binary. Therefore, using the probabilities displayed in Table 1.3 yields

$$E^{X=0}(Y) \;=\; P^{X=0}(Y{=}1) \;=\; P(Y{=}1\,|\,X{=}0) \;=\; \frac{P(Y{=}1, X{=}0)}{P(X{=}0)} \;=\; \frac{.36}{.6} \;=\; .6$$

Furthermore, using the probabilities displayed in Table 1.4 and the person variable $W{=}U$, we obtain

$$\begin{aligned} E\big(E^{X=0}(Y\,|\,U)\big) &= E\big(P^{X=0}(Y{=}1\,|\,U)\big) \\ &= P^{X=0}(Y{=}1\,|\,U{=}Joe)\cdot P(U{=}Joe) \;+\; P^{X=0}(Y{=}1\,|\,U{=}Ann)\cdot P(U{=}Ann) \\ &= P(Y{=}1\,|\,X{=}0, U{=}Joe)\cdot P(U{=}Joe) \;+\; P(Y{=}1\,|\,X{=}0, U{=}Ann)\cdot P(U{=}Ann) \\ &= \frac{.336}{.48}\cdot .5 \;+\; \frac{.024}{.12}\cdot .5 \;=\; .45. \end{aligned}$$

Hence, $E^{X=0}(Y) = E\big(E^{X=0}(Y\,|\,W)\big)$ does not hold in this example for $W{=}\,U$. Therefore, in this example, $E^{X=0}(Y)$ is neither unconfounded nor unbiased [see Defs. 10.7 (i) and 6.3 (i)], and this implies that $E(Y\,|\,X)$ is neither unconfounded nor unbiased [see Defs. 10.7 (ii) and 6.3 (ii)].

▷ **Solution 10-2** If Z is a covariate of X, W a potential confounder of X, and τ_x is P-unique, then $\sigma(Z), \sigma(Z,W) \subset \sigma(D_X)$, and $E^{X=x}(Y|Z)$ and $E^{X=x}(Y|Z,W)$ are P-unique, as well. Hence,

$$\begin{aligned} & E^{X=x}(Y|Z) \underset{P}{=} E\big(E^{X=x}(Y\,|\,Z,W)\,\big|\,Z\big) \\ \Leftrightarrow\quad & E^{X=x}(Y|Z{=}z) = E\big(E^{X=x}(Y\,|\,Z,W)\,\big|\,Z{=}z\big), \quad \text{for } P_Z\text{-a.a. } z \in \Omega'_Z \qquad \text{[RS-(4.19)]} \\ \Leftrightarrow\quad & E^{X=x}(Y|Z{=}z) = \int E^{X=x}(Y\,|\,Z{=}z, W{=}w)\,P_{W|Z=z}(dw), \quad \text{for } P_Z\text{-a.a. } z \in \Omega'_Z. \qquad \text{[SN-Th. 17.60]} \end{aligned}$$

▷ **Solution 10-3** We check the implications listed in Table 10.1 row-wise.

(1) Under the Assumptions 10.1 (a) to (d), (g), and (i): $E(Y\,|\,X) \Vdash D_X \;\Rightarrow\; E^{X=x}(Y) \Vdash D_X$. This is Proposition (10.16).

(2) Under the Assumptions 10.1 (a) to (f) and (h): $E^{X=x}(Y|Z) \Vdash D_X \Rightarrow E^{X=x}(Y|Z{=}z) \Vdash D_X$. This is Proposition (10.18).

(3) Under the Assumptions 10.1 (a) to (g) and (i): $E(Y|X,Z) \Vdash D_X \Rightarrow E^{X=x}(Y|Z{=}z) \Vdash D_X$. This is Proposition (10.19)

(4) Under the Assumptions 10.1 (a) to (e), (g), and (i): $E^{X=x}(Y|Z{=}z) \Vdash D_X \Rightarrow E^{X=x}(Y|Z) \Vdash D_X$. This is Proposition (10.17).

▷ **Solution 10-4** We assume that Z is a covariate and W a potential confounder of X, and that $\tau_x = E^{X=x}(Y|D_X)$ is $P^{Z=z}$-unique. This implies $\sigma(Z,W) \subset \sigma(D_X)$ and that $E^{X=x}(Y|Z,W)$ is $P^{Z=z}$-unique as well [see RS-Box 5.1 (iv)]. Assuming $0 < P(X{=}x, Z{=}z) < 1$ implies $P^{X=x}(Z{=}z) > 0$ and that the $(Z{=}z)$-conditional expectation value of $E^{X=x}(Y|Z,W)$ is uniquely defined [see RS-Box 5.1 (vii)]. Hence,

$$
\begin{aligned}
& E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y) \\
\Rightarrow\quad & E^{X=x}(Y|Z,W) \underset{P^{X=x}}{=} E^{X=x}(Y) && [\sigma(Z,W) \subset \sigma(D_X), \text{RS-Th. } 4.39] \\
\Rightarrow\quad & E\big(E^{X=x}(Y|Z,W)\,\big|\,Z{=}z\big) = E\big(E^{X=x}(Y)\,\big|\,Z{=}z\big) && [E^{X=x}(Y|Z,W) \text{ is } P^{Z=z}\text{-unique}] \\
\Rightarrow\quad & E\big(E^{X=x}(Y|Z,W)\,\big|\,Z{=}z\big) = E^{X=x}(Y). && [\text{RS-Box } 3.1 \text{ (i)}]
\end{aligned}
$$

▷ **Solution 10-5** The assumptions that $\tau_x = E^{X=x}(Y|D_X)$ is P-unique, that Z is a covariate and W a potential confounder of X imply $\sigma(Z,W) \subset \sigma(D_X)$ and that $E^{X=x}(Y|Z,W)$ is P-unique as well [see RS-Box 5.1 (iv)]. Hence,

$$
\begin{aligned}
& E^{X=x}(Y|D_X) \underset{P}{=} E^{X=x}(Y) \\
\Rightarrow\quad & E^{X=x}(Y|Z,W) \underset{P}{=} E^{X=x}(Y) && [\sigma(Z,W) \subset \sigma(D_X), \text{RS-Th. } 4.39] \\
\Rightarrow\quad & E\big(E^{X=x}(Y|Z,W)\,\big|\,Z\big) \underset{P}{=} E\big(E^{X=x}(Y)\,\big|\,Z\big) && [\text{RS-Box } 4.1 \text{ (xiv)}] \\
\Rightarrow\quad & E\big(E^{X=x}(Y|Z,W)\,\big|\,Z\big) \underset{P}{=} E^{X=x}(Y). && [\text{RS-Box } 4.1 \text{ (i)}]
\end{aligned}
$$

▷ **Solution 10-6** For $Z{=}m$ and units u_1, u_2 and u_3, these conditional probabilities can be computed as follows:

$$
\begin{aligned}
P(U{=}u_i\,|\,X{=}0, Z{=}m) &= \frac{P(X{=}0\,|\,Z{=}m, U{=}u_i)\cdot P(Z{=}m, U{=}u_i)}{P(X{=}0\,|\,Z{=}m)\cdot P(Z{=}m)} \\
&= \frac{2/6 \cdot 1/6}{2/6 \cdot 1/2} = \frac{1}{3},
\end{aligned}
$$

For the units u_4, u_5 and u_6, these conditional probabilities are $P(U{=}u_i\,|\,X{=}0, Z{=}m) = 0$.

For $Z{=}f$, these conditional probabilities are:

$$
\begin{aligned}
P(U{=}u_4\,|\,X{=}0, Z{=}f) &= \frac{P(X{=}0\,|\,Z{=}f, U{=}u_4)\cdot P(Z{=}f, U{=}u_4)}{P(X{=}0\,|\,Z{=}f)\cdot P(Z{=}f)} \\
&= \frac{2/12 \cdot 1/6}{2/12 \cdot 1/2} = \frac{2}{6}, \\
P(U{=}u_5\,|\,X{=}0, Z{=}f) &= \frac{P(X{=}0\,|\,Z{=}f, U{=}u_5)\cdot P(Z{=}f, U{=}u_5)}{P(X{=}0\,|\,Z{=}f)\cdot P(Z{=}f)} \\
&= \frac{1/12 \cdot 1/6}{2/12 \cdot 1/2} = \frac{1}{6}, \\
P(U{=}u_6\,|\,X{=}0, Z{=}f) &= \frac{P(X{=}0\,|\,Z{=}f, U{=}u_6)\cdot P(Z{=}f, U{=}u_6)}{P(X{=}0\,|\,Z{=}f)\cdot P(Z{=}f)} \\
&= \frac{3/12 \cdot 1/6}{2/12 \cdot 1/2} = \frac{3}{6},
\end{aligned}
$$

and $P(U{=}u_i\,|\,X{=}0, Z{=}f) = 0$ for units u_1, u_2, and u_3.

▷ **Solution 10-7** According to RS-Box 3.2 (ii),

$$E(Y\,|\,X{=}0,Z{=}z) \;=\; \sum_u E(Y\,|\,X{=}0,Z{=}z,U{=}u)\cdot P(U{=}u\,|\,X{=}0,Z{=}z).$$

For $Z{=}m$ (males) this yields

$$\begin{aligned} E(Y\,|\,X{=}0,Z{=}m) \;&=\; \sum_u E(Y\,|\,X{=}0,Z{=}m,U{=}u)\cdot P(U{=}u\,|\,X{=}0,Z{=}m) \\ &=\; 82\cdot\frac{1}{3}+89\cdot\frac{1}{3}+95\cdot\frac{1}{3} \;=\; 88.\overline{6}, \end{aligned}$$

using the results of Exercise 10-6. For $Z{=}f$ (females) we receive

$$\begin{aligned} E(Y\,|\,X{=}0,Z{=}f) \;&=\; \sum_u E(Y\,|\,X{=}0,Z{=}f,U{=}u)\cdot P(U{=}u\,|\,X{=}0,Z{=}f) \\ &=\; 102\cdot\frac{2}{6}+102\cdot\frac{1}{6}+102\cdot\frac{3}{6} \;=\; 102, \end{aligned}$$

again, using the results of Exercise 10-6.

▷ **Solution 10-8** Applying the theorem of total probability to $P(Z{=}z\,|\,X{=}0) = P^{X=0}(Z{=}z)$ yields

$$P(Z{=}z\,|\,X{=}0) \;=\; \sum_u P(Z{=}z,U{=}u\,|\,X{=}0).$$

Hence, for $Z{=}m$ (males) we receive

$$P(Z{=}m\,|\,X{=}0) \;=\; \sum_u P(Z{=}m,U{=}u\,|\,X{=}0) \;=\; \frac{4}{18}+\frac{4}{18}+\frac{4}{18} \;=\; \frac{2}{3},$$

whereas for $Z{=}f$ (females) we receive

$$P(Z{=}f\,|\,X{=}0) \;=\; \sum_u P(Z{=}f,U{=}u\,|\,X{=}0) \;=\; \frac{2}{18}+\frac{1}{18}+\frac{3}{18} \;=\; \frac{1}{3}.$$

▷ **Solution 10-9** In the example of Table 10.4, the random variable $U\colon \Omega \to \Omega_U$, where $\Omega_U :=$ $\{u_1,\dots,u_6\}$, is defined by $U(\omega) = u_i$, where $\omega = (u_i,\omega_X,\omega_Y)$ is an element of the set $\Omega = \Omega_U \times \Omega_X \times \mathbb{R}$ of possible outcomes of the random experiment considered. Such an outcome (u_i,ω_X,ω_Y) indicates that unit u_i is sampled, then assigned (or self-assigned) to treatment condition ω_X, and outcome ω_Y is observed. Therefore, we can now define the map $h\colon \Omega_U \to \{m,f\}$ with

$$h(u_i) = \begin{cases} m, & \text{if } u_i \text{ is male} \\ f, & \text{if } u_i \text{ is female.} \end{cases}$$

The random variable Z = *sex* with values m and f is the composition of this map h with U, that is, $Z = h(U)$. Both, U and Z, are maps on the set Ω of possible outcomes of the single-unit trial, that is, $U:\Omega \to \Omega_U$ and $Z:\Omega \to \{m,f\}$.

▷ **Solution 10-10** The first example is $W_1 := U$, the second is $W_2 := Z :=$ *sex*. A third example can be defined by $W_3(\omega) = 1$ if $\tau_1(\omega) = E^{X=1}(Y\,|\,U)(\omega) = 100$ and $W_3(\omega) = 0$, otherwise. A fourth example is $W_4(\omega) = 1$, if $\tau_0(\omega) < 90$, $W_4(\omega) = 2$, if $90 \le \tau_0(\omega) < 100$, and $W_4(\omega) = 3$, if $\tau_0(\omega) \ge 100$.

The partition of Ω_U associated with U is the following set of subsets of Ω_U: $S_1 = \{\{u_1\},\{u_2\},\dots,\{u_6\}\}$. The partition of Ω_U associated with Z is: $S_2 = \{\{u_1,u_2,u_3\},\,\{u_4,u_5,u_6\}\}$. The partition of Ω_U associated with W_3 is again S_2. Finally, $S_4 = \{\{u_1,u_2\},\{u_3\},\,\{u_4,u_5,u_6\}\}$ is the partition of Ω_U associated with W_4.

▹ **Solution 10-11** There are 122 possible partitions of the total population Ω_U into subsets. It is tedious to obtain this number by writing out all possible partitions. However, it can be obtained by the formula for the *Stirling numbers of the second kind*, which gives the number of possible partitions of a set with n elements into k distinct, non-empty and exhaustive sets. Stirling numbers of the second kind can be calculated by:

$$S(n,k) \; = \; \frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n,$$

where $k! := \; k\cdot(k-1)\cdot\ldots\cdot 1$ and

$$\binom{k}{i} := \frac{k!}{i!\cdot(k-i)!}.$$

The sum $\sum_{j=2}^{6} S(6,j)$ gives the number of different partitions of $\Omega_U := \{u_1, u_2, \ldots, u_6\}$ into two or more subsets, that is, excluding the partition $\{\Omega_U\}$. Hence, in our case this sum is $31+10+65+15+1 = 122$.

▹ **Solution 10-12** Conditional independence of X and U given $Z = sex$ does not hold in this example because in treatment condition 0, the individual treatment probabilities $P(X{=}0 \mid Z{=}f, U{=}u)$ are not the same for each unit u in the female subpopulation, that is, $P(X{=}0 \mid Z{=}f, U{=}u) = P(X{=}0 \mid Z{=}f)$ for all $u \in \{Ann, Sue, Eva\}$ does not hold.

▹ **Solution 10-13** The SR-condition $Y \vDash D_X|(X,Z)$ does not hold in this example, because, for example, in treatment 0, the male units u_1, u_2, and u_3 do not have the same true outcomes $E^{X=0}(Y|U{=}u)$. Instead, these true outcomes are 82, 89, and 95, respectively.

Chapter 11
Identification of Causal Effects Using Propensities

In this chapter, we study how to use conditional treatment probabilities in the identification of causal conditional and average total effects. A first possibility is to use the *true propensities* $\varphi_x = P(X{=}x \mid D_X)$, $x = 1, \ldots, J$, as covariates. We show that conditioning on the multivariate true propensity $\varphi = (\varphi_1, \ldots, \varphi_J)$, the conditional expectations $E^{X=x}(Y \mid \varphi)$, $x \in X(\Omega) = \{0, 1, \ldots, J\}$, are always unbiased, provided that we can assume that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique. In this case we can also identify the causal conditional total effect function $CTE_{\varphi;xx'}$ and the causal average total effect $ATE_{xx'}$.

A second possibility is to choose the Z*-conditional propensities*, that is, the conditional probabilities $\pi_x = P(X{=}x \mid Z)$, $x = 1, \ldots, J$, as new covariates and assume that the original covariate Z is selected such that the multivariate true outcome variable $\tau = (\tau_0, \tau_1, \ldots, \tau_J)$ and X are Z-conditionally independent, that is, assume $\tau \perp\!\!\!\perp X \mid Z$. This conditional independence can be achieved (a) by selecting Z such that X and the global confounder D_X are Z-conditionally independent or (b) by selecting Z such that Y is Z-conditionally mean independent from the global confounder D_X. Defining $\pi = (\pi_1, \ldots, \pi_J)$, we can compute the causal conditional total effect function $CTE_{\pi;xx'}$ and the causal average total effect $ATE_{xx'}$ in both cases, (a) and (b), provided that we assume that the true outcome variables τ_x, $x \in X(\Omega)$, are P-unique.

The last procedure treated in this chapter consists of weighting the outcome variable with weights involving the Z-conditional propensities π_x. This procedure can be used in the true experiment, in which the true propensities are fixed by the experimenter as a function of a covariate Z of X so that $\pi_x = \varphi_x$, $x = 1, \ldots, J$. However, weighting the outcome variable can also be used in a quasi-experiment if we can assume that the covariate Z is selected such that the conditional expectations $E^{X=x}(Y \mid Z)$, $x \in X(\Omega)$, are unbiased.

Notation and Assumptions 11.1

(a) *Let* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ *be a regular probabilistic causality setup and let* D_X *denote a global potential confounder of* X.

(b) *Let* $X(\Omega) = \{0, 1, \ldots, J\}$ *denote the image of* Ω *under* X. *For all* $x \in X(\Omega)$, *define* $\{X{=}x\} := \{\omega \in \Omega \colon X(\omega) = x\}$, *and let* $1_{X=x}$ *denote the indicator of* $\{X{=}x\}$.

(c) *Let* $\varphi_x := P(X{=}x \mid D_X)$, $x \in X(\Omega)$, *denote a (version of the) true propensity of* x.

(d) *Let* $\varphi := (\varphi_1, \ldots, \varphi_J)$ *denote the* J*-variate random variable consisting of the true propensities* $\varphi_1, \ldots, \varphi_J$.

(e) *Let* Y *be real-valued with positive variance.*

(f) *For* $x \in X(\Omega)$, *assume* $0 < P(X{=}x) < 1$ *and define the probability measure* $P^{X=x} \colon \mathscr{A} \to [0, 1]$ *by* $P^{X=x}(A) = P(A \mid X{=}x)$ *for all* $A \in \mathscr{A}$.

(g) *For* $x \in X(\Omega)$*, let* $\tau_x = E^{X=x}(Y|D_X)$ *denote a (version of the) true outcome variable of* Y *given* x *and* $\mathscr{E}^{X=x}(Y|D_X)$ *the set of all such versions.*
(h) *For all* $x \in X(\Omega)$*, assume* $0 < P(X=x) < 1$ *and let* $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$ *denote the* $(J+1)$*-variate random variable consisting of the true outcome variables* $\tau_0, \tau_1, \ldots, \tau_J$ *of* Y*. Furthermore, for* $x, x' \in X(\Omega)$*, let* τ_x *and* $\tau_{x'}$ *denote two such true outcome variables.*
(i) *Let* Z *be a covariate of* X*, that is, let* $\sigma(Z) \subset \sigma(D_X)$*.*
(j) *Let* $z \in Z(\Omega)$ *be a value of* Z*, let* $\{Z=z\} = \{\omega \in \Omega : Z(\omega) = z\}$*, assume that* $0 < P(X=x, Z=z) < 1$*, and define the probability measure* $P^{Z=z}: \mathscr{A} \to [0,1]$ *by* $P^{Z=z}(A) = P(A|Z=z)$*, for all* $A \in \mathscr{A}$*.*
(k) *Assume that* τ_x *is* P*-unique.*
(l) *Assume that all* τ_x*,* $x \in X(\Omega)$*, are* P*-unique.*

11.1 True Propensities

In this section, we show that a conditional probability $\varphi_x := P(X=x|D_X)$, which is also called a *true propensity of* x, may replace the covariate Z of X in the theorems on the identification of causal effects and effect functions (see sections 6.4.1 to 6.4.3). The name 'true propensity of x' is chosen in order to emphasize that it is the conditional probability of treatment x, conditioning *on all* potential confounders of X (see Def. 4.11).

Remark 11.2 [Conditional Randomization] Replacing the (possibly multivariate) covariate Z of X by a true propensity can be useful, and is easily achieved, in the *Z-conditionally randomized experiment*, in which by definition, the true propensity of a value x of X is under the control of the experimenter, who (a) choses a (possibly multivariate) covariate Z of X, (b) fixes the assignment probabilities such that

$$\forall\, x \in X(\Omega):\ P(X=x|D_X) \underset{P}{=} P(X=x|Z), \tag{11.1}$$

and (c) for all $x \in X(\Omega)$, assigns an observational unit obtaining a value z on Z to treatment x with probability $P(X=x|Z=z)$ that is fixed such that $0 < P(X=x|Z=z) < 1$. ◁

Remark 11.3 [Treatment-Specific Covariates] Fixing the conditional assignment probabilities for treatment x, the experimenter can also use a covariate Z_x that is specific for treatment x. If X has more than two values $x_0, x_1, \ldots, x_J$ that all have a positive probability, then the covariate Z_x may differ from the covariate $Z_{x'}$ used for the assignment to two different treatments x and x'. If this is the case, then we still can always consider the multivariate covariate $Z = (Z_1, \ldots, Z_J)$. For this covariate the following theorem holds. ◁

Theorem 11.4 [Z-Conditional Independence]
Let the Assumptions 11.1 (a) and (b) hold and assume that Z_x*,* $x \in \{1, \ldots, J\}$*, are covariates of* X*. Then* $Z = (Z_1, \ldots, Z_J)$ *is a covariate of* X *and*

$$\forall\, x \in X(\Omega):\ P(X=x|D_X) \underset{P}{=} P(X=x|Z_x) \tag{11.2}$$

implies Equation (11.1).

Proof:

The fact that Z is a covariate of X follows from Definition 4.11 (iv), Remark 4.16, and RS-Theorem 2.16. Furthermore, for all $x \in X(\Omega)$,

$$\begin{aligned}
P(X{=}x \mid Z) &\underset{P}{=} E(1_{X=x} \mid Z) && [RS\text{-}(4.10)]\\
&\underset{P}{=} E\big(E(1_{X=x} \mid D_X) \,\big|\, Z\big) && [RS\text{-}Box\ 4.1\ (xiii)]\\
&\underset{P}{=} E\big(P(X{=}x \mid D_X) \,\big|\, Z\big) && [RS\text{-}(4.10)]\\
&\underset{P}{=} E\big(P(X{=}x \mid Z_x) \,\big|\, Z\big) && [(11.2)]\\
&\underset{P}{=} P(X{=}x \mid Z_x) && \big[\sigma\big(P(X{=}x \mid Z_x)\big) \subset \sigma(Z), RS\text{-}Box\ 4.1\ (xi)\big]\\
&\underset{P}{=} P(X{=}x \mid D_X). && [(11.2)]
\end{aligned}$$

Remark 11.5 [A Methodological Implication] Practically speaking, Theorem 11.4 means that the experimenter has the choice to use either treatment-specific covariates Z_x or the same (uni- or multivariate) covariate Z of X (e. g., psychological stress) in order to fix the conditional treatment probabilities $P(X{=}x \mid D_X)$, that is, to fix the true propensities. In both cases, there is a covariate Z of X for which Equation (11.1) holds. ◁

Remark 11.6 [Z-Conditional Independence of D_X and X] If the Assumptions 11.1 (a) and (b) hold, then, according to Proposition (8.14),

$$\forall\, x \in X(\Omega):\ P(X{=}x \mid D_X) \underset{P}{=} P(X{=}x \mid Z) \quad \Leftrightarrow \quad D_X \perp\!\!\!\perp X \mid Z \tag{11.3}$$

and, according to Proposition (8.15),

$$P(X{=}x \mid D_X) \underset{P}{=} P(X{=}x \mid Z) \quad \Leftrightarrow \quad D_X \perp\!\!\!\perp 1_{X=x} \mid Z. \tag{11.4}$$

◁

Remark 11.7 [Conditionally Randomized Experiment vs. Quasi-Experiment] Defining

$$\varphi_x := P(X{=}x \mid D_X), \quad x \in X(\Omega), \tag{11.5}$$

we can summarize that the true treatment probabilities φ_x, $x \in \{1, \ldots, J\}$, and with it $\varphi = (\varphi_1, \ldots, \varphi_J)$ are known if they are chosen by the experimenter such that Equation (11.1) holds. The validity of this equation distinguishes a *Z-conditionally randomized experiment* from a *quasi-experiment* with $J+1$ treatment conditions. For details of fixing treatment probabilities see section 11.2.1. ◁

Remark 11.8 [Covariate Selection] Beyond the conditionally randomized experiment, the multivariate true propensity $\varphi = (\varphi_1, \ldots, \varphi_J)$ can be estimated if a (possibly multivariate) covariate Z *is selected* such that Equation (11.1) holds. This strategy is used in various techniques in the analysis of observational studies. Details on covariate selection are treated in section 11.2.2. ◁

Remark 11.9 [True Propensities are Covariates of X] The propensities φ_x, $x \in \{1,\dots,J\}$, and the J-variate propensity $\varphi = (\varphi_1,\dots,\varphi_J)$ are covariates of X because $\sigma(\varphi_x) \subset \sigma(\varphi) \subset \sigma(D_X)$ [see Def. 4.11 (iv), Rem. 4.16, RS-Def. 4.4, and RS-Rem. 4.12]. $\lhd$

Now we show that, under the Assumptions 11.1 (a) to (c), φ_x-conditional independence of D_X and the indicator $1_{X=x}$, denoted $D_X \perp\!\!\!\perp 1_{X=x} \,|\, \varphi_x$, always holds. Note that

$$D_X \perp\!\!\!\perp 1_{X=x} \,|\, \varphi_x \quad \Leftrightarrow \quad P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x \,|\, \varphi_x) \tag{11.6}$$

[see again Prop. (8.15)]. Furthermore, under the Assumptions 11.1 (a) to (d), we show that $D_X \perp\!\!\!\perp X \,|\, \varphi$ always holds for $\varphi := (\varphi_1,\dots,\varphi_J)$. Note that, if X is discrete, then

$$D_X \perp\!\!\!\perp X \,|\, \varphi \quad \Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x \,|\, \varphi) \tag{11.7}$$

$$\Leftrightarrow \quad \forall\, x \in X(\Omega)\colon\; D_X \perp\!\!\!\perp 1_{X=x} \,|\, \varphi \tag{11.8}$$

[see again Props. (8.14) and (8.15)].

Theorem 11.10 [φ-Conditional Independence of X and D_X]
Let the Assumptions 11.1 (a) to (c) hold. Then

(i) $\forall\, x \in X(\Omega)\colon D_X \perp\!\!\!\perp 1_{X=x} \,|\, \varphi_x$.

If we additionally assume 11.1 (d), then

(ii) $D_X \perp\!\!\!\perp X \,|\, \varphi$
(iii) $\forall\, x \in X(\Omega)\colon D_X \perp\!\!\!\perp 1_{X=x} \,|\, \varphi$.

Proof

Proposition (i). Note that φ_x is D_X-measurable and $P(X{=}x|\varphi_x)$ is a notation for the conditional expectation $E(1_{X=x}|\varphi_x)$ [see RS-(4.10)]. Hence,

$$\forall\, x \in X(\Omega)\colon P(X{=}x|\varphi_x) \underset{P}{=} E\big(P(X{=}x \,|\, D_X) \,\big|\, \varphi_x\big) \qquad [\sigma(\varphi_x) \subset \sigma(D_X),\ \textit{RS-Box 4.1 (xiii)}]$$
$$\underset{P}{=} P(X{=}x \,|\, D_X)\,. \qquad [(11.5),\ \textit{RS-Box 4.1 (xi)}]$$

According to Proposition (11.6), this proves Proposition (i).
Proposition (ii). Similarly,

$$\forall\, x \in X(\Omega)\colon P(X{=}x|\varphi) \underset{P}{=} E\big(P(X{=}x \,|\, D_X) \,\big|\, \varphi\big) \qquad [\sigma(\varphi) \subset \sigma(D_X),\ \textit{RS-Box 4.1 (xiii)}]$$
$$\underset{P}{=} P(X{=}x \,|\, D_X)\,. \qquad [(11.5),\ \textit{RS-Box 4.1 (xi)}]$$

According to Proposition (11.7), this proves Proposition (ii).
Proposition (iii). This follows from propositions (ii), (11.7), and (11.8).

According to Proposition (i) of Theorem 11.10, φ_x-conditional independence of D_X and $1_{X=x}$, $x \in X(\Omega)$, does not rest on any restrictive assumptions. We only assume that

$$\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t\in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$$

is a regular probabilistic causality setup, D_X a global potential confounder of X, and X is discrete with a finite number of values $x = 0, 1, \dots, J$. Under these assumptions, φ_x-conditional independence of D_X and $1_{X=x}$ always holds. Defining $\varphi := (\varphi_1, \dots, \varphi_J)$, these assumptions also imply φ-conditional independence of D_X and X as well as $D_X \perp\!\!\!\perp 1_{X=x} \,|\, \varphi$, $x \in X(\Omega)$, because $\sigma(1_{X=x}) \subset \sigma(X)$ [see Props. (ii) and (iii) of Th. 11.10].

Remark 11.11 [Methodological Consequences] To emphasize, if we consider the specific covariates φ_x and φ, then conditional independence of D_X and X given φ_x or φ are not assumptions. Instead these independencies necessarily hold. Therefore, we can apply all theorems and corollaries that we already studied in section 8.3.2 dealing with Z-conditional independence of D_X and $1_{X=x}$ or X, simply replacing Z by φ_x or φ. This will be detailed in the following subsections. $\lhd$

11.1.1 True Propensities and True Outcome Variables

According to the following corollary, the putative cause variable X and the multivariate true outcome variable $\tau = (\tau_0, \tau_1, \dots, \tau_J)$ are conditionally independent given φ. This implies that X and all τ_x, $x \in X(\Omega)$, are *φ-conditionally independent* which in turn implies that all τ_x, $x \in X(\Omega)$, are *φ-conditionally mean-independent* from X. Remember, *φ-conditional mean-independence* of all true outcome variables τ_x, $x \in X(\Omega)$, is defined by

$$\forall\, x \in X(\Omega)\colon\; E(\tau_x | X, \varphi) \underset{P}{=} E(\tau_x | \varphi) \tag{11.9}$$

and abbreviated $\forall\, x \in X(\Omega)\colon \tau_x \vDash X \,|\, \varphi$.

Corollary 11.12 [Conditional Independence of τ and X Given φ]
Let the Assumptions 11.1 (a) to (h) hold. Then

(i) $\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp X \,|\, \varphi$
(ii) $\forall\, x \in X(\Omega)\colon \tau_x \vDash X \,|\, \varphi$
(iii) $\tau \perp\!\!\!\perp X \,|\, \varphi$.

Proof:
Proposition (i). This follows from Theorem 11.10 (iii), $\sigma(\tau_x) \subset \sigma(D_X)$ *for all* $x \in X(\Omega)$*, and RS-Box 6.1* (vi).
Proposition (ii). This follows from (i) and RS-Theorem 6.8.
Proposition (iii). This follows from Theorem 11.10 (iii), $\sigma(\tau) \subset \sigma(D_X)$*, and RS-Box 6.1* (vi).

Note that Propositions (i) and (iii) of Corollary 11.12 are not equivalent to each other because $\tau \perp\!\!\!\perp X \,|\, \varphi$ implies $\tau_x \perp\!\!\!\perp X \,|\, \varphi$ for all $x \in X(\Omega)$, but not vice versa (see Exercise 11-2).

In the following corollary we turn to some conditional independencies of the true outcome variables and indicators of an event $\{X=x\}$.

Corollary 11.13 [Conditional Independence of τ_x and $1_{X=x}$ Given φ_x or φ]
Let the Assumptions 11.1 (a) to (g) hold and let $\tau_{x'}$, $x' \in X(\Omega)$*, denote a true outcome variable of* Y *given* x'*. Then*

(i) $\forall\, x \in X(\Omega)\colon \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, \varphi_x$
(ii) $\forall\, x \in X(\Omega)\colon \tau_x \vDash 1_{X=x} \,|\, \varphi_x$

(iii) $\forall\, x, x' \in X(\Omega)\colon \tau_{x'} \perp\!\!\!\perp 1_{X=x} \,|\, \varphi$
(iv) $\forall\, x, x' \in X(\Omega)\colon \tau_{x'} \vDash 1_{X=x} \,|\, \varphi$.

Proof
Proposition (i). This follows from Theorem 11.10 (i) and $\sigma(\tau_x) \subset \sigma(D_X)$.
Proposition (ii). This follows from (i) and RS-Theorem 6.8.
Proposition (iii). This follows from Theorem 11.10 (iii) and $\sigma(\tau_{x'}) \subset \sigma(D_X)$.
Proposition (iv). This follows from (iii) and RS-Theorem 6.8.

Hence, if the Assumptions 11.1 (a) to (f) hold, then, according to the first proposition of Corollary 11.13, the true outcome variable τ_x and the indicator $1_{X=x}$ are φ_x-conditionally independent and this holds for all $x \in X(\Omega)$. According to the second proposition of this corollary, for all $x \in X(\Omega)$, τ_x is φ_x-conditionally mean-independent from $1_{X=x}$, that is,

$$\forall\, x \in X(\Omega)\colon\; E(\tau_x \,|\, 1_{X=x}, \varphi_x) \underset{P}{=} E(\tau_x \,|\, \varphi_x). \tag{11.10}$$

Furthermore, according to the third proposition of Corollary 11.13, all true outcome variables $\tau_{x'}$, $x' \in X(\Omega)$, and all indicators $1_{X=x}$, $x \in X(\Omega)$, are φ-conditionally independent. Finally, according to the fourth proposition of this corollary, all $\tau_{x'}$, $x' \in X(\Omega)$, are φ-conditionally mean-independent from all indicators $1_{X=x}$, $x \in X(\Omega)$, and this implies

$$\forall\, x \in X(\Omega)\colon\; E(\tau_x \,|\, 1_{X=x}, \varphi) \underset{P}{=} E(\tau_x \,|\, \varphi). \tag{11.11}$$

Remark 11.14 [Consequences for Unbiasedness] If, additional to the Assumptions 11.1 (a) to (g), we assume that all τ_x, $x \in X(\Omega)$, are P-unique, then, according to Equations (11.10), (11.11), and Theorem 6.20 (i), we can conclude that all conditional expectations $E^{X=x}(Y \,|\, \varphi_x)$ and $E^{X=x}(Y \,|\, \varphi)$, $x \in X(\Omega)$, are unbiased, that is,

$$\forall\, x \in X(\Omega)\colon\; E^{X=x}(Y \,|\, \varphi_x) \underset{P}{=} E(\tau_x \,|\, \varphi_x) \tag{11.12}$$

and

$$\forall\, x \in X(\Omega)\colon\; E^{X=x}(Y \,|\, \varphi) \underset{P}{=} E(\tau_x \,|\, \varphi) \tag{11.13}$$

[see Def. 6.13 (i)]. Remember, according to Definition 6.18 (i), the conditional expectation $E(Y \,|\, X, \varphi)$ is called *unbiased*, denoted $E(Y \,|\, X, \varphi) \vdash D_X$, if all $E^{X=x}(Y \,|\, \varphi)$, $x \in X(\Omega)$, are unbiased, that is, if Equation (11.13) holds. Hence, Theorem 6.20 (i), Corollary 11.13, and Definition 6.18 (i) imply the following corollary. ◁

Corollary 11.15 [Unbiasedness of $E^{X=x}(Y \,|\, \varphi_x)$ and $E^{X=x}(Y \,|\, \varphi)$]
Let the Assumptions 11.1 (a) to (g) hold and assume that τ_x, $x \in X(\Omega)$, is P-unique. Then $E^{X=x}(Y \,|\, \varphi_x)$ and $E^{X=x}(Y \,|\, \varphi)$, $x \in X(\Omega)$, are unbiased. Furthermore, if all τ_x, $x \in X(\Omega)$, are P-unique, then $E(Y \,|\, X, \varphi)$ is unbiased as well.

Remark 11.16 [Unbiasedness of $E^{X=x}(Y \,|\, \varphi_x = p_x)$ and $E^{X=x}(Y \,|\, \varphi = p)$] Let $p_x \in [0,1]$ be a value of the true propensity φ_x and $p \in [0,1]^J$ be a value of the J-variate propensity $\varphi = (\varphi_1, \ldots, \varphi_J)$. If $P^{X=x}(\varphi_x = p_x) > 0$, then, according to Definition 6.13 (ii), the conditional expectation value $E^{X=x}(Y \,|\, \varphi_x = p_x)$ is called *unbiased*, denoted $E^{X=x}(Y \,|\, \varphi_x = p_x) \vdash D_X$, if τ_x is $P^{\varphi_x = p_x}$-unique, and

$$E^{X=x}(Y\,|\,\varphi_x = p_x) \;=\; E(\tau_x|\,\varphi_x = p_x). \tag{11.14}$$

Correspondingly, if $P^{X=x}(\varphi = p) > 0$, then, according to Definition 6.13 (ii), the conditional expectation value $E^{X=x}(Y\,|\,\varphi = p)$ is called *unbiased*, denoted $E^{X=x}(Y\,|\,\varphi = p) \vdash D_X$, if τ_x is $P^{\varphi=p}$-unique, and

$$E^{X=x}(Y\,|\,\varphi = p) \;=\; E(\tau_x|\,\varphi = p). \tag{11.15}$$

Furthermore, remember that $P^{\varphi_x=p_x}$-uniqueness and $P^{\varphi=p}$-uniqueness of τ_x follows from τ_x being P-unique, but not vice versa [see RS-Box 5.1 (v)]. $\triangleleft$

Corollary 11.17 [Unbiasedness of $E^{X=x}(Y\,|\,\varphi_x = p_x)$ and $E^{X=x}(Y\,|\,\varphi = p)$]
Let the Assumptions 11.1 (a) to (g) hold. Let $p_x \in [0,1]$ be a value of φ_x and $p \in [0,1]^J$ a value of $\varphi = (\varphi_1,\ldots,\varphi_J)$.

(i) If $P^{X=x}(\varphi_x{=}p_x) > 0$ and τ_x is $P^{\varphi_x=p_x}$-unique, then $E^{X=x}(Y\,|\,\varphi_x = p_x)$ is unbiased.
(ii) If $P^{X=x}(\varphi = p) > 0$ and τ_x is $P^{\varphi=p}$-unique, then $E^{X=x}(Y\,|\,\varphi = p)$ is unbiased.

Proof:
These two propositions follow from Corollary 11.15, Equations (11.12), (11.13)*, RS-Remark 5.31, and RS-Remark 2.55.*

11.1.2 Identification of Total Effect Functions

Corollary 11.15 has immediate implications on unbiasedness of the φ-conditional total effect functions, provided that we additionally assume P-uniqueness of τ_x and $\tau_{x'}$ in order to make sure that unbiasedness of $E^{X=x}(Y\,|\,\varphi)$ and $E^{X=x'}(Y\,|\,\varphi)$ is defined.

Corollary 11.18 [Identification of the φ-Conditional Total Effect Function]
Let the Assumptions 11.1 (a) to (g) hold, let $\tau_{x'}$, $x' \in X(\Omega)$, denote a true outcome variable of Y given x', assume that τ_x and $\tau_{x'}$ are P-unique, and let $CTE_{\varphi;xx'}$ denote the causal φ-conditional total effect function comparing x to x' with respect to Y.

(i) Then

$$CTE_{\varphi;xx'}(\varphi) \underset{P}{=} E^{X=x}(Y\,|\,\varphi) \;-\; E^{X=x'}(Y\,|\,\varphi) \tag{11.16}$$

holds for the composition of φ and $CTE_{\varphi;xx'}$.
(ii) If p is a value of φ such that $P(X{=}x, \varphi{=}p), P(X{=}x', \varphi{=}p) > 0$, then

$$CTE_{\varphi;xx'}(p) \;=\; E^{X=x}(Y\,|\,\varphi{=}p) \;-\; E^{X=x'}(Y\,|\,\varphi{=}p)\,. \tag{11.17}$$

Proof
Proposition (i). According to Corollary 11.15, $E^{X=x}(Y\,|\,\varphi)$ and $E^{X=x'}(Y\,|\,\varphi)$ are unbiased. Hence,

$$\begin{aligned} E^{X=x}(Y|\varphi) - E^{X=x'}(Y|\varphi) &\underset{P}{=} E(\tau_x|\varphi) - E(\tau_{x'}|\varphi) && [\textit{Def. } 6.13\ (i)]\\ &\underset{P}{=} E(\tau_x - \tau_{x'}|\varphi) && [\textit{RS-Box } 4.1\ (\text{xviii})] \end{aligned}$$

$$\underset{P}{=} \; CTE_{\varphi;\,xx'}(\varphi). \qquad \text{[Def. 5.38 (i)]}$$

Proposition (ii). *Note that* $P(X{=}x, \varphi{=}p) > 0$ *implies* $P(\varphi{=}p) > 0$ *and* P*-uniqueness of* τ_x *and* $\tau_{x'}$ *implies that* τ_x *and* $\tau_{x'}$ *are also* $P^{\varphi=p}$*-unique [see RS-Box 5.1 (v)]. Finally, according to Corollary 11.17, the conditional expectation values* $E^{X=x}(Y\,|\,\varphi{=}p)$ *and* $E^{X=x'}(Y\,|\,\varphi{=}p)$ *are unbiased. Hence,*

$$\begin{aligned} E^{X=x}(Y\,|\,\varphi{=}p) - E^{X=x'}(Y\,|\,\varphi{=}p) &= E(\tau_x\,|\,\varphi{=}p) - E(\tau_{x'}\,|\,\varphi{=}p) && \text{[Def. 6.13 (ii)]} \\ &= E(\tau_x - \tau_{x'}\,|\,\varphi{=}p) && \text{[RS-(3.35)]} \\ &= CTE_{\varphi;\,xx'}(p). && \text{[Def. 5.38 (ii)]} \end{aligned}$$

11.1.3 Identifying the Expectation of a True Outcome Variable

According to the following theorem, we can use the conditional expectations $E^{X=x}(Y\,|\,\varphi_x)$ and $E^{X=x}(Y\,|\,\varphi)$ for identifying the expectation of a true outcome variable τ_x.

Corollary 11.19 [Identifying the Expectation of a True Outcome Variable]
Let the Assumptions 11.1 (a) to (g) hold and assume that τ_x *is* P*-unique. Then*

$$E(\tau_x) \;=\; E\big(E^{X=x}(Y\,|\,\varphi)\big) \;=\; E\big(E^{X=x}(Y\,|\,\varphi_x)\big). \tag{11.18}$$

Proof:
According to Corollary 11.15, $E^{X=x}(Y|\varphi_x)$ *and* $E^{X=x}(Y|\varphi)$ *are unbiased. Assuming that* τ_x *is* P*-unique implies that the expectation* $E(\tau_x)$ *is uniquely defined [see RS-Box 5.1 (vii) for* $C = \Omega$*] and that the conditional expectation* $E(\tau_x|\,\varphi_x)$ *is* P*-unique as well [see RS-Box 4.1 (xiv) for* $Y = \tau_x$ *and* $Z = \tau_x'$*, where* $\tau_x, \tau_x' \in \mathscr{E}^{X=x}(Y|D_X)$ *and RS-Rem. 4.10]. Hence,*

$$\begin{aligned} E\big(E^{X=x}(Y|\varphi_x)\big) &= E\big(E(\tau_x|\,\varphi_x)\big) && \text{[Cor. 11.15, (11.12)]} \\ &= E(\tau_x). && \text{[RS-Box 4.1 (iv)]} \end{aligned}$$

The corresponding arguments apply if we replace φ_x *by* φ.

Now we consider the implications of Corollary 11.19 on the identification of a causal average total effect. According to Corollary 11.20, the causal average total effect $ATE_{xx'}$ can be computed from the expectations of the conditional expectations $E^{X=x}(Y|\varphi_x)$ and $E^{X=x'}(Y|\varphi_{x'})$ and from the expectations of the conditional expectations $E^{X=x}(Y|\varphi)$ and $E^{X=x'}(Y|\varphi)$, provided that we can presume that τ_x and $\tau_{x'}$ are P-unique.

Corollary 11.20 [Identification of the Causal Average Total Effect]
Let the Assumptions 11.1 (a) to (h) hold and assume that τ_x *and* $\tau_{x'}$ *are* P*-unique. Then*

$$ATE_{xx'} \;=\; E\big(E^{X=x}(Y|\varphi_x)\big) \;-\; E\big(E^{X=x'}(Y|\varphi_{x'})\big) \tag{11.19}$$

$$= E\big(E^{X=x}(Y|\varphi)\big) - E\big(E^{X=x'}(Y|\varphi)\big). \tag{11.20}$$

Proof:

$$\begin{aligned} ATE_{xx'} &= E(\tau_x) - E(\tau_{x'}) && [\text{Def. 5.38 (i)}] \\ &= E\big(E^{X=x}(Y|\varphi_x)\big) - E\big(E^{X=x'}(Y|\varphi_{x'})\big). && [\text{Cor. 11.19}] \end{aligned}$$

The corresponding argument applies if we replace φ_x and $\varphi_{x'}$ by φ.

11.1.4 Identifying a Conditional Expectation of a True Outcome Variable

Equation (11.18) shows how to identify the expectation $E(\tau_x)$ of a true outcome variable τ_x via the expectation of the φ-conditional expectation $E^{X=x}(Y|\varphi)$ and via the expectation of the φ_x-conditional expectation $E^{X=x}(Y|\varphi_x)$. Of course, $E(\tau_x)$ is only a very coarse information as compared to the true outcome variable τ_x itself.

The next theorem shows how to identify the conditional expectation $E(\tau_x|V_x)$, where V_x is any random variable that is measurable with respect to $(1_{X=x}, \varphi_x)$. The random variable V_x can be chosen such that the degree of dissolution in the reaggregation is as fine-grained as desired. Of course, V_x cannot contain more information (in terms of the σ-algebra it generates) than the random variable $(1_{X=x}, \varphi_x)$.

Corollary 11.21 [Identifying a V_x-Conditional Expectation of τ_x]
Let the Assumptions 11.1 (a) to (g) hold and assume that τ_x is P-unique. Furthermore, let V_x be a random variable on $(\Omega, \mathscr{A}, P)$ such that $\sigma(V_x) \subset \sigma(1_{X=x}, \varphi_x)$.

(i) Then

$$E(\tau_x|V_x) \underset{P}{=} E\big(E^{X=x}(Y|\varphi_x)\,\big|\,V_x\big). \tag{11.21}$$

(ii) If v_x is a value of V_x such that $P(X=x, V_x=v_x) > 0$, then

$$E(\tau_x|V_x=v_x) = E\big(E^{X=x}(Y|\varphi_x)\,\big|\,V_x=v_x\big). \tag{11.22}$$

Proof:
Proposition (i).

$$\begin{aligned} E\big(E^{X=x}(Y|\varphi_x)\,\big|\,V_x\big) &\underset{P}{=} E\big(E(\tau_x|\varphi_x)\,\big|\,V_x\big) && [\text{Cor. 11.15, (11.12)}] \\ &\underset{P}{=} E\big(E(\tau_x|1_{X=x}, \varphi_x)\,\big|\,V_x\big) && [\text{Cor. 11.13 (ii), Def. 4.41}] \\ &\underset{P}{=} E(\tau_x|V_x). && [\text{Box RS-4.1 (xiii)}] \end{aligned}$$

Proposition (ii). This equation immediately follows from (i) and RS-Theorem 4.27 because $P(X=x, V_x=v_x) > 0$ implies $P(V_x=v_x) > 0$.

The following theorem shows how to use the conditional expectation $E^{X=x}(Y|\varphi)$ for identifying the conditional expectation $E(\tau_x|V)$ of a true outcome variable τ_x given a random variable V that is (X, φ)-measurable.

Corollary 11.22 [Identifying a V-Conditional Expectation of τ_x]
Let the Assumptions 11.1 (a) to (h) hold and assume that all τ_x, $x \in X(\Omega)$, are P-unique. Furthermore, let V be a random variable on $(\Omega, \mathscr{A}, P)$ such that $\sigma(V) \subset \sigma(X, \varphi)$.

(i) Then

$$\forall\, x \in X(\Omega)\colon\; E\big(E^{X=x}(Y \,|\, \varphi) \,\big|\, V\big) \underset{P}{=} E(\tau_x | V). \tag{11.23}$$

(ii) If v is a value of V such that $P(X{=}x, V{=}v) > 0$, then

$$E\big(E^{X=x}(Y \,|\, \varphi) \,\big|\, V{=}v\big) \;=\; E(\tau_x | V{=}v). \tag{11.24}$$

Proof:
Proposition (i). For all $x \in X(\Omega)$,

$$\begin{aligned}
E\big(E^{X=x}(Y \,|\, \varphi) \,\big|\, V\big) &\underset{P}{=} E\big(E(\tau_x | \varphi) \,\big|\, V\big) && [\textit{Cor. } 11.15, (11.13)]\\
&\underset{P}{=} E\big(E(\tau_x | X, \varphi) \,\big|\, V\big) && [\textit{Cor. } 11.12\ (ii), \textit{Def. } 4.41]\\
&\underset{P}{=} E(\tau_x | V). && [\textit{Box RS-}4.1\ (\text{xiii})]
\end{aligned}$$

Proposition (ii). This equation immediately follows from (i) and RS-Theorem 4.27 because $P(X{=}x, V{=}v) > 0$ implies $P(V{=}v) > 0$.

Remark 11.23 [A Caveat] Note that Corollary 11.21 (i) is restricted to a $(1_{X=x}, \varphi_x)$-measurable random variable V_x. Correspondingly, Corollary 11.22 (i) is restricted to a (X, φ)-measurable random variable V. Hence, these corollaries do not allow to identify $E(\tau_x | Z)$ via the conditional expectations $E^{X=x}(Y | \varphi_x)$ or $E^{X=x}(Y | \varphi)$ for *every* covariate Z of X. $\triangleleft$

11.1.5 Identifying a Causal Conditional Total Effect Function

Corollary 11.20 shows how to identify the causal average total effect. In contrast, Corollary 11.24 shows how to identify a causal conditional total effect function that is less fine-grained than $CTE_{\varphi;xx'}(\varphi)$ but still more informative than the causal average total effect. This can be useful, for example, in statistical analysis if sample size is too small for a reliable estimation of $CTE_{\varphi;xx'}(\varphi)$.

Corollary 11.24 [Identifying a Causal Conditional Total Effect Function]
Let the Assumptions 11.1 (a) to (h) hold and assume that τ_x and $\tau_{x'}$ are P-unique. Furthermore, let V be a random variable on $(\Omega, \mathscr{A}, P)$ such that $\sigma(V) \subset \sigma(X, \varphi)$.

(i) Then

$$CTE_{V;xx'}(V) \underset{P}{=} E\big(E^{X=x}(Y \,|\, \varphi) \,\big|\, V\big) - E\big(E^{X=x'}(Y \,|\, \varphi) \,\big|\, V\big). \tag{11.25}$$

(ii) If $v \in \Omega'_V$ is a value of V such that $P(X{=}x, V{=}v)$, $P(X{=}x', V{=}v) > 0$, then

$$CTE_{V;xx'}(v) \;=\; E\big(E^{X=x}(Y \,|\, \varphi) \,\big|\, V{=}v\big) - E\big(E^{X=x'}(Y \,|\, \varphi) \,\big|\, V{=}v\big). \tag{11.26}$$

Proof:

Proposition (i).

$$
\begin{aligned}
CTE_{V;xx'}(V) &\underset{P}{=} E(\tau_x - \tau_{x'} | V) && [\textit{Def. } 5.38\ (i)] \\
&\underset{P}{=} E(\tau_x | V) \;-\; E(\tau_{x'} | V) && [\textit{RS-Box } 4.1\ (\text{xviii})] \\
&\underset{P}{=} E\big(E^{X=x}(Y\,|\,\varphi)\,\big|\,V\big) \;-\; E\big(E^{X=x'}(Y\,|\,\varphi)\,\big|\,V\big). && [\textit{Cor. } 11.22\ (i)]
\end{aligned}
$$

Proposition (ii). This equation immediately follows from (i) and RS-Theorem 4.27 because $P(X{=}x, V{=}v) > 0$ *implies* $P(V{=}v) > 0$.

Note that $CTE_{V;xx'}(V)$ is still a *causal* conditional total effect variable. In general, $CTE_{V;xx'}(V)$ is not identical to the prima facie total effect variable

$$
PFE_{V;xx'}(V) \underset{P}{=} E^{X=x}(Y|V) \;-\; E^{X=x'}(Y|V).
$$

In contrast to $CTE_{V;xx'}(V)$, the random variable $PFE_{V;xx'}(V)$ has no causal meaning unless it happens to be identical to $CTE_{V;xx'}(V)$.

11.1.6 Numerical Example

Example 11.25 [Z-Conditional Independence of X and τ — continued] We use the example presented in Table 7.5 that also displays the values of the true propensity

$$
\varphi \;=\; P(X{=}1\,|\,D_X) \;=\; P(X{=}1\,|\,U). \tag{11.27}
$$

In this specific example, there are three different true propensity scores for treatment 1: The first unit has probability $P(X{=}1\,|\,U{=}u_1) = 7/8$ for treatment 1, the next three units have probability $5/8$, and the last two units have probability $2/8$ for treatment 1. These conditional probabilities are the values of the true propensity $\varphi = \varphi_1$ of treatment 1 because, in this example, U is a global potential confounder of X. Because X is dichotomous, we only need this single propensity φ_1. The other one can be computed via $P(X{=}0\,|\,U) = \varphi_0 = 1 - \varphi_1$.

First of all, we check if the true outcome variables $\tau_0 = E^{X=0}(Y|D_X) = E^{X=0}(Y|U)$ and $\tau_1 = E^{X=1}(Y|D_X) = E^{X=1}(Y|U)$ are P-unique. According to Theorem 5.27 (ii), this is the case, because $P(X{=}0\,|\,U), P(X{=}1\,|\,U) > 0$. According to Corollary 11.15, this implies that the conditional expectations $E^{X=x}(Y|\varphi)$ of the outcome variable Y given treatment x and the true propensity φ are unbiased for both values, 0 and 1, of the treatment variable X.

Table 11.1 displays the conditional expectation values $E^{X=x}(Y\,|\,\varphi{=}p)$, which can be computed from the parameters in Table 7.5 using

$$
\begin{aligned}
E^{X=x}(Y|\varphi{=}p) \;&=\; \sum_u E^{X=x}(Y|\varphi{=}p, U{=}u)\cdot P^{X=x}(U{=}u\,|\,\varphi{=}p) \\
&=\; \sum_u E^{X=x}(Y|U{=}u)\cdot P^{X=x}(U{=}u\,|\,\varphi{=}p)
\end{aligned} \tag{11.28}
$$

(see Exercise 11-4).

Because φ has only three different values, the conditional expectations $E^{X=x}(Y|\varphi)$ can be parameterized as

Table 11.1. Conditional expectation values $E^{X=x}(Y \mid \varphi = p)$ of the outcome variable Y given treatment x and true propensity p in the example of Table 7.5

	$\varphi = P(X=1 \mid U)$		
Treatment	7/8	5/8	2/8
$x=0$	72	89	111
$x=1$	83	96	122

$$E^{X=x}(Y \mid \varphi) = \Upsilon_0^{(x)} + \Upsilon_1^{(x)} \cdot 1_{\varphi=5/8} + \Upsilon_2^{(x)} \cdot 1_{\varphi=2/8}, \tag{11.29}$$

where $1_{\varphi=5/8}$ and $1_{\varphi=2/8}$ denote the indicator variables for the events $\{\varphi = 5/8\}$ and $\{\varphi = 2/8\}$, respectively. For treatment $x = 0$, this conditional expectation is

$$\begin{aligned} E^{X=0}(Y \mid \varphi) &= \Upsilon_0^{(0)} + \Upsilon_1^{(0)} \cdot 1_{\varphi=5/8} + \Upsilon_2^{(0)} \cdot 1_{\varphi=2/8} \\ &= 72 + 17 \cdot 1_{\varphi=5/8} + 39 \cdot 1_{\varphi=2/8}, \end{aligned} \tag{11.30}$$

and for treatment $x = 1$ it is

$$\begin{aligned} E^{X=1}(Y \mid \varphi) &= \Upsilon_0^{(1)} + \Upsilon_1^{(1)} \cdot 1_{\varphi=5/8} + \Upsilon_2^{(1)} \cdot 1_{\varphi=2/8} \\ &= 83 + 13 \cdot 1_{\varphi=5/8} + 39 \cdot 1_{\varphi=2/8}. \end{aligned} \tag{11.31}$$

(see Exercise 11-5).

Because the conditional expectations $E^{X=x}(Y \mid \varphi)$ are unbiased, the prima facie effect variable

$$\begin{aligned} PFE_{\varphi;10}(\varphi) &= E^{X=1}(Y \mid \varphi) - E^{X=0}(Y \mid \varphi) \\ &= 83 + 13 \cdot 1_{\varphi=5/8} + 39 \cdot 1_{\varphi=2/8} - (72 + 17 \cdot 1_{\varphi=5/8} + 39 \cdot 1_{\varphi=2/8}) \\ &= 11 - 4 \cdot 1_{\varphi=5/8} \end{aligned} \tag{11.32}$$

is also the *causal φ-conditional total effect variable* $CTE_{\varphi;10}(\varphi)$. That is, $PFE_{\varphi;10}(\varphi) = CTE_{\varphi;10}(\varphi)$. Hence, the causal conditional effects $CTE_{\varphi;10}(p)$ of treatment 1 compared to control given $\varphi = p$ are

$$CTE_{\varphi;10}(7/8) = PFE_{\varphi;10}(7/8) = 11 - 4 \cdot 0 = 11,$$

$$CTE_{\varphi;10}(5/8) = PFE_{\varphi;10}(5/8) = 11 - 4 \cdot 1 = 7,$$

and

$$CTE_{\varphi;10}(2/8) = PFE_{\varphi;10}(2/8) = 11 - 4 \cdot 0 = 11.$$

Now we illustrate how to use the conditional expectations $E^{X=x}(Y \mid \varphi)$ in order to compute the *Z_1-conditional effect function $CTE_{Z_1;10}$* for the covariate $Z_1 = Sex$. Note that, in contrast to $Z_2 = College$ and $Z = (Z_1, Z_2)$, the covariate Z_1 is φ-measurable. In fact, Z_1 is the composition of φ and a map g, because

$$Z_1(\omega) = g\big(\varphi(\omega)\big) = \begin{cases} m, & \text{if } \varphi(\omega) = 7/8 \text{ or } \varphi(\omega) = 5/8 \\ f, & \text{if } \varphi(\omega) = 2/8, \end{cases}$$

(see RS-Lemma 2.34). Hence,

$$\begin{aligned} CTE_{Z_1;10}(Z_1) &\underset{P}{=} E\big(E^{X=1}(Y\,|\,\varphi) - E^{X=0}(Y\,|\,\varphi)\,\big|\,Z_1\big) && [(11.25)] \\ &\underset{P}{=} E(11 - 4\cdot 1_{\varphi=5/8}\,|\,Z_1) && [(11.32)] \\ &\underset{P}{=} 11 - 4\cdot E(1_{\varphi=5/8}\,|\,Z_1) && [\text{RS-Box 4.1 (xviii)}] \\ &\underset{P}{=} 11 - 4\cdot E(1_{\varphi=5/8}\,|\,1_{Z_1=f}) && [\sigma(1_{Z_1=f}) = \sigma(Z_1)] \\ &\underset{P}{=} 11 - 4\cdot\left(\frac{3}{4} - \frac{3}{4}\cdot 1_{Z_1=f}\right) && [\text{see Exercise 11-6}] \\ &\underset{P}{=} 8 + 3\cdot 1_{Z_1=f}\,. \end{aligned}$$

Hence, the causal conditional effects $CTE_{Z_1;10}(z)$ of treatment 1 compared to control given the value z of Z_1 are

$$CTE_{Z_1;10}(m) \;=\; 8+3\cdot 0 \;=\; 8$$

for males and

$$CTE_{Z_1;10}(f) \;=\; 8+3\cdot 1 \;=\; 11$$

for females. These effects are in fact identical to the two conditional expectation values $E(\tau_1-\tau_0\,|\,Z_1{=}m)$ and $E(\tau_1-\tau_0\,|\,Z_1{=}f)$ [see Def. 5.38 (ii)] that can be computed directly from Table 7.5.

Finally, according to Theorem 11.20, the *causal average total effect* is

$$\begin{aligned} ATE_{10} &= E[CTE_{\varphi;10}(\varphi)] \;=\; E[PFE_{\varphi;10}(\varphi)] \\ &= E(11-4\cdot 1_{\varphi=5/8}) \;=\; 11-4\cdot E(1_{\varphi=5/8}) \;=\; 11-4\cdot 3/6 \;=\; 9. \end{aligned}$$

This effect can also be directly computed from Table 7.5 as the expectation $E(\tau_1-\tau_0)$. Another alternative is to compute it from the Z_1-conditional effect variable $CTE_{Z_1;10}(Z_1)$ via

$$\begin{aligned} E\big(CTE_{Z_1;10}(Z_1)\big) &= E\big(8+3\cdot 1_{Z_1=f}\big) \\ &= 8+3\cdot E\big(1_{Z_1=f}\big) \\ &= 8+3\cdot\frac{2}{6} \;=\; 9. \end{aligned}$$

To summarize, this example shows how we can use the true propensity φ as a covariate for identifying average and conditional total effects, where we can condition on φ or on a measurable function of φ. The values of the causal total effect variable $CTE_{\varphi;10}(\varphi)$ are not only the conditional total effects given a value of the propensity φ, but we can also apply Equation (11.25) in order to compute a V-conditional effect function provided that V is a measurable function of φ. Furthermore, Theorem 11.20 allows to compute the causal average total effect from the φ-conditional effect variable. ◁

11.2 Methodological Implications of the Theory of True Propensities

Now we discuss the methodological implications of the theory of true propensities presented in section 11.1 for the analysis of causal effects. We distinguish two cases: in the first one, the true propensities are known, in the second they are unknown.

11.2.1 Known True Propensities

Remark 11.26 [Conditional Randomization] In the *true experiment*, the true propensities are fixed by the experimenter, for example, as a function of a measured covariate Z, such as the rated *neediness* of a person for a therapy. The experimenter may decide, for example, that a person with a *high* need receives the treatment with probability 5/6, a person with a *medium* need receives the treatment with probability 3/6, and person with a *low* need receives the treatment with probability 1/6. If a person is rated to have a *high* need, then the experimenter tosses a dice and treats the subject unless he tosses a six, and he proceeds in the same way with persons belonging to the other two categories, except for the different treatment probabilities and associated decision rules: If a person is categorized to have a *medium* need, then the experimenter treats the subject unless he tosses a four, five or six. If a person is rated to have low need, then the experimenter treats the person if he tosses a one. Such a procedure is called *conditional randomization.* ◁

Remark 11.27 [True Experiment] More generally, conditional randomization ensures

$$\forall\, x \in X(\Omega)\colon\ \varphi_x \underset{P}{=} P(X{=}x \,|\, D_X) \underset{P}{=} P(X{=}x \,|\, Z)\,. \tag{11.33}$$

Under Equation (11.33), the true propensities φ_x and the Z-conditional treatment probabilities $P(X{=}x \,|\, Z)$ are identical (at least P-almost surely).

To emphasize, according to Equation (11.33), the treatment probabilities do not have to be the same for each observational unit. Instead, the experimenter may determine the treatment probabilities according to his diagnostic criteria of severity of symptoms, his rated probability of success for the specific patient, or other optimality criteria (see, e. g., Pukelsheim, 2006; Berger & Wong, 2005). Of course, if several, say m, criteria are used for fixing the treatment probabilities, then the covariate $Z = (Z_1, \dots, Z_m)$ would be an m-variate random variable. Note that one or all of these covariates might also be continuous. If $Z_1, \dots, Z_m$ are real-valued, then one possibility is to fix the true treatment probabilities by

(a) choosing coefficients $\alpha_0, \alpha_1, \dots, \alpha_m$ according to the desired relevance for the treatment probability to be fixed,
(b) computing the conditional treatment probability according to the linear logistic equation
$$P(X{=}x \,|\, Z) \;=\; \frac{\exp(\alpha_0 + \alpha_1 \cdot Z_1 + \ldots + \alpha_m \cdot Z_m)}{1 + \exp(\alpha_0 + \alpha_1 \cdot Z_1 + \ldots + \alpha_m \cdot Z_m)}, \tag{11.34}$$
(c) assign an observational unit (person) to treatment x with (conditional) probability $P(X{=}x \,|\, Z_1{=}z_1, \dots, Z_m{=}z_m)$, which itself depends on the values $z_1, \dots, z_m$ of the unit on these covariates.

◁

Remark 11.28 [Identifying Conditional and Average Total Effects] If the true propensities $\varphi_x := P(X{=}x \,|\, D_X)$, $x \in X(\Omega)$, are known as described for a true experiment, then we may use them in order to identify conditional and average total effects. If, for example, there are only two values of X, say 0 and 1, and the true outcome variables τ_0 and τ_1 are P-unique, then the conditional expectations $E^{X=0}(Y \,|\, \varphi_0)$ and $E^{X=1}(Y \,|\, \varphi_1)$ are unbiased (see Cor. 11.15). Because $\varphi_0 = P(X{=}0 \,|\, D_X) = 1 - \varphi_1$, the σ-algebras generated by φ_0 and φ_1 are identical, which implies $E^{X=0}(Y \,|\, \varphi_0) \underset{P}{=} E^{X=0}(Y \,|\, \varphi_1)$ and $E^{X=1}(Y \,|\, \varphi_0) \underset{P}{=} E^{X=1}(Y \,|\, \varphi_1)$.

Note that, using $E^{X=0}(Y|\varphi_1)$ and $E^{X=1}(Y|\varphi_1)$ in the identification of conditional and average total effects, *no restrictive assumptions are necessary*. When determining the treatment probabilities, the experimenter only has to make sure that

$$P(X{=}0\,|\,D_X) \underset{P}{=} P(X{=}0\,|\,Z) \underset{P}{>} 0 \quad \text{and} \quad P(X{=}1\,|\,D_X) \underset{P}{=} P(X{=}1\,|\,Z) \underset{P}{>} 0. \tag{11.35}$$

Remember, according to Theorem RS-5.27, this condition is equivalent to P-uniqueness of τ_0 and τ_1. Also note that (11.35) implies $0 < P(X{=}1) < 1$, which is equivalent to the conjunction of $P(X{=}0) > 0$ and $P(X{=}1) > 0$ [see SN-Th. 3.52 (ii)].

According to Corollary 11.19, the expectations of the conditional expectations $E^{X=x}(Y\,|\,\varphi)$ and $E^{X=x}(Y\,|\,\varphi_x)$ are identical to the expectation of the true outcome variable τ_x [see Eq. (11.18)]. Hence, because $ATE_{xx'} = E(\tau_x) - E(\tau_{x'})$, the expectations of the conditional expectations $E^{X=x}(Y\,|\,\varphi_x)$, $E^{X=x}(Y\,|\,\varphi)$, $E^{X=x'}(Y\,|\,\varphi_x)$, and $E^{X=x'}(Y\,|\,\varphi)$ can be also be used for identifying the *causal average total effect* $ATE_{xx'}$. That is,

$$ATE_{xx'} = E\big(E^{X=x}(Y\,|\,\varphi_x)\big) - E\big(E^{X=x'}(Y\,|\,\varphi_{x'})\big) \tag{11.36}$$

$$= E\big(E^{X=x}(Y\,|\,\varphi)\big) - E\big(E^{X=x'}(Y\,|\,\varphi)\big). \tag{11.37}$$

Furthermore, the conditional expectations $E^{X=x}(Y|\varphi)$ and $E^{X=x'}(Y|\varphi)$ can also be used to compute the *φ-conditional total effect function $CTE_{\varphi;xx'}$*, that is,

$$CTE_{\varphi;xx'}(\varphi) \underset{P}{=} E^{X=x}(Y|\varphi) \;-\; E^{X=x'}(Y|\varphi) \tag{11.38}$$

[see Eq. (11.16)].

Finally, we can also compute the conditional expectations $E(\tau_x|V)$ and $E(\tau_{x'}|V)$ of the true outcome variables τ_x and $\tau_{x'}$ given an (X,φ)-measurable map V [see Eq. (11.23)]. Because

$$CTE_{V;xx'}(V) \underset{P}{=} E(\tau_x|V) - E(\tau_{x'}|V) \tag{11.39}$$

$$\underset{P}{=} E\big(E^{X=x}(Y\,|\,\varphi)\,\big|\,V\big) \;-\; E\big(E^{X=x'}(Y\,|\,\varphi)\,\big|\,V\big) \tag{11.40}$$

[see Eq. (11.25)], the conditional expectations $E^{X=x}(Y|\varphi)$ can be also be used for identifying the *causal V-conditional total effect function $CTE_{V;xx'}$* via Equation (11.40). $\lhd$

Gehört das Folgende nicht nach hinten?

11.2.2 Unknown True Propensities

Remark 11.29 [Quasi-Experiment] By definition, in a *quasi-experiment*, the true propensities can only be estimated. In this case, we need two kinds of assumptions, which entails that there are also two sources of errors. The *first* source of error is in the selection of the covariates. The critical question is: 'Did we select those covariates in $Z = (Z_1,\ldots,Z_m)$ for which the treatment variable X and a global potential confounder D_X are independent given these covariates of X?' [see Eq. (11.1)]. The *second* source of error is in the specification of the parameterization of the conditional probability $P(X{=}x\,|\,Z)$. If Z has only a few values such as *neediness* in the example above, then it is easy to specify a saturated parameterization of $P(X{=}x\,|\,Z)$, that is, a parameterization that always holds for the selected covariate Z. However, if Z has many values, is multivariate, or even continuous, then a correct parameterization may be difficult. If, for example, we base our analysis on the assumption that the parameterization specified in Equation (11.34) holds, then the critical question is, whether or not this assumption is correct, that is, 'Did we correctly specify the parameterization of $P(X{=}x\,|\,Z)$?' Of course, there are many other types of parametric functions for $P(X{=}x\,|\,Z)$. These functions are not necessarily logistic functions or they may include quadratic and higher order terms as well as product terms representing interactions between different covariates. Furthermore, it is also possible that there is no simple parametric function at all that describes $P(X{=}x\,|\,Z)$. In fact, there are many alternatives to logistic regression for estimating $P(X{=}x\,|\,Z)$, such as neuronal networks (e. g., King & Zeng, 2001), boosted regressions (e. g., McCaffrey, Ridgeway, & Morral, 2004), classification trees (e. g., Lee, Lessler, & Stuart, 2010), and smoothing splines generalized additive models (e. g., Keele, 2008; Woo, Reiter, & Karr, 2008). ◁

In the sequel we treat some theorems that are useful in covariate selection. They cannot only be used in selecting a covariate $Z = (Z_1, \ldots, Z_m)$ for which the treatment variable X and the global potential confounder D_X are Z-conditionally independent [see Eq. (11.1)], but also for finding a correct parameterization of $P(X{=}x\,|\,Z)$.

Remark 11.30 [Balancing] If W is a potential confounder of X, then according to Theorem 11.31, there is φ_x-conditional independence of W and the indicator $1_{X=x}$ and φ-conditional independence of W and X. ◁

Corollary 11.31 [Conditional Independence of X and a Potential Confounder W]
Let the Assumptions 11.1 (a) to (c) hold and let W be a potential confounder of X. Then

(i) $W \perp\!\!\!\perp 1_{X=x} \,|\, \varphi_x$.

If we additionally assume 11.1 (d), then

(ii) $W \perp\!\!\!\perp X \,|\, \varphi$

(iii) $W \perp\!\!\!\perp 1_{X=x} \,|\, \varphi$.

Proof

Proposition (i). This follows from Theorem 11.10 (i), Definition 4.11 (iv), and RS-Box 6.1 (vi).

Proposition (ii). This follows from Theorem 11.10 (ii), Definition 4.11 (iv), and RS-Box 6.1 (vi).

Proposition (iii). This follows from Theorem 11.10 (iii), Definition 4.11 (iv), and RS-Box 6.1 (vi).

Remark 11.32 [Identical Conditional Distributions] If x and x' are two values of X for which $P(X{=}x, \varphi{=}p), P(X{=}x', \varphi{=}p) > 0$, then $W \perp\!\!\!\perp X \,|\, \varphi$ implies

$$P_{W\,|\,X=x,\varphi=p} \;=\; P_{W\,|\,X=x',\varphi=p} \tag{11.41}$$

(see SN-Rem. 17.47). Hence, if $P(X{=}x, \varphi{=}p) > 0$ for all $x \in X(\Omega)$, then φ-conditional independence of X and D_X implies that the $(X{=}x,\, \varphi{=}p)$-conditional distributions of a potential confounder W of X are identical for all $x \in X(\Omega)$. Note that, by definition, a covariate Z of X is also a potential confounder of X. Therefore, Equation (11.41) also holds if W is one of the covariates in $Z = (Z_1, \ldots, Z_m)$. ◁

Remark 11.33 [Covariate Seclection] The fact that $W \perp\!\!\!\perp X \mid \varphi$ implies Equation (11.41) is often used in covariate selection aiming at finding a covariate Z of X such that

$$\forall x \in \{1, \ldots, J\}\colon\ P(X{=}x \mid Z) \underset{P}{=} P(X{=}x \mid D_X) \underset{P}{=} \varphi_x. \tag{11.42}$$

If we define

$$\forall\, x \in X(\Omega)\colon\ \pi_x := P(X{=}x \mid Z) \qquad \text{and} \qquad \pi := (\pi_1, \ldots, \pi_J), \tag{11.43}$$

and Equation (11.42) holds, then $W \perp\!\!\!\perp X \mid \varphi$ and $W \perp\!\!\!\perp X \mid \pi$ are equivalent. Hence, if x and x' are two values of X and p is a value of π for which $P(X{=}x, \pi{=}p), P(X{=}x', \pi{=}p) > 0$, then $W \perp\!\!\!\perp X \mid \pi$ implies

$$P_{W \mid X=x, \pi=p} \;=\; P_{W \mid X=x', \pi=p} \tag{11.44}$$

(see Rem. 11.32). Now, if, in an analysis of empirical data, we have to reject the hypothesis that these two conditional distributions of the potential confounder W of X are identical, then it follows that Equation (11.42) does not hold and we may look for another covariate Z^* of X satisfying

$$\forall x \in \{1, \ldots, J\}\colon\ P(X{=}x \mid Z^*) \underset{P}{=} P(X{=}x \mid D_X) \underset{P}{=} \varphi_x. \tag{11.45}$$

$\lhd$

Remark 11.34 [π-Conditional Mean-Independence of W from X] Note that $W \perp\!\!\!\perp X \mid \pi$ implies

$$E(W \mid X, \pi) \underset{P}{=} E(W \mid \pi) \tag{11.46}$$

(see RS-Th. 6.8). Furthermore, if x and x' are two values of X for which $P(X{=}x, \pi{=}p)$, $P(X{=}x', \pi{=}p) > 0$, then this implies

$$E(W \mid X{=}x, \pi{=}p) \;=\; E(W \mid X{=}x', \pi{=}p). \tag{11.47}$$

These equations can be used as well to check if Equations (11.42) and (11.43) actually hold for the selected covariate $Z = (Z_1, \ldots, Z_m)$. $\lhd$

Remark 11.35 [Did we Select the Right Covariates?] If, for example, X is dichotomous, we define $\pi := P(X{=}x \mid Z)$, and find that

$$E(W \mid X, \pi) \underset{P}{=} E(W \mid \pi)$$

does not hold, then we can conclude that $\pi \underset{P}{=} \varphi = P(X{=}1 \mid D_X)$ does not hold as well. That is, we did not select Z such that Equations (11.42), (11.43), and all their implications — such as unbiasedness of $E(Y \mid X, \pi)$ (see Cor. 11.15) — hold. $\lhd$

Remark 11.36 [Did we Consider a Correct Parameterization?] Even if we are sure that Equation (11.42) holds for the covariate Z of X, then each of the equations (11.44) to (11.47) can also be used to check if we consider a correct parameterization of the conditional probabilities $P(X{=}x \mid Z)$, $x \in X(\Omega)$. For example, if X is binary, we define the random variable

$$\rho := \frac{\exp(\alpha_0 + \alpha_1 \cdot Z_1 + \ldots + \alpha_m \cdot Z_m)}{1 + \exp(\alpha_0 + \alpha_1 \cdot Z_1 + \ldots + \alpha_m \cdot Z_m)}, \tag{11.48}$$

and find in an empirical data analysis that the hypothesis

$$E(W|X,\rho) \underset{P}{=} E(W|\rho)$$

does not hold, then we can conclude that either $P(X{=}x|Z)$ is not a version of the true propensity φ — which means that $Z = (Z_1, \ldots, Z_m)$ does not contain the right covariates in the sense that Equation (11.42) holds — or Equation (11.48) does not specify a (correct) parameterization of the conditional probability $P(X{=}x|Z)$. ◁

11.3 Z-Conditional Propensities for Total Effects

If the true outcome variables $\tau_0, \tau_1, \ldots, \tau_J$ are P-unique, then, according to Corollary 11.15, the conditional expectations $E^{X=x}(Y|\varphi_x)$, $x \in X(\Omega)$, and $E(Y|X,\varphi)$ are unbiased. Hence, in section 10.2 we discussed fixing the true propensities $\varphi_x = P(X{=}x|D_X)$ or selecting covariates such that the *true* propensities are estimated. The criterion for covariate selection discussed in section 11.1 has been presented in Theorem 11.4, namely

$$\forall\, x \in X(\Omega)\colon\ P(X{=}x|D_X) \underset{P}{=} P(X{=}x|Z).$$

This equation can not only be used to fix the true propensities as functions of the covariate Z; instead, its implications can also serve for covariate selection (see Remarks 11.30 to 11.36).

Remark 11.37 [Another Criterion for Covariate Selection] Another criterion for covariate selection does not aim at estimating the *true* propensities. Instead, it suffices to select a (possibly multivariate) covariate Z of X such that

$$\forall\, x \in X(\Omega)\colon\ P(X{=}x|Z) \underset{P}{=} P(X{=}x|\tau_x, Z). \tag{11.49}$$

This equation is equivalent to Z-conditional independence of $1_{X=x}$ and the true outcome variable τ_x, for each value x of X. In Box 7.2 we introduced the shortcut $\forall x : \tau_x \perp\!\!\!\perp 1_{X=x} | Z$ for this equation. Unfortunately, in practice, Equation (11.49) does not lend itself for covariate selection, because it is not empirically testable (see sect. 7.5).

However, Equation (11.49) does not only follow from $D_X \perp\!\!\!\perp X | Z$ (see Th. 8.31). If all τ_x, $x \in X(\Omega)$, are P-unique, then, according to Theorem 9.35, it also follows from

$$\forall\, x \in X(\Omega)\colon\ E^{X=x}(Y|D_X) \underset{P^{X=x}}{=} E^{X=x}(Y|Z). \tag{11.50}$$

This equation is abbreviated by $\forall x\colon Y \vDash D_X | (X{=}x, Z)$. In fact, if all τ_x, $x \in X(\Omega)$, are P-unique, and Equation (11.50) holds, then

$$\forall\, x \in X(\Omega)\colon\ \tau_x \underset{P}{=} E^{X=x}(Y|Z). \tag{11.51}$$

Hence, if we can assume that all τ_x, $x \in X(\Omega)$, are P-unique, then another strategy of covariate selection aims at finding a covariate Z of X such that Equation (11.50) holds. ◁

Remark 11.38 [Another Strategy for Covariate Selection] Because

$$E(Y|X,Z) \underset{P}{=} \sum_{x \in X(\Omega)} E^{X=x}(Y|Z) \cdot 1_{X=x} \tag{11.52}$$

[see RS-Eq. (5.36)], we can also summarize the alternative strategy of data analysis as follows: Find a (possibly multivariate) covariate $Z = (Z_1, \ldots, Z_m)$ of X such that

$$E(Y|X,D_X) \underset{P}{=} E(Y|X,Z). \tag{11.53}$$

If Equation (11.53) holds, we define $\pi = (\pi_1, \ldots, \pi_J)$ with $\pi_x = P(X{=}x|Z)$, $x = 1, \ldots, J$, and we assume that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique, then we can identify the causal π-conditional total effect variables by

$$CTE_{\pi;xx'}(\pi) \underset{P}{=} E^{X=x}(Y|\pi) - E^{X=x'}(Y|\pi), \tag{11.54}$$

and the causal average total effect by

$$ATE_{xx'} = E\big(CTE_{\pi;xx'}(\pi)\big) \underset{P}{=} E\big(E^{X=x}(Y|\pi)\big) - E\big(E^{X=x'}(Y|\pi)\big). \tag{11.55}$$

In general, the conditional total effect function $CTE_{\pi;xx'}$ is less fine-grained than the conditional total effect function $CTE_{Z;xx'}$, which, under Equation (11.53), is also a version of the true total effect function specified by $CTE_{D_X;xx'}(D_X) \underset{P}{=} E^{X=x}(Y|D_X) - E^{X=x'}(Y|D_X)$. ◁

Remark 11.39 [Z-Conditional Propensity] Let $x \in X(\Omega)$ be a value of X and Z be a (possibly multivariate) covariate of X.Then a version

$$\pi_x := P(X{=}x|Z) \tag{11.56}$$

of the Z-conditional probability of the event $\{X{=}x\}$ is called a *Z-conditional propensity of x*. Hence, π_x is a random variable that is measurable with respect to Z and because Z is a covariate of X, the propensity π_x and Z itself are measurable with respect to D_X [see Def. 4.11 (iv) and Rem. 4.16]. Furthermore, if $z \in Z(\Omega)\,\Omega'_Z$ is a value of Z such that $P(Z{=}z) > 0$, then $P(X{=}x|Z{=}z)$ is also called the *propensity score of x given the value z of* Z (cf. Rosenbaum & Rubin, 1983a, 1984). Note that π_x is not necessarily identical to the true propensity φ_x that has been defined in Equation (11.5), that is, Equation (11.53) does not necessarily hold. Defining φ_x we condition on a global confounder D_X of X, whereas defining π_x we condition on a covariate Z of X. While $\sigma(Z) \subset \sigma(D_X)$, the two σ-algebras are not necessarily identical. ◁

Remark 11.40 [Only J-Propensities] Note that we only need J such propensities if there are $J+1$ different values of X, because

$$\pi_0 \underset{P}{=} 1 - (\pi_1 + \cdots + \pi_J) \tag{11.57}$$

(see Exercise 11-3). Hence, we can gather the propensities in the J-dimensional vector $\pi := (\pi_1, \ldots, \pi_J)$. Note that π_0 is function of Z because each π_x is a function of Z, $x \in X(\Omega)$. Hence, according to RS-Lemma 2.35, π_0 is measurable with respect to Z. If X is binary with values 0 and 1, then

$$P(X{=}0|Z) \underset{P}{=} 1 - P(X{=}1|Z) \tag{11.58}$$

(see again Exercise 11-3). ◁

Remark 11.41 [The Propensity Given Z is a Covariate] Also note that $\pi = (\pi_1, \dots, \pi_J)$ as well as each π_x, $x = 1, \dots, J$, are covariates of X, that is, π and each π_x are measurable with respect to D_X (see RS-Th. 2.16), provided that the random variable Z is a covariate of X and therefore is D_X-measurable [see again Def. 4.11 (iv) and Rem. 4.16]. Following the same arguments, according to RS-Theorem 2.34, π_0 is also a covariate of X. ◁

In the following lemma we show that $P(X{=}x \,|\, \pi_x)$ and $P(X{=}x \,|\, \pi)$ are versions of the conditional probability $P(X{=}x \,|\, Z)$. This lemma prepares Theorems 11.44 and 11.48.

Lemma 11.42 [$P(X{=}x \,|\, \pi_x)$ **and** $P(X{=}x \,|\, \pi)$ **are Versions of** $P(X{=}x \,|\, Z)$]
Let X and Z be random variables on the probability space $(\Omega, \mathscr{A}, P)$, *let* $\pi_x := P(X{=}x \,|\, Z)$, *for all* $x \in X(\Omega)$, *and* $\pi := (\pi_1, \dots, \pi_J)$. *Then*

$$\forall\, x \in X(\Omega){:}\;\; P(X{=}x \,|\, Z) \underset{P}{=} P(X{=}x \,|\, \pi_x) \underset{P}{=} P(X{=}x \,|\, \pi). \tag{11.59}$$

Proof

According to RS-Remark 4.12, $P(X{=}x \,|\, Z)$ *is another notation for the conditional expectation* $E(1_{X=x} \,|\, Z)$, *where* $1_{X=x}$ *denotes the indicator variable of the event* $\{X{=}x\}$. *Hence, for all* $x \in X(\Omega)$,

$$\begin{aligned} P(X{=}x \,|\, Z) &\underset{P}{=} E\big(P(X{=}x \,|\, Z) \,\big|\, \pi_x\big) && \big[\sigma\big(P(X{=}x \,|\, Z)\big) \subset \sigma(\pi_x),\ \textit{RS-Box } 4.1\ (\text{xi})\big] \\ &\underset{P}{=} P(X{=}x \,|\, \pi_x). && [\sigma(\pi_x) \subset \sigma(Z),\ \textit{RS-Box } 4.1\ (\text{xiii})] \end{aligned}$$

The proof of the second equation is analog. We simply have to replace π_x *by* π *and then use transitivity of the relation* $\underset{P}{=}$ *(see SN-Remark 2.73).*

Remark 11.43 [Z**-Conditional Independence of** X **and True Outcome Variables**] Note that

$$\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z \quad \Leftrightarrow \quad P(X{=}x \,|\, Z, \tau_x) \underset{P}{=} P(X{=}x \,|\, Z) \tag{11.60}$$

(see RS-Th. 6.6). According to the following theorem, this condition has an implication for the conditional probability $P(X{=}x \,|\, \pi_x)$. ◁

Theorem 11.44 [An Implication of Z-Conditional Independence of τ_x and $1_{X=x}$]
Let the Assumptions 11.1 (a), (b), (e) to (g), and (i) hold, and define $\pi_x \underset{P}{=} P(X{=}x \,|\, Z)$. *Then*

$$\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z \quad \Rightarrow \quad P(X{=}x \,|\, \pi_x, \tau_x) \underset{P}{=} P(X{=}x \,|\, \pi_x). \tag{11.61}$$

Proof:

$$\begin{aligned} & P(X{=}x \,|\, \pi_x, \tau_x) \\ \underset{P}{=}\ & E\big(P(X{=}x \,|\, Z, \tau_x) \,\big|\, \pi_x, \tau_x\big) && [\sigma(\pi_x, \tau_x) \subset \sigma(Z, \tau_x),\ \textit{RS-Box } 4.1\ (\text{xiii})] \\ \underset{P}{=}\ & E\big(P(X{=}x \,|\, Z) \,\big|\, \pi_x, \tau_x\big) && [\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z,\ (11.60)] \\ \underset{P}{=}\ & P(X{=}x \,|\, Z) && \big[\sigma(P(X{=}x \,|\, Z) \subset \sigma(\pi_x, \tau_x),\ \textit{RS-Box } 4.1\ (\text{xi})\big] \\ \underset{P}{=}\ & P(X{=}x \,|\, \pi_x). && [(11.59)] \end{aligned}$$

Remark 11.45 [An Implication of Z_x-Conditional Independence of τ_x and $1_{X=x}$] According to RS-Lemma 6.6, Proposition (11.61) is equivalent to

$$\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, Z \quad \Rightarrow \quad \tau_x \perp\!\!\!\perp 1_{X=x} \,|\, \pi_x\,. \tag{11.62}$$

$\triangleleft$

If Z is a covariate of X, then $\pi_x = P(X{=}x \,|\, Z)$ is a covariate of X as well. Therefore, the second part of Theorem 7.19 immediately yields the following corollary.

Corollary 11.46 [$\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, \pi_x$ Implies Unbiasedness of $E^{X=x}(Y \,|\, \pi_x)$]
Let the Assumptions 11.1 (a), (b), (e) to (g), and (i) hold and assume that τ_x, $x \in X(\Omega)$, is P-unique. Furthermore, for $x \in X(\Omega)$, define $\pi_x := P(X{=}x \,|\, Z)$. Then

$$\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, \pi_x \quad \Rightarrow \quad E^{X=x}(Y \,|\, \pi_x) \vdash D_X\,. \tag{11.63}$$

Hence, according to this corollary, if the true outcome variable τ_x is P-unique, and $1_{X=x}$ and the true outcome variable τ_x are π_x-conditionally independent (i. e., if $\tau_x \perp\!\!\!\perp 1_{X=x} \,|\, \pi_x$), then $E^{X=x}(Y \,|\, \pi_x)$ is unbiased.

Remark 11.47 [Strong Ignorability] If we consider $\tau := (\tau_0, \tau_1, \ldots, \tau_J)$, then

$$\tau \perp\!\!\!\perp X \,|\, Z \quad \Leftrightarrow \quad \forall\, x \in X(\Omega){:}\;\; P(X{=}x \,|\, Z, \tau) \underset{P}{=} P(X{=}x \,|\, Z) \tag{11.64}$$

(see RS-Th. 6.5). Remember, if we assume $\tau \perp\!\!\!\perp X \,|\, Z$ and that each τ_x, $x \in X(\Omega)$, is P-unique, then this is the assumption of strong ignorability given Z (see Rem. 7.16). $\triangleleft$

Theorem 11.48 [π-Conditional Independence of True Outcome Variables and X]
Let the Assumptions 11.1 (a), (b), and (e) to (i) hold. Furthermore, for all $x \in X(\Omega)$, define $\pi_x := P(X{=}x \,|\, Z)$ and $\pi := (\pi_1, \ldots, \pi_J)$. Then

(i) $D_X \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, Z \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, \pi$

(ii) If all τ_x, $x \in X(\Omega)$, are P-unique, then $Y \vDash D_X \,|\, (X, Z) \;\Rightarrow\; \tau \perp\!\!\!\perp X \,|\, \pi$.

Proof:

Proposition (i). The first implication in this proposition is Theorem 8.31 (i). For the proof of the second implication, we have to show

$$\tau \perp\!\!\!\perp X \,|\, Z \quad \Rightarrow \quad \forall\, x \in X(\Omega){:}\;\; P(X{=}x \,|\, \pi, \tau) \underset{P}{=} P(X{=}x \,|\, \pi)$$

(see RS-Th. 6.5). Hence, for all $x \in X(\Omega)$,

$$\begin{aligned}
&P(X{=}x \,|\, \pi, \tau) \\
\underset{P}{=}\; &E\big(P(X{=}x \,|\, Z, \tau) \,\big|\, \pi, \tau\big) && [\sigma(\pi,\tau) \subset \sigma(Z,\tau),\ \textit{RS-Box } 4.1\ \text{(xiii)}] \\
\underset{P}{=}\; &E\big(P(X{=}x \,|\, Z) \,\big|\, \pi, \tau\big) && [\tau \perp\!\!\!\perp X \,|\, Z,\ (11.64)] \\
\underset{P}{=}\; &P(X{=}x \,|\, Z) && \big[\sigma\big(P(X{=}x \,|\, Z)\big) \subset \sigma(\pi,\tau),\ \textit{RS-Box } 4.1\ \text{(xi)}\big] \\
\underset{P}{=}\; &P(X{=}x \,|\, \pi). && [(11.59)]
\end{aligned}$$

Proposition (ii). According to Theorem 9.38, $Y \vDash D_X \,|\, (X,Z) \Rightarrow \tau \perp\!\!\!\perp X \,|\, Z$, if all τ_x, $x \in X(\Omega)$, are P-unique. Hence, Proposition (ii) follows from (i).

According to Theorem 11.48, Z-conditional independence of the putative cause variable X and the global potential confounder D_X of X as well as the conjunction of (X,Z)-conditional mean-independence of the outcome variable Y from D_X and P-uniqueness of all true outcome variables τ_x, $x \in X(\Omega)$, imply π-conditional independence of X and the multivariate true outcome variable τ.

If Z is a covariate of X, then $\pi_x = P(X{=}x \,|\, Z)$ is a covariate of X as well. Therefore, Theorem 7.21 yields the following corollary.

Corollary 11.49 [$\tau \perp\!\!\!\perp X \,|\, \pi$ **Implies Unbiasedness of** $E^{X=x}(Y \,|\, \pi)$]
Let the Assumptions 11.1 (a), (b), and (e) to (i) hold and assume that all τ_x, $x \in X(\Omega)$, are P-unique. Furthermore, for all $x \in X(\Omega)$, define $\pi_x := P(X{=}x \,|\, Z)$ and $\pi := (\pi_1, \dots, \pi_J)$. Then

$$\tau \perp\!\!\!\perp X \,|\, \pi \;\Rightarrow\; E(Y \,|\, X, \pi) \vdash D_X. \tag{11.65}$$

11.3.1 Methodological Implications

The conclusion of Theorem 11.48, conditional independence of X and the multivariate true outcome variable τ given π, does not only imply unbiasedness of the conditional expectation $E(Y \,|\, X, \pi)$. It also has other important methodological implications.

Remark 11.50 [Reduction of Dimensionality] Conditioning on $\pi := (\pi_1, \dots, \pi_J)$ instead of the original covariate $Z := (Z_1, \dots, Z_m)$ can be advantageous. For a small number $J+1$ of values of X and a large number of covariates $Z_1, \dots, Z_m$ this may simplify the conditional expectation $E(Y \,|\, X, \pi)$ considerably compared to the conditional expectation $E(Y \,|\, X, Z)$. If, for instance, X represents a treatment variable and there are $J+1 = 2$ treatment conditions and $m = 5$ covariates $Z_1, \dots, Z_5$, then conditional independence of $\tau = (\tau_0, \tau_1)$ and X given the covariate $Z := (Z_1, \dots, Z_5)$ implies that τ and X are also conditionally independent given $\pi_1 := P(X{=}1 \,|\, Z)$. In contrast to Z, the propensity π_1 is a *univariate* random variable. This fact allows us to compute the average total effect ATE_{10} not only from the conditional expectations $E^{X=x}(Y \,|\, Z_1, \dots, Z_5)$, $x = 0, 1$, but also from the conditional expectations $E^{X=x}(Y \,|\, \pi_1)$ $x = 0, 1$, and from the π_1-conditional prima facie effect function $PFE_{\pi_1;10}$ (see Th. 6.34). Furthermore, the causal π_1-conditional total effect function $CTE_{\pi_1;10}$ is identical to the π_1-conditional prima facie effect function $PFE_{\pi_1;10}$ (see Cor. 6.37).

Note, however, that the Z-conditional total effect function $CTE_{Z;10}$ may be more informative than the π_1-conditional total effect function $CTE_{\pi_1;10}$, because $\pi_1 := P(X{=}1 \,|\, Z)$ may not be a one-to-one function of Z. Furthermore, the interpretation of π_1 may not be as easy as the interpretation of the covariate Z. If, for example, Z is a pretest, then we exactly know the meaning of the $(Z{=}z)$-conditional treatment effect of a person scoring z on the pretest. These conditional effects immediately answer questions such as 'Is the treatment effect on the outcome variable Y higher for those having a low score on the pretest as compared to those having a high pretest score?' In contrast, the meaning of the $(\pi_1 = p)$-conditional treatment effect of a person scoring p on the propensity π_1 usually is less clear.

Finally, if there are many values of X and few covariates it will be more convenient and more informative to use the conditional expectations $E^{X=x}(Y|Z)$ rather than the conditional expectations $E^{X=x}(Y|\pi)$ in identifying causal total effects and causal conditional effect functions. For example, if there are six treatment conditions and only one single univariate covariate Z for which $\tau \perp\!\!\!\perp X | Z$ holds, then we would need five different propensities $\pi_x = P(X=x|Z)$, $x = 1,\dots,5$, and six conditional expectations $E^{X=x}(Y|\pi_x)$, $x \in X(\Omega)$, instead of just six conditional expectations $E^{X=x}(Y|Z)$ for identifying all causal total effects and causal conditional effect functions. In this case, using the six conditional expectations $E^{X=x}(Y|Z)$ would be preferable to using five propensities $\pi_x = P(X=x|Z)$ plus six conditional expectations $E^{X=x}(Y|\pi_x)$. Furthermore, using the six conditional expectations $E^{X=x}(Y|Z)$ would yields even more informative causal conditional effect functions because the propensities π_x are measurable with respect to Z.

Alternatively, we may use the conditional expectations $E^{X=x}(Y|\pi)$ with $\pi = (\pi_1,\dots,\pi_5)$, a five-dimensional random variable, although Z is just one-dimensional. Because, in empirical applications, the propensities $\pi_x = P(X=x|Z)$ usually have to be estimated, using propensities instead of covariates is much more error-prone in such a case with many values of X and only few covariates. $\lhd$

Remark 11.51 [Fraud Prevention] If $Z = (Z_1,\dots,Z_m)$ is chosen such that $D_X \perp\!\!\!\perp X | Z$ holds, then the outcome variable Y is not involved in the selection of the covariates $Z_1,\dots,Z_m$. In other words, in empirical applications, we can choose Z without knowing the outcome variable Y and even before Y is assessed. This can be used to prevent that a covariate Z is selected such that $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$ describes the desired total effect function and $E\big(E^{X=x}(Y|Z)\big) - E\big(E^{X=x'}(Y|Z)\big)$ the desired average effect of the treatment. Covariate selection before the outcome variable Y is assessed does not imply that we have to conduct a propensity score analysis estimating the π-conditional expectations $E^{X=x}(Y|\pi)$ and the π-conditional effect functions. Instead or in addition, we may also estimate the Z-conditional expectations $E^{X=x}(Y|Z)$ and the Z-conditional effect functions $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$. $\lhd$

11.3.2 Numerical Examples

Table 7.5 gives an example for conditional independence of the true outcome variable $\tau = (\tau_0, \tau_1)$ and the treatment variable X given the covariate $Z := (Z_1, Z_2)$, where Z_1 = *sex* and Z_2 = *college*. In this table, the covariate Z is *two-dimensional*. Remember, this example is constructed such that the person variable U is a global potential confounder implying

$$E^{X=x}(Y|D_X) \underset{P}{=} E^{X=x}(Y|U) \tag{11.66}$$

and

$$P(X=1|D_X) \underset{P}{=} P(X=1|U). \tag{11.67}$$

Note that, in this example, there is also conditional independence of τ and X given the *one-dimensional* propensity $\pi_1 = P(X=1|Z)$. In section 7.4.2, we already checked conditional independence of τ and X given the two-dimensional covariate Z, which implies

Table 11.2. Conditional expectation values $E^{X=x}(Y|\pi_1=p)$ in the example of Table 7.5

	Propensity score		
Treatment	$p=6/8$	$p=5/8$	$p=2/8$
$x=0$	72.00	97.50	111.00
$x=1$	83.00	102.50	122.00

$\tau \perp\!\!\!\perp X \mid \pi_1$ [see Th. 11.48 (i)]. According to Theorem 7.21, this in turn implies that the conditional expectations $E^{X=0}(Y|\pi_1)$, $E^{X=1}(Y|\pi_1)$, and $E(Y|X,\pi_1)$ are unbiased, provided that the true outcome variables τ_0 and τ_1 are P-unique. According to RS-Theorem 5.27, P-uniqueness of τ_0 and τ_1 follows from $P(X=0|U)>0$ and $P(X=1|U)>0$, respectively.

Also note that, in this example, the Z-conditional propensity π_1 differs from the true propensity φ_1. This entails that the conditional total effect functions $CTE_{\varphi;10}$ and $CTE_{\pi;10}$ differ from each other. However, both are causal total effect functions and both can be used for identifying the causal average total effect ATE_{10}.

In this example, there are three different Z-conditional propensity scores for treatment 1. The first two units have the Z-conditional propensity score $P(X=1|Z_1=m, Z_2=no)$ = 6/8, whereas the next two units have the propensity score $P(X=1|Z_1=m, Z_2=yes)$ = 5/8, and the last two units have the propensity score $P(X=1|Z_1=f, Z_2=yes) = 2/8$ for treatment 1 (see Table 7.5). These propensity scores are the values of the propensity $\pi_1 = P(X=1|Z)$. Because X is dichotomous with values 0 and 1, we only need one a single propensity $\pi_1 = P(X=1|Z_1, Z_2)$. The second propensity is $\pi_0 = P(X=0|Z_1,Z_2) = 1 - P(X=1|Z_1,Z_2)$.

In this example, π_1 has only three different values. Therefore, the conditional expectations $E^{X=x}(Y|\pi_1)$, $x=0,1$, can be parameterized by

$$E^{X=x}(Y|\pi_1) = \gamma_0^{(x)} + \gamma_1^{(x)} \cdot 1_{\pi_1=5/8} + \gamma_2^{(x)} \cdot 1_{\pi_1=2/8}, \quad x=0,1, \tag{11.68}$$

where $1_{\pi_1=5/8}$ and $1_{\pi_1=2/8}$ denote the indicator variables for the events $\{\pi_1=5/8\}$ and $\{\pi_1=2/8\}$, respectively.

For treatment $x=0$,

$$\begin{aligned} E^{X=0}(Y|\pi_1) &= \gamma_0^{(0)} + \gamma_1^{(0)} \cdot 1_{\pi_1=5/8} + \gamma_2^{(0)} \cdot 1_{\pi_1=2/8} \\ &= 72 + 25.5 \cdot 1_{\pi_1=5/8} + 39 \cdot 1_{\pi_1=2/8}, \end{aligned} \tag{11.69}$$

and for treatment $x=1$

$$\begin{aligned} E^{X=1}(Y|\pi_1) &= \gamma_0^{(1)} + \gamma_1^{(1)} \cdot 1_{\pi_1=5/8} + \gamma_2^{(1)} \cdot 1_{\pi_1=2/8} \\ &= 83 + 19.5 \cdot 1_{\pi_1=5/8} + 39 \cdot 1_{\pi_1=2/8}. \end{aligned} \tag{11.70}$$

The γ-coefficients in these equations can be computed from the conditional expectation values $E^{X=x}(Y|\pi_1=p)$ displayed in Table 11.2 (see Exercise 11-5). These conditional expectation values $E^{X=x}(Y|\pi_1=p)$ have be computed from the parameters displayed in Table 7.5 using

$$\begin{aligned} E^{X=x}(Y|\pi_1=p) &= \sum_u E^{X=x}(Y|\pi_1=p, U=u) \cdot P^{X=x}(U=u|\pi_1=p) \\ &= \sum_u E^{X=x}(Y|U=u) \cdot P^{X=x}(U=u|\pi_1=p), \end{aligned}$$

just in the same way as in Example 11.25.

Because the conditional expectations $E^{X=x}(Y|\pi_1)$ are unbiased, the prima facie effect variable

$$\begin{aligned} PFE_{\pi_1;10}(\pi_1) &= E^{X=1}(Y|\pi_1) - E^{X=0}(Y|\pi_1) \\ &= 83 + 19.5 \cdot 1_{\pi_1=5/8} + 39 \cdot 1_{\pi_1=2/8} - (72 + 25.5 \cdot 1_{\pi_1=5/8} + 39 \cdot 1_{\pi_1=2/8}) \\ &= 11 - 6 \cdot 1_{\pi_1=5/8} \end{aligned} \tag{11.71}$$

is also the causal conditional total effect variable $CTE_{\pi_1;10}(\pi_1)$. Hence, the causal conditional total effect $CTE_{\pi_1;10}(6/8)$ of treatment 1 compared to the control on the outcome variable Y given $\pi_1 = 6/8$ is

$$CTE_{\pi_1;10}(6/8) = PFE_{\pi_1;10}(6/8) = 11 - 6 \cdot 0 = 11,$$

the causal conditional total effect $CTE_{\pi_1;10}(5/8)$ of treatment 1 compared to the control on the outcome variable Y given $\pi_1 = 5/8$ is

$$CTE_{\pi_1;10}(5/8) = PFE_{\pi_1;10}(5/8) = 11 - 6 \cdot 1 = 5,$$

and the causal conditional total effect $CTE_{\pi_1;10}(2/8)$ of treatment 1 compared to the control on the outcome variable Y given $\pi_1 = 2/8$ is

$$CTE_{\pi_1;10}(2/8) = PFE_{\pi_1;10}(2/8) = 11 - 6 \cdot 1 = 11.$$

According to Theorem 6.34 and using Equation (11.71), the causal average total effect is

$$E[PFE_{\pi_1;10}(\pi_1)] = E(11 - 6 \cdot 1_{\pi_1=5/8}) = 11 - 6 \cdot E(1_{\pi_1=5/8}) = 11 - 6 \cdot 2/6 = 9,$$

which is the same result as already obtained in Example 11.25.

11.4 Weighting the Outcome Variable

Now we will show how to compute the average total effects by choosing appropriate weights for the outcome variable Y. The only assumption required is that the conditional expectation $E(Y|X,Z)$ is unbiased.

Remark 11.52 [Implications of Unbiasedness of $E(Y|X,Z)$] According to Definition 6.13, unbiasedness of the conditional expectation $E(Y|X,Z)$ implies that all true outcome variables τ_x, $x \in X(\Omega)$, are P-unique, which in turn implies $P(X=x\,|\,D_X) \underset{P}{>} 0$, for all $x \in X(\Omega)$ (see RS-Th. 5.27). If Z is a covariate of X, then $\sigma(Z) \subset \sigma(D_X)$ [see Def. 4.11 (iv), Rem. 4.16], and, according to RS-Box 5.1 (iv), this implies $P(X=x|Z) \underset{P}{>} 0$. Finally, note that $P(X=x|Z) \underset{P}{>} 0$ implies that there is a version $\pi_x \in \mathscr{P}(X=x|Z)$, that is, a Z-conditional probability of $\{X=x\}$, such that $\pi_x > 0$ (see Exercise ?? ?). Hence, if Z is a covariate of X, then

$$E(Y|X,Z) \text{ is unbiased} \quad \Rightarrow \quad \forall\, x = 0,1,\ldots,J\colon\ \exists\, \pi_x \in \mathscr{P}(X=x|Z)\colon\ \pi_x > 0. \tag{11.72}$$

$\triangleleft$

Now, for all $x = 0, 1, \ldots, J$, let $\pi_x \in \mathscr{P}(X{=}x \,|\, Z)$ be such that $\pi_x > 0$. Then the *weighted outcome variable* referred to in Lemma 11.53 is:

$$Y_w \;:=\; Y \cdot \left(\sum_{x=0}^{J} 1_{X=x} \cdot \frac{P(X{=}x)}{\pi_x} \right), \tag{11.73}$$

Also note that the weighting variable, by which Y is multiplied, is a function of X and Z (see Exercise 11-10). The crucial ingredients of this weighting variable are the Z-conditional propensities $\pi_x = P(X{=}x \,|\, Z)$, the values of which are the $(Z{=}z)$-conditional probabilities $P(X{=}x \,|\, Z{=}z) = E(1_{X=x} \,|\, Z{=}z)$ [see the general definition of these conditional expectation values by RS-Eq. (4.24)].

The following lemma will be used in proving Theorem 11.54. In this lemma $E^{X=x}(Y_w \,|\, Z)$ denotes the Z-conditional expectation of Y_w with respect to the probability measure $P^{X=x}$.

Lemma 11.53 [The Conditional Expectation $E^{X=x}(Y_w \,|\, Z)$]
Let the Assumptions 11.1 (a) to (d) and (g) hold, let $\pi_x \in \mathscr{P}(X{=}x \,|\, Z)$ such that $\pi_x > 0$, and let Y_w denote the weighted outcome variable defined in Equation (11.73). Then

$$Y_w \underset{P^{X=x}}{=} Y \cdot \frac{P(X{=}x)}{\pi_x} \tag{11.74}$$

and

$$E^{X=x}(Y_w \,|\, Z) \underset{P^{X=x}}{=} E^{X=x}(Y \,|\, Z) \cdot \frac{P(X{=}x)}{\pi_x}. \tag{11.75}$$

Proof:

For all $x' \in X(\Omega)$, let $\pi_{x'} \in \mathscr{P}(X{=}x' \,|\, Z)$ such that $\pi_{x'} > 0$. Because $P^{X=x}(X{\neq}x) = 0$, we can conclude: $P^{X=x}(1_{X=x'}{=}0) = 1$, for all $x' \neq x$. Hence,

$$\begin{aligned} Y_w &= Y \cdot \left(\sum_{x'=0}^{J} 1_{X=x'} \cdot \frac{P(X{=}x')}{\pi_{x'}} \right) && [(11.73)] \\ &\underset{P^{X=x}}{=} Y \cdot \frac{P(X{=}x)}{\pi_x}, \end{aligned}$$

which proves Equation (11.74). Using this equation yields

$$\begin{aligned} E^{X=x}(Y_w \,|\, Z) &\underset{P^{X=x}}{=} E^{X=x}\left(Y \cdot \frac{P(X{=}x)}{\pi_x} \,\middle|\, Z \right) && [(11.74), \textit{RS-Box } 4.1 \text{ (xiv)}] \\ &\underset{P^{X=x}}{=} E^{X=x}(Y \,|\, Z) \cdot \frac{P(X{=}x)}{\pi_x}. && [\textit{RS-Box } 4.1 \text{ (xvi)}] \end{aligned}$$

If $P^{X=x}(Z{=}z) > 0$, then Equation (11.75) implies

$$E^{X=x}(Y_w \,|\, Z{=}z) \;=\; E^{X=x}(Y \,|\, Z{=}z) \cdot \frac{P(X{=}x)}{P(X{=}x \,|\, Z{=}z)}. \tag{11.76}$$

In Theorem 11.54 we present an identification equation for the expectation $E(\tau_x)$ of a true outcome variable τ_x.

Theorem 11.54 [Identification of $E(\tau_x)$]
Let the Assumptions 11.1 (a) to (d) and (g) hold, assume that $E^{X=x}(Y|Z)$ is unbiased, and let $\pi_x \in \mathscr{P}(X{=}x|Z)$ such that $\pi_x > 0$. Then

$$E(\tau_x) = E^{X=x}\Big(E^{X=x}(Y|Z) \cdot \frac{P(X{=}x)}{\pi_x}\Big) \tag{11.77}$$

$$= E^{X=x}(Y_w). \tag{11.78}$$

Proof:
Note that unbiasedness of $E^{X=x}(Y|Z)$ implies that $\tau_x = E^{X=x}(Y|D_X)$ is P-unique, which implies that $E^{X=x}(Y|Z)$ is P-unique [see RS-Box 5.1 (iv)]. Hence,

$$
\begin{aligned}
& E^{X=x}(Y_w) \\
&= E^{X=x}\big(E^{X=x}(Y_w|Z)\big) && [\textit{RS-Box } 4.1 \text{ (iv)}] \\
&= E^{X=x}\Big(E^{X=x}(Y|Z) \cdot \frac{P(X{=}x)}{\pi_x}\Big) && [(11.75)] \\
&= \int E^{X=x}(Y|Z) \cdot \frac{P(X{=}x)}{\pi_x}\, dP^{X=x} && [\textit{RS-}(3.17), (3.21)] \\
&= \frac{1}{P(X{=}x)} \int 1_{X=x} \Big(E^{X=x}(Y|Z) \cdot \frac{P(X{=}x)}{\pi_x}\Big)\, dP && [\textit{SN-Cor. } 9.5] \\
&= \int E^{X=x}(Y|Z) \cdot \frac{1_{X=x}}{\pi_x}\, dP \\
&= \int E^{X=x}(Y|Z{=}z) \cdot \frac{1_{\{x\}}(x')}{P(X{=}x|Z{=}z)}\, P_{X,Z}\big(d(x',z)\big) && [\textit{RS-}(3.12)] \\
&= \int\!\!\int E^{X=x}(Y|Z{=}z) \cdot \frac{1_{\{x\}}(x')}{P(X{=}x|Z{=}z)}\, P_{X|Z=z}(dx')\, P_Z(dz) && [\textit{SN-}(17.43)] \\
&= \int E^{X=x}(Y|Z{=}z) \cdot \frac{1}{P(X{=}x|Z{=}z)} \int 1_{\{x\}}(x') P_{X|Z=z}(dx')\, P_Z(dz) \\
&= \int E^{X=x}(Y|Z{=}z) \cdot \frac{1}{P(X{=}x|Z{=}z)}\, P_{X|Z=z}(\{x\})\, P_Z(dz) && [\textit{SN-}(17.30)] \\
&= \int E^{X=x}(Y|Z{=}z) \cdot \frac{1}{P(X{=}x|Z{=}z)}\, P(X{=}x|Z{=}z)\, P_Z(dz) && [\textit{SN-}(17.14)] \\
&= \int E^{X=x}(Y|Z{=}z)\, P_Z(dz) \\
&= E\big(E^{X=x}(Y|Z)\big) && [\textit{RS-}(3.15)] \\
&= E\big(E(\tau_x|Z)\big) && [E^{X=x}(Y|Z) \vdash D_X, \textit{Def. } 6.13] \\
&= E(\tau_x). && [\textit{RS-Box } 4.1 \text{ (iv)}]
\end{aligned}
$$

Remark 11.55 [Identification of the Causal Average Total Effect] Under the assumptions of Theorem 11.54, the expectation $E(\tau_x)$ of the true outcome variable $\tau_x = E^{X=x}(Y|D_X)$ can be identified by the conditional expectation $E^{X=x}(Y_w)$ of the *weighted outcome variable* Y_w [see Eq. (11.78)]. Furthermore, according to Equation (11.77), we can identify the expectation $E(\tau_x)$ of the true outcome variable τ_x also by taking the $(X{=}x)$-conditional expectation of the conditional expectation $E^{X=x}(Y|Z)$ of the outcomes on the covariate Z *weighted* by the function $P(X=x)/\pi_x$. Because these equations hold for each value of X, this yields

$$\begin{aligned}
ATE_{xx'} &= E(\tau_x) - E(\tau_{x'}) \\
&= E^{X=x}\left(E^{X=x}(Y|Z) \cdot \frac{P(X=x)}{\pi_x}\right) - E^{X=x'}\left(E^{X=x'}(Y|Z) \cdot \frac{P(X=x')}{\pi_{x'}}\right) \qquad (11.79)\\
&= E^{X=x}(Y_w) - E^{X=x'}(Y_w), \qquad (11.80)
\end{aligned}$$

provided that $E^{X=x}(Y|Z)$ and $E^{X=x'}(Y|Z)$ are unbiased. $\triangleleft$

Remark 11.56 [Prerequisites for Using the Theorem] A prerequisite for using Theorem 11.54 is to know the $(Z=z)$-conditional propensity scores $P(X=x|Z=z)$. In applications, these functions are known if the experimenter determines these conditional treatment probabilities, for example, as a function of the covariate Z, or as a function of the observational-unit variable U. In quasi-experiments, in which by definition the propensities are *not* under the experimenter's control, the propensities may also be estimated. However, in this case we need (a) a correct choice of the relevant (possibly multivariate) covariate Z such that the conditional expectation $E(Y|X,Z)$ is unbiased, and (b) a correct estimate of the propensities $P(X=x|Z)$, $x \in X(\Omega)$. $\triangleleft$

Remark 11.57 [A Caveat Concerning Efficiency] Using the weighting procedure neither involves the conditional expectation $E(Y|X,Z)$ nor the conditional expectation $E(Y|X,\pi)$ or the expectations of these conditional expectations. Therefore this procedure may be an option whenever estimating these conditional expectations or their expectations is difficult. What is needed, instead, are the propensities $P(X=x|Z)$. It should be emphasized, however, that Theorem 11.54 does not say anything about efficiency of estimators. Hence, in empirical applications one should also compare the standard errors of the average total effects estimated via Theorem 11.54 to the average total effects estimated via adjusting procedures based on Theorems 6.34 and 11.20. $\triangleleft$

Example 11.58 [Z-Conditional Independence of X and ? — continued] We use Table 7.5 to illustrate weighting the outcome variable using the Z-conditional propensities $P(X=x|Z)$, where $Z := (Z_1, Z_2)$ with $Z_1 :=$ *sex* and $Z_2 :=$ *college*. Remember,

$$E^{X=x}(Y_w) = \sum_z E^{X=x}(Y_w|Z=z) \cdot P^{X=x}(Z=z), \qquad (11.81)$$

is always true if $E^{X=x}(Y_w)$ exists and Z is discrete with $P(X=x, Z=z) > 0$ for all values $z \in \Omega'_Z$ [see RS-Box 3.2 (ii)]. Inserting Equation (11.76) yields

$$\begin{aligned}
E^{X=x}(Y_w) &= \sum_z E^{X=x}(Y_w|Z=z) \cdot P^{X=x}(Z=z)\\
&= \sum_z E^{X=x}(Y|Z=z) \cdot \frac{P(X=x)}{P(X=x|Z=z)} \cdot P^{X=x}(Z=z)
\end{aligned}$$

(see also Exercise 11-8). All terms occurring in this equation can be computed from the parameters displayed in Tables 7.5 and 11.2.

For *treatment* 0 we obtain

$$
\begin{aligned}
E^{X=0}(Y_w) &= \sum_z E^{X=0}(Y\,|\,Z{=}z)\cdot\frac{P(X{=}0)}{P(X{=}0\,|\,Z{=}z)}\cdot P^{X=0}(Z{=}z)\\
&= 72\cdot\frac{1/2}{2/8}\cdot 1/6+97.50\cdot\frac{1/2}{3/8}\cdot 1/4+111\cdot\frac{1/2}{6/8}\cdot 1/2 \;=\; 93.50,
\end{aligned}
$$

and for *treatment* 1,

$$
\begin{aligned}
E^{X=1}(Y_w) &= \sum_z E^{X=1}(Y\,|\,Z{=}z)\cdot\frac{P(X{=}1)}{P(X{=}1\,|\,Z{=}z)}\cdot P^{X=1}(Z{=}z)\\
&= 83\cdot\frac{1/2}{6/8}\cdot 1/2+102.50\cdot\frac{1/2}{5/8}\cdot 5/12+122\cdot\frac{1/2}{2/8}\cdot 1/6 \;=\; 102.50.
\end{aligned}
$$

Hence, according to Equation (11.77), the causal average total effect is

$$
ATE_{10} \;=\; E^{X=1}(Y_w)-E^{X=0}(Y_w) \;=\; 102.50-93.50 \;=\; 9.
$$

This result is exactly what we obtained before using the true propensity (see Example 11.25). ◁

11.5 Summary and Conclusions

In this chapter, we treated several ways to use conditional treatment probabilities in the identification of causal conditional and average total effects. A first possibility is to use the true propensities $\varphi_x = P(X{=}x\,|\,D_X)$, $x = 1,\ldots,J$, as covariates. Conditioning on $\varphi = (\varphi_1,\ldots,\varphi_J)$, the conditional prima facie effect function $PFE_{\varphi;xx'}$ and the conditional expectation $E(Y\,|\,X,\varphi)$ are always unbiased. The limitation of this procedure is that the true propensities have to be known. This means that this procedure can only be applied if the true propensities are fixed by the experimenter or if the (possibly multivariate) covariate Z is selected such that $\varphi_x = P(X{=}x\,|\,D_X) = P(X{=}x\,|\,Z)$.

A second possibility is to choose the Z-conditional propensities, that is, the conditional probabilities $\pi_x = P(X{=}x\,|\,Z)$, $x = 1,\ldots,J$, as new covariates, and assume that the original covariate Z is selected such that $\tau = (\tau_0,\tau_1,\ldots,\tau_J)$ and X are conditionally mean-independent given Z. According to Theorem 11.48, this can be achieved (a) by selecting Z such that X and a global confounder D_X are Z-conditionally independent or, (b) by selecting Z such that Y is Z-conditionally independent from a global confounder D_X. In both cases we can compute the conditional total effect function $CTE_{\pi;xx'}$ and its expectation, which is the causal average total effect $ATE_{xx'}$. Brookhart et al. (2006) and Wyss, Girman, LoCasale, Brookhart, and Stürmer (2013) compare the statistical implications of procedures (a) and (b) to each other.

The last procedure treated in this chapter consists of weighting the outcome variable with weights using the Z-conditional propensities $P(X{=}x\,|\,Z)$. Using this procedure either presumes that these probabilities are known — this defines the true experiment — or that they are estimated in quasi-experiments. In a quasi-experiment, we have to assume that Z is selected such that the conditional expectations $E^{X=x}(Y\,|\,Z)$, $x \in X(\Omega)$, are unbiased.

Table 11.3. Implications among some causality conditions involving propensities

	$\tau_x \perp\!\!\!\perp 1_{X=x} \,\vert\, \pi_x$	$\tau \perp\!\!\!\perp X \,\vert\, \pi$	$\tau_x \perp\!\!\!\perp 1_{X=x} \,\vert\, Z$	$\tau_x \perp\!\!\!\perp X \,\vert\, Z$	$\tau \perp\!\!\!\perp 1_{X=x} \,\vert\, Z$	$\tau \perp\!\!\!\perp X \,\vert\, Z$	$Y \vDash D_X \vert (X,Z)$
$\tau \perp\!\!\!\perp X \,\vert\, \pi$	(a)-(d)						
$\tau_x \perp\!\!\!\perp 1_{X=x} \,\vert\, Z$	(a)-(d)						
$\tau_x \perp\!\!\!\perp X \,\vert\, Z$	(a)-(d)		(a)-(d)				
$\tau \perp\!\!\!\perp 1_{X=x} \,\vert\, Z$	(a)-(d)		(a)-(d)				
$\tau \perp\!\!\!\perp X \,\vert\, Z$	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)		
$Y \vDash D_X \vert (X,Z)$	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)	
$Y \perp\!\!\!\perp D_X \vert (X,Z)$	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)	(a)-(d)

Note. We define $\pi_x = P(X{=}x \,\vert\, Z)$, $x \in X(\Omega)$, and $\pi = (\pi_1, \dots, \pi_J)$. An entry (a)-(d) means that the condition in the row implies the condition in the column, provided that the Assumptions 11.1(a) to (d) hold. Furthermore, we assume that, for all $x \in X(\Omega)$, the true outcome variables τ_x are P-unique. The symbols involving $\vDash$ and $\perp\!\!\!\perp$ have been explained in Box 7.2. The implications of $\tau_x \perp\!\!\!\perp 1_{X=x} \,\vert\, \pi_x$ and of $\tau \perp\!\!\!\perp X \,\vert\, \pi$ on other causality conditions have been summarized in Table 7.2.

11.6 Exercises

▷ **Exercise 11-1** Is it possible in applications to know the true treatment probability $P(X{=}x \,\vert\, D_X)$ and why is it useful to know it?

▷ **Exercise 11-2** Why does $\tau \perp\!\!\!\perp X \,\vert\, \varphi$ imply $\tau_x \perp\!\!\!\perp X \,\vert\, \varphi$ for all $x \in \{0, 1, \dots, J\}$?

▷ **Exercise 11-3** Show that Equation (11.57) holds.

▷ **Exercise 11-4** Compute the conditional expectation values displayed in Table 11.1 using Equation (11.71).

▷ **Exercise 11-5** Compute the coefficients of Equation (11.70) from the conditional expectations presented in Table 11.1.

▷ **Exercise 11-6** Consider Table 7.5 and show that

$$E(1_{\varphi=5/8} \,\vert\, 1_{Z_1=f}) = \frac{3}{4} - \frac{3}{4} \cdot 1_{Z_1=f}$$

holds for $\varphi = P(X{=}1 \,\vert\, U)$.

▷ **Exercise 11-7** Show that the strong ignorability condition $\tau \perp\!\!\!\perp X \,\vert\, Z$ holds in the example presented in Table 7.5.

▷ **Exercise 11-8** Show that $E(Y_w \,\vert\, X{=}x) = E(\tau_x)$ for Y_w defined in Equation (11.73).

▷ **Exercise 11-9** Use Equation (11.19) to compute the average total effect in the total population from the average total effects in the two subpopulations for the random experiment presented in Table 7.5.

▷ **Exercise 11-10** Why is the weight variable occurring in Equation (11.73) a function of X and Z?

▷ **Exercise 11-11** Prove the implications among the causality conditions shown in Table 11.3.

Solutions

▷ **Solution 11-1** In applications, it *is* possible to know the true treatment probability function $P(X{=}x\,|\,D_X)$. For instance, if the experimenter fixes the individual treatment probabilities $P(X{=}x\,|\,U{=}u)$ himself and makes sure that no other potential confounder determines the treatment probability. These probabilities can be fixed following ethical or economic considerations. They do not have to be identical for all individuals in order to allow for valid causal inference. If the individual treatment probabilities are known, we can use the functions $\varphi_x := P(X{=}x\,|\,D_X) = P(X{=}x\,|\,U)$ as covariates in the adjustment theorems. The conditional expectations $E(Y\,|\,X,\varphi)$, with $\varphi := (\varphi_1,\ldots,\varphi_J)$, are always unbiased. Alternatively, we can compute the weighted outcome variable [see Eq. (11.73)], where D_X takes the role of the covariate Z. The expectation of this weighted outcome variable in treatment x is equal to the expectation $E(\tau_x)$ of the true outcome variable under treatment x. Furthermore, the difference between the expectation of the weighted outcomes in treatment x and the expectation of the weighted outcomes in control 0 is the average causal effect of treatment x compared to the control (see Theorem 11.54).

▷ **Solution 11-2** Because $\sigma(\tau_x) \subset \sigma(\tau)$, for all $x \in \{0,1,\ldots,J\}$, we can conclude from SN-Box 16.3 (vi) that $\tau \perp\!\!\!\perp X\,|\,\varphi$ implies $\tau_x \perp\!\!\!\perp X\,|\,\varphi$ for all $x \in \{0,1,\ldots,J\}$.

▷ **Solution 11-3** The indicator 1_Ω is (identical to) the constant 1 and $\pi_x = P(X{=}x\,|\,Z) \underset{P}{=} E(1_{X=x}\,|Z)$, $x \in \{0,1,\ldots,J\}$ [see SN-Eq. (10.4)]. Hence,

$$
\begin{aligned}
1 &= 1_\Omega \\
&\underset{P}{=} E(1_\Omega\,|\,Z) && \text{[SN-Box 10.2 (i)]} \\
&\underset{P}{=} E(1_{X=0} + 1_{X=1} + \ldots + 1_{X=J}\,|\,Z) && [1_\Omega = 1_{X=0} + 1_{X=1} + \ldots + 1_{X=J}] \\
&\underset{P}{=} E(1_{X=0}\,|\,Z) + E(1_{X=1}\,|\,Z) + \ldots + E(1_{X=J}\,|\,Z). && \text{[SN-Box 10.2 (xvi)]}
\end{aligned}
$$

Inserting the shortcut π_x, $x \in \{0,1,\ldots,J\}$, and subtracting $\pi_1 + \ldots + \pi_J$ on both sides yields Equation (11.57).

▷ **Solution 11-4** Fehlt noch

$$E^{X=x}(Y\,|\,\varphi{=}p) = \sum_u E^{X=x}(Y\,|\,\varphi{=}p, U{=}u) \cdot P^{X=x}(U{=}u\,|\,\varphi{=}p) \tag{11.82}$$

$$= \sum_u E^{X=x}(Y\,|\,U{=}u) \cdot P^{X=x}(U{=}u\,|\,\varphi{=}p) \tag{11.83}$$

▷ **Solution 11-5** According to Equations (11.68) and (11.70), the conditional expectation $E^{X=1}(Y\,|\,\varphi)$ is

$$E^{X=1}(Y\,|\,\varphi) = \gamma_0^{(1)} + \gamma_1^{(1)} \cdot 1_{\varphi=5/8} + \gamma_2^{(1)} \cdot 1_{\varphi=2/8}. \tag{11.84}$$

For $\varphi = 7/8$, this equation and the conditional expectations in Table 11.1 yield

$$E^{X=1}(Y\,|\,\varphi{=}7/8) = \gamma_0^{(1)} = 83.00,$$

and for $\varphi = 5/8$ they yield

$$E^{X=1}(Y\,|\,\varphi{=}5/8) = \gamma_0^{(1)} + \gamma_1^{(1)} = 96.00.$$

Hence $\gamma_1^{(1)} = 96 - 83 = 13$. Finally, for $\varphi = 2/8$, Equation (11.84) and Table 11.1 yield

$$E^{X=1}(Y\,|\,\varphi{=}2/8) = \gamma_0^{(1)} + \gamma_2^{(1)} = 122.00.$$

Hence $\gamma_2^{(1)} = 122 - 83 = 39$.

▷ **Solution 11-6** Fehlt noch

▷ **Solution 11-7** If X is binary and each of the values 0 and 1 has a positive probability, then $\tau \perp\!\!\!\perp X \mid Z$ is equivalent to

$$P(X{=}1 \mid Z, \tau) = P(X{=}1 \mid Z) \tag{11.85}$$

(see SN-Rem. 16.25 and Eq. (11.57)). Looking at Table 7.5 shows that, given the value (m, no) of Z, the bivariate true outcome variable $\tau = (\tau_0, \tau_1)$ is constant, taking on the value $(72, 83)$. This implies

$$P\big(X{=}1 \,\big|\, Z{=}(m, no), \tau{=}(72,83)\big) = P\big(X{=}1 \,\big|\, Z{=}(m, no)\big).$$

Furthermore, given the value (m, yes) of Z,

$$P\big(X{=}1 \,\big|\, Z{=}(m, yes), \tau{=}(95,100)\big) = P\big(X{=}1 \,\big|\, Z{=}(m, yes), \tau{=}(100,105)\big) = P\big(X{=}1 \,\big|\, Z{=}(m, yes)\big).$$

Finally, given the value (f, yes) of Z,

$$P\big(X{=}1 \,\big|\, Z{=}(f, yes), \tau{=}(106,114)\big) = P\big(X{=}1 \,\big|\, Z{=}(f, yes), \tau{=}(116,130)\big) = P\big(X{=}1 \,\big|\, Z{=}(f, yes)\big).$$

This shows that Equation (11.85) holds.

▷ **Solution 11-8**

$$\begin{aligned}
E(Y_w \mid X{=}x) &= E^{X=x}(Y_w) && [(3.24)]\\
&= \sum_z E^{X=x}(Y_w \mid Z{=}z) \cdot P^{X=x}(Z{=}z) && \text{[SN-Eq. (9.23)]}\\
&= \sum_z E(Y \mid X{=}x, Z{=}z) \cdot \frac{P(X{=}x)}{P(X{=}x \mid Z{=}z)} \cdot P^{X=x}(Z{=}z) && [(11.76)]\\
&= \sum_z E(Y \mid X{=}x, Z{=}z) \cdot \frac{P(X{=}x) \cdot P(Z{=}z)}{P(X{=}x, Z{=}z)} \cdot \frac{P(X{=}x, Z{=}z)}{P(X{=}x)} && \text{[SN-Eqs. (9.13) to (9.15)]}\\
&= \sum_z E(Y \mid X{=}x, Z{=}z) \cdot P(Z{=}z)\\
&= \sum_z E(\tau_x \mid Z{=}z) \cdot P(Z{=}z) && \text{[Def. 6.13 (ii)]}\\
&= E(\tau_x). && \text{[SN-Eq. (9.23)]}
\end{aligned}$$

▷ **Solution 11-9** According to Equation (11.19) and because the two-dimensional covariate Z is discrete, we can use

$$ATE_{10} = E\big(PFE_{\varphi;10}(\varphi)\big) = \sum_p PFE_{\varphi;10}(p) \cdot P(\varphi{=}p),$$

where the summation is over all values p of φ. For the example presented in Table 7.5 this yields:

$$\begin{aligned}
ATE_{10} &= PFE_{\varphi;10}(7/8) \cdot 1/6 + PFE_{\varphi;10}(5/8) \cdot 3/6 + PFE_{\varphi;10}(2/8) \cdot 2/6\\
&= 11.00 \cdot 1/6 + 7.00 \cdot 3/6 + 11.00 \cdot 2/6 = 9.
\end{aligned}$$

▷ **Solution 11-10** The weight variable in Equation (11.73) is a function of X and Z because it depends on X via the indicator $1_{X=x}$ and on the covariate Z via the Z-conditional propensity $P(X{=}x \mid Z)$.

▷ **Solution 11-11** The sequence of proofs of the propositions of Table 11.3 is row wise.

(a) $\tau \perp\!\!\!\perp X \mid \pi \;\Rightarrow\; \tau_x \perp\!\!\!\perp 1_{X=x} \mid \pi_x$. According to Lemma 6.6, we have to show that $\tau \perp\!\!\!\perp X \mid \pi$ implies

$$P(X{=}x \mid \pi_x, \tau_x) \underset{P}{=} P(X{=}x \mid \pi_x), \quad \forall\, x = 0, 1, \ldots, J.$$

Hence,

$$P(X{=}x \mid \pi_x, \tau_x)$$

$$\underset{P}{=} E(1_{X=x} \mid \pi_x, \tau_x) \qquad \text{[SN-Eq. (10.4)]}$$

$$\underset{P}{=} E\big(E(1_{X=x} \mid \pi, \tau) \,\big|\, \pi_x, \tau_x\big) \qquad \text{[SN-Box 10.2 (v)]}$$

$$\underset{P}{=} E\big(E(1_{X=x} \mid \pi) \,\big|\, \pi_x, \tau_x\big) \qquad [\tau \perp\!\!\!\perp X \mid \pi, \text{SN-Box 16.3 (vi), Lem. 6.6, SN-Eq. (10.4)}]$$

$$\underset{P}{=} E\big(E(1_{X=x} \mid Z) \,\big|\, \pi_x, \tau_x\big) \qquad \text{[Th. 11.48 ??, SN-Eq. (10.4)]}$$

$$\underset{P}{=} E\big(E(1_{X=x} \mid \pi_x) \,\big|\, \pi_x, \tau_x\big) \qquad \text{[Th. 11.44 ??, SN-Eq. (10.4)]}$$

$$\underset{P}{=} E(1_{X=x} \mid \pi_x) \qquad \text{[SN-Box 10.2 (vii)]}$$

$$\underset{P}{=} P(X{=}x \mid \pi_x). \qquad \text{[SN-Eq. (10.4)]}$$

(b) $\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid \pi_x$. This is Theorem 11.44 **??** (see Rem. 11.45).

(c) $\tau_x \perp\!\!\!\perp X \mid Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid \pi_x$. According to SN-Box 16.3 (vi), $\tau_x \perp\!\!\!\perp X \mid Z$ implies $\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$, because $\sigma(1_{X=x}) \subset \sigma(X)$. Now the proposition follows from (b).

(d) $\tau_x \perp\!\!\!\perp X \mid Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$. This proposition immediately follows from SN-Box 16.3 (vi), because $\sigma(1_{X=x}) \subset \sigma(X)$.

(e) $\tau \perp\!\!\!\perp 1_{X=x} \mid Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid \pi_x$. According to SN-Box 16.3 (vi), $\tau \perp\!\!\!\perp 1_{X=x} \mid Z$ implies $\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$, because $\sigma(\tau_x) \subset \sigma(\tau)$. Now the proposition follows from (b).

(f) $\tau \perp\!\!\!\perp 1_{X=x} \mid Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$. This proposition immediately follows from SN-Box 16.3 (vi), because $\sigma(\tau_x) \subset \sigma(\tau)$.

(g) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid \pi_x$. According to SN-Box 16.3 (vi), $\tau \perp\!\!\!\perp X \mid Z$ implies $\tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$, because $\sigma(\tau_x) \subset \sigma(\tau)$ and $\sigma(1_{X=x}) \subset \sigma(X)$. Now the proposition follows from (b).

(h) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow \tau \perp\!\!\!\perp X \mid \pi$. This is Theorem 11.48 (i).

(i) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$. This proposition immediately follows from SN-Box 16.3 (vi), because $\sigma(\tau_x) \subset \sigma(\tau)$ and $\sigma(1_{X=x}) \subset \sigma(X)$.

(j) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow \tau_x \perp\!\!\!\perp X \mid Z$. This proposition immediately follows from SN-Box 16.3 (vi), because $\sigma(\tau_x) \subset \sigma(\tau)$.

(k) $\tau \perp\!\!\!\perp X \mid Z \Rightarrow \tau \perp\!\!\!\perp 1_{X=x} \mid Z$. This proposition immediately follows from SN-Box 16.3 (vi), because $\sigma(1_{X=x}) \subset \sigma(X)$.

(l) $Y \vDash D_X|(X,Z) \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid \pi_x$. According to Theorem 11.48 (ii), $Y \vDash D_X|(X,Z)$ implies $\tau \perp\!\!\!\perp X \mid \pi$. Now the proposition follows from (a).

(m) $Y \vDash D_X|(X,Z) \Rightarrow \tau \perp\!\!\!\perp X \mid \pi$. This is Theorem 11.48 (ii).

(n) $Y \vDash D_X|(X,Z) \Rightarrow \tau_x \perp\!\!\!\perp 1_{X=x} \mid Z$. According to Theorem 9.38, $Y \vdash D_X|(X,Z)$ implies $\tau \perp\!\!\!\perp X \mid Z$. Now, the proposition immediately follows from SN-Box 16.3 (vi), because $\sigma(\tau_x) \subset \sigma(\tau)$ and $\sigma(1_{X=x}) \subset \sigma(X)$.

(o) $Y \vDash D_X|(X,Z) \Rightarrow \tau_x \perp\!\!\!\perp X \mid Z$. According to Theorem 9.38, $Y \vdash D_X|(X,Z)$ implies $\tau \perp\!\!\!\perp X \mid Z$. Now, the proposition immediately follows from SN-Box 16.3 (vi), because $\sigma(\tau_x) \subset \sigma(\tau)$.

(p) $Y \vDash D_X|(X,Z) \Rightarrow \tau \perp\!\!\!\perp 1_{X=x} \mid Z$. According to Theorem 9.38, $Y \vdash D_X|(X,Z)$ implies $\tau \perp\!\!\!\perp X \mid Z$. Now, the proposition immediately follows from SN-Box 16.3 (vi), because. and $\sigma(1_{X=x}) \subset \sigma(X)$.

(q) $Y \vDash D_X|(X,Z) \Rightarrow \tau \perp\!\!\!\perp X \mid Z$. This is Theorem 9.38.

(r) All seven implications in the last line of Table 11.3. According to SN-Theorem 16.34 and SN-Remark 16.35, $Y \perp\!\!\!\perp D_X|(X,Z) \Rightarrow Y \vDash D_X|(X,Z)$. Therefore all other implications in the last line of Table 11.3 follow from (l) to (q).

Chapter 12
Methodological Conclusions

In this chapter, we will summarize and discuss the most important issues treated in this book. The specific goals are:

(a) to summarize the theory,
(b) to discuss the most important consequences of the theory of causal conditional and average total effects for the design and analysis of experiments and quasi-experiments, and
(c) to reflect on other issues in causal inference that are not treated in this volume.

12.1 Summary of the Theory

In the introduction, we motivated the theory presented in this book by a version of Simpson's paradox showing that comparing conditional probabilities can lead to contradictory and seriously misleading causal inferences. There, we also showed in a second example that this also applies to comparing conditional expectation values.

The Problem

The difference between the conditional expectation values of the outcome variable Y given values x and x' of a discrete putative cause variable X is called a *prima facie effect*. This is not necessarily identical to the causal total effect on Y comparing x to x'. In fact, we presented examples, in which there are *positive* total effects of comparing treatment 1 to treatment 0 *for each and every observational unit* (person) and still a *negative* prima facie effect $E(Y|X{=}1) - E(Y|X{=}0)$ (see e. g., Tables 1.1 to 1.4 or Table 6.2).

The Kind of Empirical Phenomenon the Theory is About: Single-Unit Trials

We continued preparing the stage for the theory of causal conditional and average total effects by describing the kind of empirical phenomena the theory is dealing with. These empirical phenomena are called *single-unit trials*. One of the simplest examples of such a single-unit trial consists of drawing a single observational unit (e. g., a person) from the set of units, exposing or observing exposure to one of at least two treatment conditions and, finally, observing the outcome variable Y, the distribution of which might be affected by the treatment variable X. In terms of statistics, one might also say that the theory deals with a sample of size $n = 1$. It neither deals with a sample of size $n > 1$ nor with a concrete data sample or data set that we use in data analysis. Note that this kind of empirical phenomenon is not only of interest to the scientist but also to the practitioner who has to choose a treatment for the person seeking his help.

The Mathematical Framework of the Theory: A Regular Probabilistic Causality Setup

The description of a single-unit trial already shows that there is a time order between drawing an observational unit, its treatment, and the outcome. This is one of the properties distinguishing *causal* dependencies from *ordinary* probabilistic dependencies. This time order is represented by a *filtration* $(\mathscr{F}_t)_{t \in T}$, which is not only used to define an *outcome variable* Y but also plays an important role in specifying the *cause* σ*-algebra* $\mathscr{C}$ and the *confounder* σ*-algebra* $\mathscr{D}_{\mathscr{C}}$ used to define the *putative cause variable* X and the *global potential confounder* D_X *of* X, respectively. The confounder σ-algebra $\mathscr{D}_{\mathscr{C}}$ is also used to define a *potential confounder of* X, which is a random variable on $(\Omega, \mathscr{A}, P)$ that is measurable with respect to $\mathscr{D}_{\mathscr{C}}$. Such a potential confounder of X is also called a *covariate of* X, often denoted by Z, in particular if used in a conditional expectation such as $E(Y|X,Z)$.

The terms mentioned above are the additional components that distinguish a *regular probabilistic causality setup* $\big((\Omega, \mathscr{A}, P), (\mathscr{F}_t)_{t \in T}, \mathscr{C}, \mathscr{D}_{\mathscr{C}}, X, Y\big)$ from a *probability space* $(\Omega, \mathscr{A}, P)$. Within the latter, we can define random variables and deal with ordinary probabilistic dependencies among random variables and among events. The additional structural components of a regular probabilistic causality setup, however, do not yet distinguish between causal and ordinary probabilistic dependencies. They just provide the mathematical framework in which we can define *causality conditions*, and it is them that are crucial for the distinction between ordinary and causal probabilistic dependencies and effects.

Causality Conditions

The logically weakest of these causality conditions are the *unbiasedness conditions*. Unbiasedness of the conditional expectation value $E^{X=x}(Y) = E(Y|X{=}x)$ is defined by postulating that the true outcome variable $\tau_x = E^{X=x}(Y|D_X)$ is P-unique (implying that its expectation is identical for all its versions) and that $E^{X=x}(Y)$ is identical to the expectation of $\tau_x = E^{X=x}(Y|D_X)$. The conditional expectation $E(Y|X)$ is then called unbiased if all its values $E^{X=x}(Y) = E(Y|X{=}x)$ are unbiased. Similarly, unbiasedness of the conditional expectation $E^{X=x}(Y|Z)$ is defined by assuming P-uniqueness of τ_x and requiring that $E^{X=x}(Y|Z)$ is identical almost surely to the Z-conditional expectation of τ_x. The conditional expectation $E(Y|X,Z)$ is called unbiased if each $E^{X=x}(Y|Z)$ is unbiased (see ch. 6).

The term $E^{X=x}(Y)$ used above is the expectation of Y with respect to the probability measure $P^{X=x}$. Presuming $P(X{=}x) > 0$, this measure is defined by $P^{X=x}(A) = P(A|X{=}x)$, for all $A \in \mathscr{A}$. Similarly, $E^{X=x}(Y|D_X)$ and $E^{X=x}(Y|Z)$ are conditional expectations with respect to $P^{X=x}$ (see RS-ch. 5).

The assumption of P*-uniqueness of* τ_x mentioned above means that whenever τ_x and τ_x^* are two versions of the D_X-conditional expectation of Y with respect to $P^{X=x}$, then $\tau_x \underset{P}{=} \tau_x^*$. According to RS-Theorem 5.27, this assumption is equivalent to $P(X{=}x|D_X) \underset{P}{>} 0$, that is, to postulating that the D_X-conditional probability of the event $\{X{=}x\}$ is positive P-almost surely. Assuming that the true outcome variables τ_x and $\tau_{x'}$ are P-unique implies that the expectations of τ_x and $\tau_{x'}$ (and of their difference) is a uniquely defined number that does not change with a different choice of the versions of the true outcome variables $\tau_x = E^{X=x}(Y|D_X)$ or $\tau_{x'} = E^{X=x'}(Y|D_X)$ (see RS-sect. 5.2 for the mathematical details).

Unfortunately, the unbiasedness conditions are not empirically testable or falsifiable. This also applies to independence and Z-conditional independence of X and the true outcome variables τ_x, as well as to mean-independence and Z-conditional mean independence of the τ_x from the putative cause variable X. These conditions, summarily called the *Rosenbaum-Rubin conditions*, have been treated in chapter 7.

Another kind of causality conditions, which have been introduced in chapter 8, are the *Fisher conditions*. They comprise independence and Z-conditional independence of the putative cause variable X and the global potential confounder D_X of X. These conditions can be tested, and if necessary be falsified, by testing independence or Z-conditional independence of X and a potential confounder of X, say W. Even more important is that, if X is a treatment variable, then independence of X and D_X as well as Z-conditional independence of X and D_X can be created by the design technique of random assignment of an observational unit to one of the treatment conditions. Similarly, Z-conditional independence of X and D_X can be created by the design technique of Z-conditional random assignment of a unit to a treatment (see sect. 8.6 for details).

Still another type of causality condition are the *Suppes-Reichenbach conditions* that have been treated in chapter 9. They comprise X-conditional and (X,Z)-conditional independence of the outcome variable Y and the global potential confounder D_X of X. These conditions, too, can be tested, and if necessary be falsified, by testing X-conditional or (X,Z)-conditional independence of Y and a potential confounder W. The Suppes-Reichenbach conditions also comprise X-conditional and (X,Z)-conditional mean-independence of Y from D_X, which can empirically be checked and if necessary be falsified via testing X-conditional or (X,Z)-conditional mean-independence of Y from a potential confounder W of X. In contrast to the Fisher conditions, the Suppes-Reichenbach conditions cannot be created by design techniques, but (X,Z)-conditional mean-independence of Y from D_X can be used for covariate selection (see sect. 9.4).

The Fisher conditions as well as the Suppes-Reichenbach conditions imply *unconfoundedness*, the last kind of causality conditions treated in this volume (see ch. 10). Unconfoundedness of the conditional expectation value $E^{X=x}(Y) = E(Y|X=x)$ is defined by postulating P-uniqueness of τ_x and that $E^{X=x}(Y)$ is identical to the expectation of the conditional expectation $E^{X=x}(Y|W)$ whenever W is a potential confounder of X. Furthermore, the conditional expectation $E(Y|X)$ is called unconfounded if all its values $E^{X=x}(Y)$ are unconfounded. Similarly, unconfoundedness of the conditional expectation $E^{X=x}(Y|Z)$ is defined by requiring P-uniqueness of τ_x and that $E^{X=x}(Y|Z)$ is identical almost surely to the Z-conditional expectation of the conditional expectation $E^{X=x}(Y|Z,W)$ whenever W is a potential confounder of X. The conditional expectation $E(Y|X,Z)$ is called unconfounded if each $E^{X=x}(Y|Z)$ is unconfounded. Unconfoundedness of the conditional expectation $E(Y|X,Z)$, too, can be used for covariate selection. To my knowledge, it is the weakest causality condition implying unbiasedness of $E(Y|X,Z)$ that it is still empirically testable (see sect. 10.7).

Causal Total Effects

The definition of causal total effects on an outcome variable Y comparing x to x', two values of a putative cause variable X, is based on the concept of a true outcome variable

$\tau_x = E^{X=x}(Y|D_X)$, a conditional expectation of Y given the global potential confounder D_X of X with respect to the probability measure $P^{X=x}$. In this conditional expectation we condition on the value x of X and the global potential confounder D_X of X, and with it, on all potential confounders of X. Hence, the *true total effect variable* $\delta_{xx'} = \tau_x - \tau_{x'}$ cannot be biased and we can define

$$ATE_{xx'} \;=\; E(\tau_x - \tau_{x'}) \;=\; E\big(E^{X=x}(Y|D_X) - E^{X=x'}(Y|D_X)\big),$$

the *causal average total effect on Y comparing x to x′*, provided that τ_x and $\tau_{x'}$ are P-unique, or equivalently, provided that $P(X{=}x|D_X)$ and $P(X{=}x'|D_X)$ are positive P-almost surely.

Similarly, a *causal Z-conditional total effect variable* is defined by the difference

$$CTE_{Z;xx'}(Z) \underset{P}{=} E(\tau_x|Z) - E(\tau_{x'}|Z)$$

between the Z-conditional expectations of the true outcome variables τ_x and $\tau_{x'}$, again presuming P-uniqueness of τ_x and $\tau_{x'}$. Here, Z is a random variable on $(\Omega, \mathscr{A}, P)$, typically a covariate of X, but it can also be X itself. Such a causal Z-conditional total effect variable may provide more detailed information than the causal average total effect $ATE_{xx'}$.

12.2 Design and Analysis of Experiments and Quasi-Experiments

True Experiments

Some of the conditions implying unbiasedness of the conditional expectation $E(Y|X)$ or unbiasedness of the conditional expectation $E(Y|X,Z)$ can be created the researcher in *true experiments*, that is, in empirical studies in which

(a) there are at least two treatment conditions,
(b) there is random assignment, or at least conditional random assignment of the observational unit to one of the treatment conditions, and
(c) assignment also means actual exposure to treatment.

Causal inference is most safe in true experiments. There, unbiasedness of the conditional expectations $E(Y|X)$ and $E(Y|X,Z)$ can be created by design, namely by some form of random assignment of the unit to one of the treatment conditions. Note, however, that such a design can be jeopardized by systematic attrition of observational units (see, e. g., Shadish et al., 2002; West & Sagarin, 2000). In this case, the study is not a randomized experiment any more, and the Fisher conditions of independence and Z-conditional independence of the treatment variable X and the global potential confounder D_X of X (see Box 8.1) may be violated.

Quasi-Experiments

The assumptions under which identification of causal effects is possible may also hold in a *quasi-experiment* if an appropriate covariate Z of X is selected. By definition, in a quasi-experiment there are at least two treatment conditions, but requirement (b) or (c) defining a true experiment does not hold. Such a uni- or multivariate covariate Z of the putative cause variable X can be manifest or latent. It can also be the Z-conditional propensity $P(X{=}x|Z)$ because, if Z is a covariate of X, then $P(X{=}x|Z)$ is a covariate of X as well.

Causality Tests and Covariate Selection

Causality tests and covariate selection are useful for two purposes. The first is to establish causality, that is, to select a (possibly multivariate) covariate Z such that $E(Y|X,Z)$ is unbiased. The second is to select a covariate Z in order to learn about conditional total effects that are as close as possible to the individual total effects. Also in the true experiment, covariate selection is meaningful. However, here the purpose is not to establish causality, but to learn about conditional total effects.

Unfortunately, the weakest causality condition, unbiasedness of the conditional expectation $E(Y|X,Z)$, is not empirically testable. What is testable, however, are some of the other causality conditions, all of which imply unbiasedness. Testing one of these sufficient conditions may be called a *causality test*. In Remarks 8.60 and 8.61 we described how covariate selection and causality tests can be conducted for the Fisher conditions. Similarly, in section 9.4 we described how covariate selection and causality tests can be conducted for the Suppes-Reichenbach conditions and in section 10.7, we described how to check the unconfoundedness conditions and how to use them for covariate selection.

In empirical causal research, covariate selection oftentimes is focussed on the Fisher condition of Z-conditional independence of X and the global potential confounder D_X of X. This is most obvious in those studies that use some kind of propensity score analysis. However, if we are interested in the causal mechanisms and the most-fine grained conditional total effects, then we might be better off with covariate selection based on the Suppes-Reichenbach condition choosing the covariate Z such that $E(Y|X,Z) \underset{P}{=} E(Y|X,D_X)$. Because

$$E(Y|X,D_X) \underset{P}{=} \sum_x E^{X=x}(Y|D_X),$$

the Z-conditional total effect variables $E^{X=x}(Y|Z) - E^{X=x'}(Y|Z)$ are identical P-almost surely to the true total effect variable $\tau_x - \tau_{x'} \underset{P}{=} E^{X=x}(Y|D_X) - E^{X=x'}(Y|D_X)$. Hence, if we can assume $E(Y|X,Z) \underset{P}{=} E(Y|X,D_X)$, then the Z-conditional total effect variable

$$E^{X=x}(Y|Z) - E^{X=x'}(Y|Z) \underset{P}{=} \tau_x - \tau_{x'}$$

yields the ultimate goal of in the search for causal total effects.

Two Threats to Causal Inference

In quasi-experiments, there are two kinds of threats to valid causal inferences. *First*, the covariate Z may not be chosen such that $E(Y|X,Z)$ is unbiased. Selecting the covariates determining the probability that a unit is exposed to treatment x or the covariates determining the outcome variable Y are of crucial importance in a quasi-experiment. Alternatively, one may select Z such that $E(Y|X,Z)$ is unconfounded.

If Z is not chosen such that $E(Y|X,Z)$ satisfies one of the causality conditions, then causal inference will fail. Choosing the appropriate covariate is not just statistical routine; it requires substantive knowledge about (a) the process of assigning units to treatment conditions and/or (b) the covariates determining the conditional expectation of the outcome variable Y.

Second, even if the covariate Z is correctly chosen such that the conditional expectation $E(Y|X,Z)$ is unbiased, the conditional expectation $E(Y|X,Z)$ may not be modeled correctly.[1] If, for example, one considers only a *linear* function of X and Z, but the (true) conditional expectation $E(Y|X,Z)$ is a *nonlinear* function of X and Z, then causal inferences may not be valid. If the propensity is used, an additional problem is the correct model for the propensity $P(X=x|Z)$, unless the propensity scores have been fixed by the experimenter and do not have to be estimated.

12.3 Other Issues of Causal Inference

In this book we studied the relations between the conditional expectations $E(Y|X)$ and $E(Y|X,Z)$, and various kinds of causal conditional and average total effects. However, as already mentioned in the preface, empirical causal research involves several inferences and substantive interpretations. Among these are:

(a) statistical inference, i. e., the inference from a sample to parameters characterizing the distributions of random variables,
(b) causal inference, i. e., the inference from parameters characterizing the distributions of random variables to causal effects,
(c) interpretation of the putative cause,
(d) interpretation of the outcome variable,
(e) interpretation of the random experiment (i. e., the empirical phenomenon) considered.

Most of these issues, which are dealt with in the theory of internal and external validity (see, e. g., Campbell & Stanley, 1963; Cook & Campbell, 1979; Shadish et al., 2002), have not been treated in this book. Instead we only focused on the second point. Briefly considering all these points may help to put the contribution of this book into perspective.

Statistical Inference

Statistical issues and models dealing with sampling processes, estimation and hypothesis testing, e. g., have not been treated. We did not present the sampling processes and models allowing to estimate and test the different kinds of causal effects. Under which assumptions is it possible, e. g., to use conditional expectation models with fixed regressors, stochastic regressors, or models with errors in the variables for estimating and testing causal effects (see, e. g., Fahrmeir, Hamerle, & Tutz, 1996)? Can we also use random coefficient models (see, e. g., Searle, Casella, & McCulloch, 1992, or models with latent variables (see, e. g., Bollen, 1995)? Which of the hypotheses relevant for the analysis of causal effects are linear, which need methods for testing nonlinear hypotheses? Can we use also nonparametric conditional expectation techniques (see, e. g., Takezawa, 2005)?

There are also specific statistical issues, e. g., dealing with *non-compliance* and *subject reassignment* (see, e. g., Angrist, Imbens, & Rubin, 1996; Barnard, Du, Hill, & Rubin, 1998; Barnard, Frangakis, Hill, & Rubin, 2003; Greenland, 2000; Jo, 2002c; J. M. Robins, 1998; West & Sagarin, 2000), *time-dependent covariates* (see, e. g., J. M. Robins, Hérnan, & Brumback, 2000), and *sensitivity analysis* (see, e. g., Greenland, 1996, 2001; Jo, 2002b; Rosenbaum, 2002b).

[1] Note again that — unlike the linear quasi-regression — the conditional expectation $E(Y|X,Z)$ is well-defined without any reference to a specific function of X and Z (see RS-sections 3.3 and 4.2).

Causal Inference

Causal inference deals with the inference from parameters characterizing the distributions of random variables (such as the expectations of an outcome variable in two treatment conditions) to causal effects. This is what the theory of causal effects in experiments and quasi-experiments presented in this book has been about. In section 12.1 we summarized the most important issues.

Interpretation of the Treatment Variable

The interpretation of the treatment variable is another point at which substantive conclusions can go wrong. Suppose X is a treatment variable. The substantive differences between the different values x of the treatment variable X are often not clear, because there usually are differences *in several aspects*. Suppose, for instance, we study the effects of 'taking drug A' as compared to 'not taking drug A' in a pharmacological experiment. The nominal value 'taking drug A' of the treatment variable consists of several components, such as 'going to a physician', 'taking a pill', and 'ingesting drug A', while the nominal value of X 'not taking drug A' may consist of 'not going to a physician', 'taking no pill', and 'not ingesting the drug'.

Which of these components is the causal agent? Is it the nominal value of X, i. e., is it 'taking drug A' or is it rather 'going to the physician', or 'taking a pill'? Of course, these questions can be answered only if studies are designed such that the interpretation of the values of the treatment variable is unambiguous, e. g., if 'taking drug A' is compared to 'taking a pill' (a placebo) that looks and tastes exactly the same without containing the drug to be investigated and that also involves 'going to a physician' and 'taking a pill', but not 'ingesting drug A'.

It should be emphasized that random assignment does not solve this interpretation problem. As argued above, a treatment usually differs from the control in several components and it is not clear which of these components is the causal agent. Similarly, random assignment neither solves the interpretation problem of the outcome variable, nor the interpretation of the random experiment considered (external validity) (see the following sections).

Interpretation of the Outcome Variable

Another point at which substantive conclusions about causal effects can go wrong is the interpretation of the outcome variable. Here, the problem is that oftentimes there is a difference between what is *intended to measure* and what is *actually measured*. A simple example is *annual income*. What is actually measured is either what subjects report as their annual income or what the tax authorities have in their records. However, the true annual income may differ from both actual measures, self report and tax authorities records. Similar problems occur if we measure *intelligence* by a specific *psychological test*, *academic achievement* by a *citation index*, *depression* or *life satisfaction* by *specific questionnaires*. In all these cases, we may erroneously believe that what is *actually measured* is what we *intended to measure* (for more details, see Shadish et al., 2002, ch. 3).

Interpretation of the Random Experiment Considered

The individual and average causal effects defined in this book refer to single-unit trials. A simple example is sampling an observational unit (e. g., a person) from a set of observational units, exposing it (or observing its exposure) to one of the treatment conditions, and finally, observing the outcome variable Y for the sampled unit (see ch. 2 for more details). In applications, such a single-unit trial always refers to a specific set of observational units, such as the set of all patients with a given diagnosis, the set of all undergraduate psychology students at your university in a given year, or the set of all consumers of a specific brand at a specific point of time. Furthermore, the single-unit trial also consists of a specific distribution of the observational-unit variable, that is, the probabilities $P(U{=}u)$ that the unit u (patient, student, consumer) is sampled. However, what is *not* explicitly represented in the mathematics treated in this book are the side conditions, i. e., the historical context, space, and time under which the single-unit trial is conducted. All these side conditions are implicitly referred to when we say that we are talking about a *specific* single-unit trial, a specific random experiment that is represented by the probability space $(\Omega, \mathscr{A}, P)$.

Consequently, all our propositions about causal effects refer only to the specific random experiment, i. e., the specific single-unit trial considered. What can go wrong in the interpretation of the random experiment considered, for instance, is that we may not be aware of some crucial side conditions without which the effect would not occur. For example, the treatment studied may exert its effects only if the subjects *know that they are a part of a scientific study*. In this case a generalization beyond the specific study would go wrong. Another point is generalizing to a population differing from the population of the units from which the actual sample was drawn. In psychology, for instance, many experiments are conducted with psychology students. The crucial issue is if and to what extent we may generalize to other populations of subjects. Problems related to the interpretation of the random experiment considered may be called problems of *external validity* (see, e. g., Shadish et al., 2002, ch. 5).

References

Abraham, W. T., & Russell, D. W. (2004). Missing data: A review of current methods and applications in epidemiological research. *Current Opinion in Psychiatry*, *17*, 315–321. pages 254, 291

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley. pages 12, 254

Aiken, L. S., & West, S. G. (1996). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, CA: Sage. pages 291

Aitkin, M. (1978). The analysis of unbalanced cross-classifications. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, *141*, 195–223. pages 13, 22

Allen, M. P. (1997). *Understanding regression analysis.* New York, NY: Plenum Press. pages 291

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444–455. pages 382

Appelbaum, M. I., & Cramer, E. M. (1974). Some problems in the nonorthogonal analysis of variance. *Psychological Bulletin*, *81*, 335–343. pages 13, 22, 23

Arbuckle, J. L. (2006). AMOS 7.0 user's guide [Computer software manual]. Chicago, IL: SPSS. pages XIII

Barnard, J., Du, J., Hill, J. L., & Rubin, D. B. (1998). A broader template for analyzing broken randomized experiments. *Sociological Methods & Research*, *27*, 285–317. pages 382

Barnard, J., Frangakis, C. E., Hill, J. L., & Rubin, D. B. (2003, June). Principal stratification approach to broken randomized experiments: A case study of school choice vouchers in new york city. *Journal of the American Statistical Association*, *98*(462), 299–323. pages 382

Bentler, P. M. (1995). EQS Structural Equations program manual [Computer software manual]. Encino, CA: Multivariate Software. pages XIII

Berger, M. P. F., & Wong, W. K. (Eds.). (2005). *Applied optimal designs.* Chichester, England: Wiley. pages 356

Biesanz, J. C., Deeb-Sossa, N., Aubrecht, A. M., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods*, *9*, 30–52. pages 42

Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least-squares estimator. *Sociological Methodology*, *25*, 223–251. pages 382

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach.* Hoboken, NJ: Wiley. pages 42

Bonney, G. E. (1987). Logistic regression for dependent binary observations. *Biometrics*, *43*, 951–973. pages 254

Borooah, V. K. (2001). *Logit and probit: Ordered and multinomial models.* Thousand Oaks, CA: Sage. pages 254

Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology, 163 (12),* 1149–1156. pages 371

Browne, M. W., & Mels, G. (1998). Path analysis: RAMONA. In *SYSTAT for Windows: Advanced Applications (Version 8) [Computer software manual].* Evanston, IL: SYSTAT. pages XIII

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching.* Chicago, IL: Rand McNally. pages VI, IX, 22, 36, 382

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Boston, MA: Houghton Mifflin. pages XIII, 121

Carlson, J. E., & Timm, N. H. (1974). Analysis of nonorthogonal fixed-effects designs. *Psychological Bulletin, 81,* 563–570. pages 13

Cartwright, N. (1979). Causal laws and effective strategies. *Noûs, 13,* 419–437. pages XII, 303

Cheng, J., & Small, D. S. (2006). Bounds on causal effects in three-arm trials with noncompliance. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 68,* 815–836. pages 23

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression / correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum. pages 23, 291

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago, IL: Rand McNally. pages VI, IX, XIII, 22, 382

Cox, D. R., & Wermuth, N. (2004). Causality: A statistical view. *International Statistical Review, 72,* 285–305. pages XIII

Draper, N. R., & Smith, H. (1998). *Applied regression analysis.* New York, NY: Wiley. pages 291

Dunn, G., Maracy, M., Dowrick, C., Ayuso-Mateos, J. L., Dalgard, O. S., Page, H., ... Wilkinson, G. (2003). Estimating psychological treatment effects from a randomised controlled trial with both non-compliance and loss to follow-up. *British Journal of Psychiatry, 183,* 323–331. pages 23

Fahrmeir, L., Hamerle, A., & Tutz, G. (1996). *Multivariate statistische Verfahren. [Multivariate statistical methods]* (2nd ed.). Berlin, Germany: de Gruyter. pages 382

Fichman, M., & Cummings, J. N. (2003). Multiple imputation for missing data: Making the most of what you know. *Organizational Research Methods, 6,* 282–308. pages 254, 291

Fisher, R. A. (1925/1946). *Statistical methods for research workers* (10th ed.). Edinburgh, England: Oliver and Boyd. pages VII, 227, 251

Gelman, A., & Hill, J. L. (2007). *Data analysis using regression and multilevel / hierarchical models.* New York, NY: Cambridge University. pages 291

Gosslee, D. G., & Lucas, H. L. (1965). Analysis of variance of disproportionate data when interaction is present. *Biometrics, 21,* 115–133. pages 13, 23

Graham, J. W., & Donaldson, S. I. (1993). Evaluating interventions with differential attrition: The importance of nonresponse mechanisms and use of follow-up data. *Journal of Applied Psychology, 78,* 119–128. pages 254, 291

Green, W. H. (2003). *Econometric analysis* (5th ed.). Upper Saddle River, NJ: Prentice-Hall. pages 254

Greenland, S. (1996). Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology*, *25*, 1107–1116. pages 382

Greenland, S. (2000). Causal analysis in the health sciences. *Journal of the American Statistical Association*, *95*, 286–289. pages 22, 382

Greenland, S. (2001). Sensitivity analysis, Monte Carlo risk analysis and Bayesian uncertainty assessment. *Risk Analysis*, *21*, 579–583. pages 382

Greenland, S. (2004). An overview of methods for causal inference from observational studies. In A. Gelman & X.-L. Meng (Eds.), *Applied bayesian modeling and causal inference from incomplete-data perspectives* (pp. 3–14). Chichester, England: Wiley. pages 22

Greenland, S., & Robins, J. M. (1986). Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology*, *15*, 413–419. pages 187

Györfi, L., Kohler, M., Krzyzak, A., & Walk, H. (2002). *A distribution-free theory of nonparametric regression.* New York, NY: Springer. pages 291

Hernán, M. A., Clayton, D., & Keiding, N. (2011). The Simpson's paradox unraveled. *International Journal of Epidemiology*, 1–6. (Advance Access published March 30, 2011) doi: 10.1093/ije/dyr041 pages 4

Höfler, M. (2005). Causal inference based on counterfactuals. *BMC Medical Research Methodology*, *5*, 1–12. pages 22

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*, 945–960. pages IX, 3, 6, 8, 121, 166

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley. pages 254

Jennings, E., & Green, J. L. (1984). Resolving nonorthogonal ANOVA disputes using cell means. *Journal of Experimental Education*, *52*, 159–162. pages 13

Jo, B. (2002a). Estimation of intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics, 27*, 385–409. pages 23, 33

Jo, B. (2002b). Model misspecification sensitivity analysis in estimating causal effects of interventions with non-compliance. *Statistics in Medicine*, *21*, 3161–3181. pages 23, 33, 382

Jo, B. (2002c). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods, 7*, 178–193. pages 23, 33, 382

Jo, B., Asparouhov, T., Muthén, B. O., Ialongo, N. S., & Brown, C. H. (2008). Cluster randomized trials with treatment noncomplicance. *Psychological Methods*, *13*, 1–18. pages 23, 33

Jöreskog, K. G., & Sörbom, D. (1996/2001). *LISREL 8: User's Reference Guide. [Computer software manual]* (2nd ed.). Lincolnwood, IL: Scientific Software International. pages XIII

Keele, L. (2008). *Semiparametric regression for the social sciences.* Chichester, England: Wiley. pages 358

Kenny, D. A. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychology Bulletin*, *82*, 887–903. pages 42

Keren, G., & Lewis, C. (1976). Nonorthogonal designs: Sample versus population. *Psychological Bulletin*, *83*, 817–826. pages 13

King, G., & Zeng, L. (2001). Improving forecasts of state failure. *World Politics*, *53*, 623–658. pages 358

Klenke, A. (2020). *Probability theory – A comprehensive course* (3rd ed.). Cham, Switzerland: Springer. doi: 10.1007/978-3-030-56402-5 pages 48

Kolmogorov, A. N. (1933/1977). *Grundbegriffe der Wahrscheinlichkeitsrechnung [Foundations of the theory of probability]* (Reprinted ed.). Berlin, Germany: Springer. pages XI, 117

Kramer, C. Y. (1955). On the analysis of variance of a two-way classification with unequal sub-class numbers. *Biometrics*, *11*, 441–452. pages 13

Langsrud, Ø. (2003). ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. *Statistics and Computing*, *13*, 163–167. pages 13

Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, *29*, 337–346. doi: 10.1002/sim.3782 pages 358

Liao, T. F. (1994). *Interpreting probability models: Logit, probit, and other generalized linear models*. London, England: Sage. pages 254

Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, *68*, 304–305. pages 23

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum. pages 23

Mayer, A., Dietzfelbinger, L., Rosseel, Y., & Steyer, R. (2016). The EffectLiteR approach for analyzing average and conditional effects. *Multivariate Behavioral Research*, *51*, 374–391. doi: 10.1080/00273171.2016.1151334 pages XVI

Mayer, A., & Thoemmes, F. (2019). Analysis of variance models with stochastic group weights. *Multivariate Behavioral Research*, *54*, 542–554. doi: 10.1080/00273171.2018.1548960 pages 13, 16

Mayer, A., Thoemmes, F., Rose, N., Steyer, R., & West, S. G. (2014). Theory and analysis of total, direct, and indirect causal effects. *Multivariate Behavioral Research*, *49*(5), 425–442. doi: 10.1080/00273171.2014.931797 pages XII, 186

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analysis. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future* (pp. 341–380). Lincolnwood, IL: Scientific Software International. pages 42

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403–425. pages 358

McCullagh, P., & Nelder, J. A. (1989). *Monographs on statistics and applied probability: Vol. 37. Generalized linear models* (2nd ed.; D. R. Cox, D. V. Hinkley, N. Reid, D. B. Rubin, & D. V. Silverman, Eds.). Chapman & Hall. pages 254

Meredith, M., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122. pages 42

Mill, J. S. (1843/1865). Of the four methods of experimental inquiry. In *A system of logic, ratiocinative and inductive: Volume 1. Being a connected view of the principles of evidence, and the methods of scientific investigation*. London, England: Longmans, Green, and Co. pages VI, 122, 145

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference. methods and principles for social research.* New York, NY: Cambridge University. pages 22

Muthén, L. K., & Muthén, B. O. (1998-2007). Mplus User's Guide (5th ed.) [Computer software manual]. Los Angeles, CA: Muthén & Muthén. pages XIII

Nagengast, B. (2009). *Causal inference in multilevel models* (Unpublished doctoral dissertation). Friedrich-Schiller-Universität Jena, Thüringen, Germany. pages 39

Nagengast, B., Kröhne, U., Bauer, M., & Steyer, R. (2007). Causal Effects Explorer: A didactic tool for teaching the theory of individual and average causal effects [Computer software manual]. University of Jena, Thüringen, Germany. pages XV

Nelder, J. A., & Lane, P. W. (1995). The computer analysis of factorial experiments: In memoriam – Frank Yates. *The American Statistician*, *49*, 382–385. pages 13

OpenMx. (2009). *OpenMx - Advanced Structural Equation Modeling [Computer Software].* Retrieved from http://openmx.psyc.virginia.edu/. pages XIII

Overall, J. E., & Spiegel, D. K. (1969). Concerning least squares analysis of experimental data. *Psychological Bulletin*, *72*(5), 311–322. pages 13

Overall, J. E., & Spiegel, D. K. (1973a). Comment on "Regression analysis of proportional cell data". *Psychological Bulletin*, *80*, 28–30. pages 13

Overall, J. E., & Spiegel, D. K. (1973b). Comments on Rawlings' nonorthogonal analysis of variance. *Psychological Bulletin*, *79*, 164–167. pages 13

Overall, J. E., Spiegel, D. K., & Cohen, J. (1975). Equivalence of orthogonal and nonorthogonal analysis of variance. *Psychological Bulletin*, *82*, 182–186. pages 13, 23

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge, UK. pages XIII, 22, 126

Pearson, K., Lee, A., & Bramley-Moore, L. (1899). Genetic (reproductive) selection: Inheritance of fertility in man, and of fecundity in thoroughbred racehorses. *Series A, Containing Papers of a Mathematical or Physical Character: Philosophical Transactions of the Royal Society of London*, *192*, 257–330. pages 3

Popper, K. (2005). *The logic of scientific discovery.* Taylor & Francis. pages 255

Porta, M. (2014). *A dictionary of epidemiology.* Oxford University Press. doi: 10.1093/acref/9780199976720.001.0001 pages 209

Pukelsheim, F. (2006). *Optimal design of experiments.* Philadelphia, PA: SIAM (Society for Industrial and Applied Mathematics). pages 356

Reichardt, C. S. (1979). The statistical analysis of data from nonequivalent group designs. In T. D. Cook & D. C. Campbell (Eds.), *Quasi-experimentation: Design and analysis issues for field settings.* Orlando, FL: Houghton Mifflin. pages 23

Reichenbach, H. (1956). *The direction of time* (M. Reichenbach, Ed.). University of California Press. pages 271

Robins, J., & Rotnitzky, A. (2004). Estimation of treatment effects in randomised trials with non-compliance and a dichotomous outcome using structural mean models. *Biometrika*, *91*, 763–783. pages 23

Robins, J. M. (1998). Correction for non-compliance in equivalence trials. *Statistics in Medicine*, *17*, 269–302. pages 23, 382

Robins, J. M., Hérnan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, *11*, 550–560. pages 382

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*, 245–258. pages 42

Rosenbaum, P. R. (2002a). Attributing effects to treatment in matched observational studies. *Journal of the American Statistical Association*, *97*(457), 183–192. doi: 10.1198/016214502753479329 pages 22, 33

Rosenbaum, P. R. (2002b). *Observational studies* (2nd ed.). New York, NY: Springer. pages 382

Rosenbaum, P. R., & Rubin, D. B. (1983a). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *45*, 212–218. pages 361

Rosenbaum, P. R., & Rubin, D. B. (1983b). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55. doi: 10.1093/biomet/70.1.41 pages 23, 199, 207, 209

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, *79*, 516–524. pages XV, 361

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. doi: 10.18637\slashjss.v048.i02 pages XIII

Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins. pages 8

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*, 688–701. pages 22, 117, 121

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, *100*, 322–331. doi: 10.1198/016214504000001880 pages 22, 117, 121, 126

Rubin, D. B. (2006). *Matched sampling for causal effects*. New York, NY: Cambridge University. pages 22

Searle, S. R., Casella, G., & McCulloch, C. E. (1992). *Variance components*. New York, NY: Wiley. pages 382

Sengewald, M.-A., Steiner, P. M., & Pohl, S. (2019). When does measurement error in covariates impact causal effect estimates? analytic derivations of different scenarios and an empirical illustration. *British Journal of Mathematical and Statistical Psychology*, *72*, 244–270. doi: doi.org/10.1111/bmsp.12146 pages 34

Senn, S. (2006). Change from baseline and analysis of covariance revisited. *Statistics in Medicine*, *25*, 4334–4344. pages 23

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin. pages VI, IX, XIII, 22, 254, 291, 380, 382, 383, 384

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *13*, 238–241. pages 3, 4

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University. pages 42

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge, MA: MIT. pages XIII, 22

Splawa-Neyman, J. (1923/1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9 (Reprinted from roczniki nauk rolniczych tom 10, pp. 1–51, 1923). *Statistical Science*, *5*, 465–480. pages VII, VIII, 6, 22, 117

Spohn, W. (1980). Stochastic independence, causal independence, and shieldability. *Journal of Philosophical Logic*, *9*, 73–99. pages XII

Stegmüller, W. (1983). *Erklärung, Begründung, Kausalität: Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie. [Explanation, justification, causality: Problems and findings of philosophy of science and analytic philosophy]*. Berlin, Germany: Springer. pages XII

Steyer, R. (1984). Causal linear stochastic dependencies: An introduction. In J. R. Nesselroade & A. von Eye (Eds.), *Individual development and social change: Explanatory analysis.* New York, NY: Academic Press. pages 303

Steyer, R. (1992). *Theorie kausaler Regressionsmodelle [Theory of causal regression models].* Stuttgart, Germany: Fischer. pages 303

Steyer, R. (2001). Classical test theory. In C. Ragin & T. D. Cook (Eds.), *International encyclopedia of the social and behavioural sciences: Logic of inquiry and research design* (pp. 481–520). Oxford, England: Pergamon. pages 39

Steyer, R. (2005). Analyzing individual and average causal effects via structural equation models. *Methodology*, *1*, 39–54. pages 42

Steyer, R. (2024). *Introduction to probability and conditional expectation.* Jena, Germany: University of Jena. pages XIV, XV, 47, 83, 116, 156, 199, 228, 271

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research Online*, *2*, 21–33. pages 42

Steyer, R., Gabler, S., von Davier, A. A., & Nachtigall, C. (2000). Causal regression models II: Unconfoundedness and causal unbiasedness. *Methods of Psychological Research Online*, *5*, 55–87. pages 303

Steyer, R., Mayer, A., Geiser, C., & Cole, D. (2015). A theory of states and traits - revised. *Annual Review of Clinical Psychology*, *11*, 71–98. pages 31, 34, 39, 40

Steyer, R., & Nagel, W. (2017). *Probability and conditional expectation. Fundamentals for the Empirical Sciences.* Chichester, England: Wiley. pages XIV, XV, 104

Suppes, P. (1970). *A probabilistic theory of causality.* Amsterdam, Netherlands: North-Holland. pages XII, 271

Takezawa, K. (2005). *Introduction to nonparametric regression.* New York, NY: Wiley. pages 382

Tisak, J., & Tisak, M. S. (2000). Permanency and ephemerality of psychological measures with application to organizational commitment. *Psychological Methods*, *5*, 175–198. pages 42

van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, *59*, 920–925. doi: 10.1016/j.jclinepi.2006.02.007 pages 23

von Eye, A., & Schuster, C. (1998). *Regression analysis for social sciences.* San Diego, CA: Academic Press. pages 291

Wainer, H. (1991). Adjusting for differential base rates: Lord's paradox again. *Psychological Bulletin*, *109*, 147–151. pages 23

Wang, Z., & Rousseau, R. (2021). Covid-19, the yule-simpson paradox and research evaluation. *Scientometrics*, *126*, 3501–3511. doi: doi.org/10.1007/s11192-020-03830-w pages 4

Watkins, M. W., Lei, P., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence*, *35*, 59–68. pages 42

West, S. G., & Aiken, L. S. (2005). Multiple linear regression. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science.* Chichester, England: Wiley. pages 291

West, S. G., Biesanz, J. C., & Pitts, S. C. (2000). Causal inference and generalization in field settings experimental and quasi-experimental designs. In H. T. Reis & C. M. Judd

(Eds.), *Handbook of research methods in social and personality psychology* (pp. 40–84). New York, NY: Cambridge University. pages 22

West, S. G., & Sagarin, B. J. (2000). Participant selection and loss in randomized experiments. In L. Bickman (Ed.), *Research design: Donald Campbell's legacy* (pp. 117–151). Thousand Oaks, CA: Sage. pages 380, 382

Williams, J. D. (1972). Two way fixed effects analysis of variance with disproportionate cell frequencies. *Multivariate Behavioral Research*, *7*, 57–83. pages 13

Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, *25*, 659–707. pages 22

Wolf, F. M., Chandler, T. A., & Spies, C. J. (1981). A crossed-lagged panel analysis of quality of school life and achievement responsibility. *Journal of Educational Research*, *74*, 363–368. pages 42

Woo, M.-J., Reiter, J. P., & Karr, A. F. (2008). Estimation of propensity scores using generalized additive models. *Statistics in Medicine*, *27*, 3805–3816. pages 358

Woodward, J. A., & Bonett, G. B. (1991). Simple main effects in factorial designs. *Journal of Applied Statistics*, *18*(2), 255–264. doi: 10.1080/02664769100000019 pages 145

Wright, S. (1918). On the nature of size factors. *Genetics*, *3*, 367–374. pages VIII

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, *20*(7), 557–585. pages VIII

Wright, S. (1923). The theory of path coefficients: A reply to Niles's criticism. *Genetics*, *8*, 239–255. pages VIII

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, *5*, 161–215. pages VIII

Wright, S. (1960a). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics*, *16*(2), 189–202. pages VIII

Wright, S. (1960b). The treatment of reciprocal interaction, with or without lag, in path analysis. *Biometrics*, *16*(3), 423–445. pages VIII

Wyss, R., Girman, C. J., LoCasale, R. J., Brookhart, M. A., & Stürmer, T. (2013). Variable selection for propensity score models when estimating treatment effects on multiple outcomes: A simulation study. *Pharmacoepidemiology and Drug Safety*, *22*, 77–85. doi: 10.1002/pds.3356 pages 371

Yule, G. U. (1903). Notes on the theory of association of attributes of statistics. *Biometrika*, *2*, 121–134. pages 3, 4