Introduction

According to an ancient Chinese fable, a farmer who was eager to help his crops grow went into his field and pulled each seedling upward. After exhausting himself with the work, he announced to his family that they were going to have a good harvest, only to find the next morning that the plants had wilted and died. Readers with minimal agricultural knowledge may immediately point out the following: the farmer’s intervention theory was based on a correct observation that crops that grow taller tend to produce more yield. Yet his hypothesis reflects a false understanding of the cause and the effect—that seedlings pulled to be taller would yield as much as seedlings thriving on their own.

In their classic Design and Analysis of Experiments, Hinkelmann and Kempthorne (1994; updated version of Kempthorne, 1952) discussed two types of science: descriptive science and the development of theory. These two types of science are interrelated in the following sense: observations of an event and other related events, often selected and classified for description by scientists, naturally lead to one or more explanations that we call “theoretical hypotheses,” which are then screened and falsified by means of further observations, experimentation, and analyses (Popper, 1963). The experiment of pulling seedlings to be taller was costly, but did serve the purpose of advancing this farmer’s knowledge of “what does not work.” To develop a successful intervention, in this case, would require a series of empirical tests of explicit theories identifying potential contributors to crop growth. This iterative process gradually deepens our knowledge of the relationships between supposed causes and effects—that is, causality—and may eventually increase the success of agricultural, medical, and social interventions.

1.1 Concepts of moderation, mediation, and spill-over

Although the story of the ancient farmer is fictitious, numerous examples can be found in the real world in which well-intended interventions fail to produce the intended benefits or, in many cases, even lead to unintended consequences. “Interventions” and “treatments,” used interchangeably in this book, broadly refer to actions taken by agents or circumstances experienced by an individual or groups of individuals. Interventions are regularly seen in education, physical and mental health, social services, business, politics, and law enforcement. In an
education intervention, for example, teachers are typically the agents who deliver a treatment to students, while the impact of the treatment on student outcomes is of ultimate causal interest. Some educational practices such as “teaching to the test” have been criticized to be nearly as counterproductive as the attempt of helping seedlings grow by pulling them upward. “Interventions” and “treatments” under consideration do not exclude undesired experiences such as exposure to poverty, abuse, crime, or bereavement. A treatment, planned or unplanned, becomes a focus of research if there are theoretical reasons to anticipate its impact, positive or negative, on the well-being of individuals who are embedded in social settings including families, classrooms, schools, neighborhoods, and workplaces.

In social science research in general and in policy and program evaluations in particular, questions concerning whether an intervention works and, if so, which version of the intervention works, for whom, under what conditions, and why are key to the advancement of scientific and practical knowledge. Although most empirical investigations in the social sciences concentrate on the average effect of a treatment for a specific population as opposed to the absence of such a treatment (i.e., the control condition), in-depth theoretical reasoning with regard to how the causal effect is generated, substantiated by compelling empirical evidence, is crucial for advancing scientific understanding.

First, when there are multiple versions or different dosages of the treatment or when there are multiple versions of the control condition, a binary divide between “the treatment” and “the control” may not be as informative as fine-grained comparisons across, for example, “treatment version A,” “treatment version B,” “control version A,” and “control version B.” For example, expanding the federally funded Head Start program to poor children is expected to generate a greater benefit when few early childhood education alternatives are available (call it “control version A”) than when there is an abundance of alternatives including state-sponsored preschool programs (call it “control version B”).

Second, the effect of an intervention will likely vary among individuals or across social settings. A famous example comes from medical research: the well-publicized cardiovascular benefits of initiating estrogen therapy during menopause were contradicted later by experimental findings that the same therapy increased postmenopausal women’s risk for heart attacks. The effect of an intervention may also depend on the provision of some other concurrent or subsequent interventions. Such heterogeneous effects are often characterized as moderated effects in the literature.

Third, alternative theories may provide competing explanations for the causal mechanisms, that is, the processes through which the intervention produces its effect. A theoretical construct characterizing the hypothesized intermediate process is called a mediator of the intervention effect. The fictitious farmer never developed an elaborate theory as to what caused some seedlings to surpass others in growth. Once scientists revealed the causal relationship between access to chemical nutrients in soil and plant growth, wide applications of chemically synthesized fertilizers finally led to a major increase in crop production.

Finally, it is well known in agricultural experiments that a new type of fertilizer applied to one plot may spill-over to the next plot. Because social connections among individuals are prevalent within organizations or through networks, an individual’s response to the treatment may similarly depend on the treatment for other individuals in the same social setting, which may lead to possible spill-overs of intervention effects among individual human beings.

Answering questions with regard to moderation, mediation, and spill-over poses major conceptual and analytic challenges. To date, psychological research often presents well-articulated theories of causal mechanisms relating stimuli to responses. Yet, researchers often
lack rigorous analytic strategies for empirically screening competing theories explaining the observed effect. Sociologists have keen interest in the spill-over of treatment effects transmitted through social interactions yet have produced limited evidence quantifying such effects. As many have pointed out, in general, terminological ambiguity and conceptual confusion have been prevalent in the published applied research (Holmbeck, 1997; Kraemer et al., 2002, 2008).

A new framework for conceptualizing moderated, mediated, and spill-over effects has emerged relatively recently in the statistics and econometrics literature on causal inference (e.g., Abbring and Heckman, 2007; Frangakis and Rubin, 2002; Heckman and Vytlacil, 2007a, b; Holland, 1988; Hong and Raudenbush, 2006; Hudgens and Halloran, 2008; Jo, 2008; Pearl, 2001; Robins and Greenland, 1992; Sobel, 2008). The potential for further conceptual and methodological development and for broad applications in the field of behavioral and social sciences promises to greatly advance the empirical basis of our knowledge about causality in the social world.

This book clarifies theoretical concepts and introduces innovative statistical strategies for investigating the average effects of multivalued treatments, moderated treatment effects, mediated treatment effects, and spill-over effects in experimental or quasiexperimental data. Defining individual-specific and population average treatment effects in terms of potential outcomes, the book relates the mathematical forms to the substantive meanings of moderated, mediated, and spill-over effects in the context of application examples. It also explicates and evaluates identification assumptions and contrasts innovative statistical strategies with conventional analytic methods.

1.1.1 Moderated treatment effects

It is hard to accept the assumption that a treatment would produce the same impact for every individual in every possible circumstance. Understanding the heterogeneity of treatment effects therefore is key to the development of causal theories. For example, some studies reported that estrogen therapy improved cardiovascular health among women who initiated its use during menopause. According to a series of other studies, however, the use of estrogen therapy increased postmenopausal women’s risk for heart attacks (Grodstein et al., 1996, 2000; Writing Group for the Women’s Health Initiative Investigators, 2002). The sharp contrast of findings from these studies led to the hypothesis that age of initiation moderates the effect of estrogen therapy on women’s health (Manson and Bassuk, 2007; Rossouw et al., 2007). Revelations of the moderated causal relationship greatly enrich theoretical understanding and, in this case, directly inform clinical practice.

In another example, dividing students by ability into small groups for reading instruction has been controversial for decades. Many believed that the practice benefits high-ability students at the expense of low-ability peers and exacerbates educational inequality (Grant and Rothenberg, 1986; Rowan and Miracle, 1983; Trimble and Sinclair, 1987). Recent evidence has shown that, first of all, the effect of ability grouping depends on a number of conditions including whether enough time is allocated to reading instruction and how well the class can be managed. Hence, instructional time and class management are among the additional moderators to consider for determining the merits of ability grouping (Hong and Hong, 2009; Hong et al., 2012a, b). Moreover, further investigations have generated evidence that contradicts the long-held belief that grouping undermines the interests of low-ability students. Quite the contrary, researchers have found that grouping is beneficial for students...
with low or medium prior ability—if adequate time is provided for instruction and if the class is well managed. On the other hand, ability grouping appears to have a minimal impact on high-ability students’ literacy. These results were derived from kindergarten data and may not hold for math instruction and for higher grade levels. Replications with different subpopulations and in different contexts enable researchers to assess the generalizability of a theory.

In a third example, one may hypothesize that a student is unlikely to learn algebra well in ninth grade without a solid foundation in eighth-grade prealgebra. One may also argue that the progress a student made in learning eighth-grade prealgebra may not be sustained without the further enhancement of taking algebra in ninth grade. In other words, the earlier treatment (prealgebra in eighth grade) and the later treatment (algebra in ninth grade) may reinforce one another; each moderates the effect of the other treatment on the final outcome (math achievement at the end of ninth grade). As Hong and Raudenbush (2008) argued, the cumulative effect of two treatments in a well-aligned sequence may exceed the sum of the benefits of two single-year treatments. Similarly, experiencing two consecutive years of inferior treatments such as encountering an incompetent math teacher who dampens a student’s self-efficacy in math learning in eighth grade and then again in ninth grade could do more damage than the sum of the effect of encountering an incompetent teacher only in eighth grade and the effect of having such a teacher only in ninth grade.

There has been a great amount of conceptual confusion regarding how a moderator relates to the treatment and the outcome. On one hand, many researchers have used the terms “moderators” and “mediators” interchangeably without understanding the crucial distinction between the two. An overcorrective attempt, on the other hand, has led to the arbitrary recommendation that a moderator must occur prior to the treatment and be minimally associated with the treatment (James and Brett, 1984; Kraemer et al., 2001, 2008). In Chapter 6, we clarify that, in essence, the causal effect of the treatment on the outcome may depend on the moderator value. A moderator can be a subpopulation identifier, a contextual characteristic, a concurrent treatment, or a preceding or succeeding treatment. In the earlier examples, age of estrogen therapy initiation and a student’s prior ability are subpopulation identifiers, the manageability of a class which reflects both teacher skills and peer behaviors characterizes the context, literacy instruction time is a concurrent treatment, while prealgebra is a preceding treatment for algebra. A moderator does not have to occur prior to the treatment and does not have to be independent of the treatment.

Once a moderated causal relationship has been defined in terms of potential outcomes, the researcher then chooses an appropriate experimental design for testing the moderation theory. The assumptions required for identifying the moderated causal relationships differ across different designs and have implications for analyzing experimental and quasiexperimental data. Randomized block designs are suitable for examining individual or contextual characteristics as potential moderators; factorial designs enable one to determine the joint effects of two or more concurrent treatments; and sequential randomized designs are ideal for assessing the cumulative effects of consecutive treatments. Multisite randomized trials constitute an additional type of designs in which the experimental sites are often deliberately sampled to represent a population of geographical locations or social organizations. Site memberships can be viewed as implicit moderators that summarize a host of features of a local environment. Replications of a treatment over multiple sites allow one to quantify the heterogeneity of treatment effects across the sites. Chapters 7 and 8 are focused on statistical methods for moderation analyses with quasiexperimental data. In particular, these chapters demonstrate, through a number of application examples, how a nonparametric marginal mean weighting through
stratification (MMWS) method can overcome some important limitations of other existing methods.

Moderation questions are not restricted to treatment–outcome relationships. Rather, they are prevalent in investigations of mediation and spill-over. This is because heterogeneity of treatment effects may be explained by the variation in causal mediation mechanisms across subpopulations and contexts. It is also because the treatment effect for a focal individual may depend on whether there is spill-over from other individuals through social interactions. Yet similar to the confusion around the concept of moderation among applied researchers, there has been a great amount of misunderstanding with regard to mediation.

1.1.2 Mediated treatment effects

Questions about mediation are at the core of nearly every scientific theory. In its simplest form, a theory explains why treatment $Z$ causes outcome $Y$ by hypothesizing an intermediate process involving at least one mediator $M$ that could have been changed by the treatment and could subsequently have an impact on the outcome. A causal mediation analysis then decomposes the total effect of the treatment on the outcome into two parts: the effect transmitted through the hypothesized mediator, called the indirect effect, and the difference between the total effect and the indirect effect, called the direct effect. The latter represents the treatment effect channeled through other unspecified pathways (Alwin and Hauser, 1975; Baron and Kenny, 1986; Duncan, 1966; Shadish and Sweeney, 1991). Most researchers in social sciences have followed the convention of illustrating the direct effect and the indirect effect with path diagrams and then specifying linear regression models postulated to represent the structural relationships between the treatment, the mediator, and the outcome.

In Chapter 9, we point out that the approach to defining the indirect effect and the direct effect in terms of path coefficients is often misled by oversimplistic assumptions about the structural relationships. In particular, this approach typically overlooks the fact that a treatment may generate an impact on the outcome not only by changing the mediator value but also by changing the mediator–outcome relationship (Judd and Kenny, 1981). An example in Chapter 9 illustrates such a case. An experiment randomizes students to either an experimental condition which provides them with study materials and encourages them to study for a test or to a control condition which provides neither study materials nor encouragement (Holland, 1988; Powers and Swinton, 1984). One might hypothesize that encouraging experimental group members to study will increase the time they spend studying, a focal mediator in this example, which in turn will increase their average test scores. One might further hypothesize that even without a change in study time, providing study materials to experimental group members will enable them to study more effectively than they otherwise would. Consequently, the effect of time spent studying might be greater under the experimental condition than under the control condition. Omitting the treatment-by-mediator interaction effect will lead to bias in the estimation of the indirect effect and the direct effect.

Chapter 11 describes in great detail an evaluation study contrasting a new welfare-to-work program emphasizing active participation in the labor force with a traditional program that guaranteed cash assistance without a mandate to seek employment (Hong, Deutsch, and Hill, 2011). Focusing on the psychological well-being of welfare applicants who were single mothers with preschool-aged children, the researchers hypothesized that the treatment would increase employment rate among the welfare recipients and that the treatment-induced increase in employment would likely reduce depressive symptoms under the new program.
They also hypothesized that, in contrast, the same increase in employment was unlikely to have a psychological impact under the traditional program. Hence, the treatment-by-mediator interaction effect on the outcome was an essential component of the intermediate process.

We will show that by defining the indirect effect and the direct effect in terms of potential outcomes, one can avoid invoking unwarranted assumptions about the unknown structural relationships (Holland, 1988; Pearl, 2001; Robins and Greenland, 1992). The potential outcomes framework has further clarified that in a causal mediation theory, the treatment and the mediator are both conceivably manipulable. In the aforementioned example, a welfare applicant could be assigned at random to either the new program or the traditional one. Once having been assigned to one of the programs, the individual might or might not be employed due to various structural constraints, market fluctuations, or other random events typically beyond her control.

Because the indirect effect and the direct effect are defined in terms of potential outcomes rather than on the basis of any specific regression models, it becomes possible to distinguish the definitions of these causal effects from their identification and estimation. The identification step relates a causal parameter to observable population data. For example, for individuals assigned at random to an experimental condition, their average counterfactual outcome associated with the control condition is expected to be equal to the average observable outcome of those assigned to the control condition. Therefore, the population average causal effect can be easily identified in a randomized experiment. The estimation step then relates the sample data to the observable population quantities involved in identification while taking into account the degree of sampling variability.

A major challenge in identification is that the mediator–outcome relationship tends to be confounded by selection even if the treatment is randomized. Chapter 9 reviews various experimental designs that have been proposed by past researchers for studying causal mediation mechanisms. Chapter 10 compares the identification assumptions and the analytic procedures across a wide range of analytic methods. These discussions are followed by an introduction of the ratio-of-mediator-probability weighting (RMPW) strategy in Chapter 11. In comparison with most existing strategies for causal mediation analysis, RMPW relies on relatively fewer identification assumptions and model-based assumptions. Chapters 12 and 13 will show that this new strategy can be applied broadly with extensions to multilevel experimental designs and to studies of complex mediation mechanisms involving multiple mediators.

### 1.1.3 Spill-over effects of a treatment

It is well known in vaccination research that an unvaccinated person can benefit when most other people in the local community are vaccinated. This is viewed as a spill-over of the treatment effect from the vaccinated people to an unvaccinated person. Similarly, due to social interactions among individuals or groups of individuals, an intervention received by some may generate a spill-over impact on others affiliated with the same organization or connected through the same network. For example, effective policing in one neighborhood may drive offenders to operate in other neighborhoods. In evaluating an innovative community policing program, researchers found that a neighborhood not assigned to community policing tended to suffer if the surrounding neighborhoods were assigned to community policing. This is an instance in which a well-intended treatment generates an unintended negative spill-over effect. However, being assigned to community policy was found particularly beneficial when the surrounding neighborhoods also received the intervention (Verbitsky-Savitz and Raudenbush,
2012). In another example, the impact of a school-based mentoring program targeted at students displaying delinquent behaviors may be enhanced for a focal student if his or her at-risk peers are assigned to the mentoring program at the same time.

The previous examples challenge the “stable unit treatment value assumption” (SUTVA) that has been invoked in most causal inference studies. This assumption states that an individual’s potential response to a treatment depends neither on how the treatment is assigned nor on the treatment assignments of other individuals (Rubin, 1980, 1986). Rubin and others believe that, without resorting to SUTVA, the causal effect of a treatment becomes hard to define and that causal inference may become intractable. However, as Sobel (2006) has illustrated with the Moving to Opportunity (MTO) experiment that offered housing vouchers enabling low-income families to move to low-poverty neighborhoods, there are possible consequences of violating SUTVA in estimating the treatment effects on outcomes such as safety and self-sufficiency. Because social interactions among individuals may affect whether one volunteers to participate in the study, whether one moves to a low-poverty neighborhood after receiving the housing voucher, as well as housing project residents’ subjective perceptions of their neighborhoods, the program may have a nonzero impact on the potential outcome of the untreated. As a result, despite the randomization of treatment assignment, the mean difference in the observed outcome between the treated units and the untreated units may be biased for the average treatment effect. Rather, the observable quantity is the difference between the effect of treating some rather than treating none for the treated and the effect of treating some rather than treating none for the untreated, the latter being the pure spill-over effect for the untreated.

In making attempts to relax SUTVA, researchers have proposed several alternative frameworks that incorporate possible spill-over effects (see Hong and Raudenbush, 2013, for a review). Hong (2004) presented a model that involves treatments and treatment settings. A treatment setting for an individual is a local environment constituted by a set of agents and participants along with their treatment assignments. An individual’s potential outcome value under a given treatment is assumed stable when the treatment setting is fixed; the potential outcome may take different values when the treatment setting shifts. One may investigate whether the treatment effect depends on the treatment setting. Applying this framework, Hong and Raudenbush (2006) examined the effect on a child’s academic growth of retaining the child in kindergarten rather than promoting the child to first grade when a relatively small proportion of low-achieving peers in the same school are retained as opposed to when a relatively large proportion of the peers are retained. Hudgens and Halloran (2008) presented a related framework in which the effect on an individual of the treatment received by this individual is distinguished from the effect on the individual of the treatment received by others in the same local community. These effects can be identified if communities are assigned at random to different treatment assignment strategies (such as retaining a large proportion of low-achieving students as opposed to retaining a small proportion of such students) and, subsequently, individuals within a community are randomized for treatment. Chapter 14 reviews these frameworks and discusses identification and estimation strategies.

Social contagion may also serve as an important channel through which the effect of an intervention is transmitted. For example, a school-wide intervention may reduce aggressive behaviors and thereby improve students’ psychological well-being by improving the quality of interpersonal relationships in other classes as well as in one’s own class. This is because children interact not simply with their classmates but also with those from other classes in the hallways or on the playground (VanderWeele et al., 2013). In this case, the spill-over between classes becomes a part of the mediation mechanism. In a study of student mentoring, whether
an individual actually participated in the mentoring program despite the initial treatment assignment is typically viewed as a primary mediator. The proportion of one’s peers that participated in the program may act as a second mediator. A treatment may also exert its impact partly through regrouping individuals and thereby changing peer composition (Hong and Nomi, 2012). Chapter 15 discusses analytic strategies for detecting spill-over as a part of the causal mediation mechanism.

1.2 Weighting methods for causal inference

In most behavioral and social science applications, major methodological challenges arise due to the selection of treatments in a quasiexperimental study and the selection of mediator values in experimental and quasiexperimental studies. Statistical methods most familiar to researchers in social sciences are often inadequate for causal inferences with regard to multivalued treatments, moderation, mediation, and spill-over. The book offers a major revision to understanding of causality in the social world by introducing two complementary weighting strategies, both featuring nonparametric approaches to estimating these causal effects. The new weighting methods greatly simplify model specifications while enhancing the robustness of results by minimizing the reliance on some key assumptions.

The propensity score-based \( \text{MMWS} \) method removes selection bias associated with a large number of covariates by equating the pretreatment composition between treatment groups (Hong, 2010a, 2012; Huang et al., 2005). Unlike propensity score matching and stratification that are mostly restricted to evaluations of binary treatments, the MMWS method is flexible for evaluating binary and multivalued treatments by approximating a completely randomized experiment. In evaluating whether the treatment effects differ across subpopulations defined by individual characteristics or treatment settings, researchers may assign weights within each subpopulation in order to approximate a randomized block design. To investigate whether one treatment moderates the effect of another concurrent treatment, researchers may assign weights to the data to approximate a factorial randomized design. The method can also be used to assess whether the effect of an initial treatment is amplified or weakened by a subsequent treatment or to identify an optimal treatment sequence through approximating a sequential randomized experiment. Even though such analyses can similarly be conducted through inverse-probability-of-treatment weighting (IPTW) that has been increasingly employed in epidemiological research (Hernán, Brumback, and Robins, 2000; Robins, Hernán, and Brumback, 2000), IPTW is known for bias and imprecision in estimation especially when the propensity score models are misspecified in their functional forms (Hong, 2010a; Kang and Schafer, 2007; Schafer and Kang, 2008; Waernbaum, 2012). In contrast, the nonparametric MMWS method displays a relatively high level of robustness despite such misspecifications and also gains efficiency, as indicated by simulation results (Hong, 2010a).

To study causal mediation mechanisms, the \( \text{RMPW} \) method decomposes the total effect of a treatment into an “indirect effect” transmitted through a specific mediator and a “direct effect” representing unspecified mechanisms. In contrast with most existing methods for mediation analysis, the RMPW-adjusted outcome model is extremely simple and is nonparametric in nature. It generates estimates of the causal effects along with their sampling errors while adjusting for pretreatment covariates that confound the mediator–outcome relationships through weighting. The method applies regardless of the distribution of the outcome, the distribution of the mediator, or the functional relationship between the outcome and the mediator.
Moreover, the RMPW method can easily accommodate data in which the mediator effect on the outcome may depend on the treatment assignment (Hong, 2010b). This is the case when a treatment produces its effects not only through changing the mediator value but also in part by altering the mediational process that normally produces the outcome (Judd and Kenny, 1981). One may use RMPW to further investigate whether the mediation mechanism varies across subpopulations and to disentangle complex mediation mechanisms involving multiple concurrent or consecutive mediators. A combination of RMPW with MMWS enables researchers to conduct mediation analysis when the treatment is not randomized. The RMPW approach to mediation analysis will be shown to have broad applicability to single-level and multilevel data (Hong & Nomi, 2012; Hong, Deutsch, and Hill, in press) when compared with path analysis, structural equation modeling (SEM), the instrumental variable method, and their recent extensions.

1.3 Objectives and organization of the book

The book consists of four major units: an overview followed by three units focusing on moderation, mediation, and spill-over, respectively. The first unit provides an overview of the concepts of causal effects, moderation, mediation, and spill-over. After reviewing the existing research designs and analytic methods for causal effect estimation, it introduces the basic rationale of weighted estimation of population parameters in survey sampling and explains the extension of the weighting approach to causal inference. Each subsequent unit addresses research questions of increasing complexity around one of the three focal topics. Part II considers treatment effects moderated by individual or contextual characteristics, by a concurrent treatment, or by a prior or subsequent treatment. Part III shows how to investigate the case of a single mediator, of multiple mediators, and of moderated mediation effects. Part IV discusses the spill-over of treatment effects and the spill-over of mediated effects.

Throughout the book, the mathematics is kept to a minimum to ease reading. Derivations and proofs are left to appendices for readers with technical interest. Assuming that some readers may have had little prior training in basic probability theory, a topic not always covered in a systematic way in an introductory-level applied statistics course, Chapter 2 provides a glossary of the basic notation employed in causal inference. Readers who have already taken a course in causal inference or its equivalent may easily skip Chapters 2 and 3. Chapters 4–11 are suitable for use in an undergraduate- or graduate-level applied statistics course dealing with moderation and mediation analyses. Those who are working on cutting-edge problems related to causal mediation and spill-over may find the last four chapters of the book particularly engaging.

This book aims to make the new analytic methods readily accessible to a wide audience in the social and behavioral sciences. Readers with sufficient understanding of multiple regression and analysis of variance (ANOVA) will quickly grasp the logic of the weighting methods and will find them easy to implement with existing statistical software. For example, after applying MMWS to a sample, researchers can estimate the effects of multivalued treatments or a moderated treatment effect simply within the ANOVA framework. For mediational studies, the RMPW method generates parameter estimates corresponding to the direct effect and the indirect effect along with their estimated sampling variability and therefore greatly simplifies hypothesis testing. In addition to the templates, along with data examples, for implementing MMWS and RMPW in SAS, Stata, and R, the book is accompanied by stand-alone
MMWS and RMPW software programs. The program interfaces are designed not only to greatly ease computation but also to assist the applied user with analytic decision-making. All these materials are available online free of charge at the publisher’s website: http://www.wiley.com/go/social_world.

Numerous examples from education, human development, psychology, sociology, and public policy motivate the development of innovative methods and help illustrate the concepts in the book. Analytic procedures are demonstrated through a series of case studies such as those discussed earlier. The examples are chosen to represent the causal questions often raised in behavioral and social science research and hence serve as prototypes for applying the analytic methods. In each case, statistical solutions are offered in sufficient detail to allow for replications.

1.4 How is this book situated among other publications on related topics?

Many other authors have contributed to research on moderation and mediation analyses and on causal inference in general. Such work has been systematically presented in four categories, in my view. The first category is textbooks on path analysis and SEM. The most popular ones include Kenneth A. Bollen’s (1989) *Structural Equations with Latent Variables*, David P. MacKinnon’s (2008) *Introduction to Statistical Mediation Analysis*, and Andrew F. Hayes’ (2013) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Bollen’s book provides a lucid and precise introduction to SEM focusing on combining structural models with measurement models. MacKinnon’s book summarizes later developments extending SEM to multilevel data, longitudinal data, and categorical data and relaxing distributional assumptions by employing computer-intensive estimation methods. Hayes’s book offers a gentle introduction to the most basic mediation and moderation analyses with the aim of integrating the two within the regression framework. These books, however, do not place a major emphasis on clarifying the identification assumptions crucial for distinguishing between association and causation.

Books in the second category provide a healthy antidote to the standard methods that have been employed by social scientists in addressing causal questions. Among the most well known are Steven D. Levitt and Stephen J. Dubner’s (2005) *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* and Charles Manski’s (1995) *Identification Problems in the Social Sciences*. These books were aimed at raising critical awareness among social scientists and the general public with regard to the potentials and limitations of the existing methods. They were excellent in revealing analytic difficulties and uncertainties in social science research.

The fourth category is research monographs each presenting a distinct cutting-edge approach to causal analysis. In this category, one may find Paul R. Rosenbaum’s (2002) *Observational Studies* and his (2009) *Design of Observational Studies*, Donald B. Rubin’s (2006) *Matched Sampling for Causal Effects*, and Judea Pearl’s (2009) *Causality: Models, Reasoning, and Inference*. The books by Rosenbaum or Rubin provide detailed discussions of how to use propensity score matching to overcome overt biases associated with the observables and how to use sensitivity analysis to account for hidden biases associated with the unobservables. The authors restricted the discussions primarily to evaluating the average effect of a binary treatment. Pearl’s book is unique in defining direct and indirect effects as mathematical functions of counterfactual outcomes. The definitions reflect an integrated understanding of probabilistic relationships and structural relationships. The book is also unique in its extensive use of graphical models and causal diagrams for determining identification. Yet the book may not satisfy practitioners searching for step-by-step analytic solutions directly applicable to their data.

The current book is built on an accumulation of past research represented in the aforementioned publications and in numerous journal articles. However, unlike the publications in the first category, this one is exclusively based on the potential outcomes causal framework with a clear definition for each causal effect. The nonparametric weighting methods presented in this book provide useful alternatives to SEM by relaxing key assumptions and simplifying modeling specification and estimation. Unlike the books in the second category, this one not only criticizes naïve approaches to causal reasoning but also demonstrates analytic strategies enabling researchers to conduct in-depth investigations of real-world problems and arrive at meaningful substantive conclusions. Unlike those in the third category, this book is exclusively focused on moderation, mediation, and spill-over effects and makes original contributions by introducing innovative analytic methods designed for addressing these fundamental issues. This book also distinguishes itself from those in the fourth category by teaching innovative weighting methods that can be readily implemented with most existing statistical software. Finally, most of the books mentioned previously have not given particular considerations of spill-over effects. Discussions related to social interference and spill-over have been ongoing in the fields of econometrics (Durlauf and Young, 2001; Manski, 1993, 2000) and political and social network analysis (Scott and Carrington, 2011; Sinclair, 2012; Wasserman and Faust, 1994) and have increasingly drawn attention from statistics. The current book contributes to the discourse by illuminating a central role of spill-over in social intervention theories from a causal inference perspective.

In summary, this book complements yet cannot replace many other important publications mentioned earlier. Readers who need a broad overview of a wide range of research designs and causal inference methods will likely benefit from reading the third category of publications; those who seek in-depth information of a particular causal inference perspective or strategy other than the weighting approaches presented here are referred to the fourth category of publications.

References


(INTRODUCTION)


Hong, G., Pelletier, J., Hong, Y., and Corter, C. (2012b). Does literacy instruction affect kindergartners’ externalizing problem behaviors as well as their literacy learning? Taking class manageability into account. The University of Chicago and the Ontario Institute for Studies in Education of the University of Toronto.


16 CAUSALITY IN A SOCIAL WORLD


INTRODUCTION


