Bringing Context and Variability
Back in to Causal Analysis*

Stephen L. Morgan
Cornell University

Christopher Winship
Harvard University

Prepared for inclusion as Chapter 14 of
Oxford Handbook of the Philosophy of the Social Sciences,
edited by Harold Kincaid.

January 2012
(Prior versions: April 2010; August 2010; September 2010; October 2010)

Bringing Context and Variability
Back in to Causal Analysis

The methodology of causal analysis in the social sciences is often divided into two ideal type research scenarios: experimental social science and observational social science. For experimental social science, the researcher can manipulate the cause of interest. The most common research design is one where the analyst assigns values of the cause according to a randomization scheme and then calculates post-treatment differences in outcomes across levels of the assigned cause. Typically, the researcher gives little or no attention to individual-specific differences in the inferred causal effects or to the context in which the experiment is conducted.[1]

For observational social science, the analyst cannot manipulate the cause through intervention because some process outside of the analyst's control determines the pattern of causal exposure. To develop causal assertions, the analyst must adopt a model of causal exposure based on assumptions about how the cause is distributed in the population. Most commonly, a model is adopted that warrants causal inference from differences in outcomes calculated within sets of observed individuals who are exposed to alternative values of the cause but who are deemed otherwise comparable by the maintained model of causal exposure. Individual-level variation in causal effects is then presumed to exist within and across comparison sets, often arising from interactions between individuals' characteristics and the contexts within which they are exposed to the cause.

In this chapter, we will discuss methods for modeling causal effects in observational social science, giving particular attention to the capacity of new graphical methods to represent and then motivate models that can effectively deliver estimates

---

[1]This does not mean that the results imply that individual-level differences do not exist. All such differences are, by construction, balanced in expectation across arms of the randomization scheme. Moreover, for the many experiments that are performed on convenience-based collections of individuals willing to participate (e.g., college students), this balancing often does not even generate the average causal effect in a target population of fundamental interest.

of underlying heterogeneity of causal effects. We have several related goals that we will pursue in the following order: (1) explain why quantitatively oriented social science that adopted path modeling methodology became a target of critiques that it had ignored variability and context, (2) demonstrate how such effects can be expressed within a more recent methodology of causal graphs, (3) consider feasible empirical strategies to identify these effects, and (4) explain why causal graphs pose a risk of obscuring patterns of heterogeneity that deserve full scrutiny.

To set the stage for our explanations, consider some classic examples from sociology that have sought to model explicitly the effects of individual-level heterogeneity of causal effects as they interact with consequential social contexts. At least since the 1980s, sociologists have investigated the effects of neighborhoods on educational outcomes, deviance, and the transition to adulthood (for insightful reviews, see Jencks and Mayer 1990 and Harding, Gennetian, Winship, Sanbonmatsu, and Kling 2011). Because neighborhoods have many characteristics, and individuals living within them can be influenced to varying degrees by circumstances only partly under their own control, the effects of neighborhoods have proven persistently difficult to estimate. These debates have not been settled by first-rate observational data analysis or by large-scale experimentation (see Sampson 2008).

Alongside this work on neighborhoods, sociologists of education have studied the variable effects of schooling on the academic achievement of students. These studies include attempts to estimate the differential effects of public schooling on learning for students from different socioeconomic strata. For example, Downey, von Hippel, and Broh (2004) show that schools help to narrow differences in learning that would result only from differences attributable to baseline family background differences. A complementary stream of literature has shown that Catholic schooling may generate even larger "common school effects" (see Hoffer, Greeley, and Coleman 1985; Bryk, Lee, and Holland 1993), though this pattern may instead reflect

differential self-selection on the causal effect itself (see Morgan and Todd 2008).

Sociologists have also considered the differential consequences of labor market conditions and training opportunities for young adults. For example, Mare and Winship (1984) studied the extent to which changes in the unemployment rate for black youths can be considered differential responses to labor market conditions across youths who have differential propensities to enter the military or postsecondary schooling. More recently, Brand and Xie (2010) have studied the differential payoff of college across different types of students, challenging the position implicitly maintained by many economists that college provides the greatest benefits to those most likely to enter college.

Our focal example in this chapter will be the contentious research on charter schooling in the United States that has been the subject of substantial and recent public debate. In an excellent book on these debates, Henig (2008:2) introduces and defines charter schools in the following way:

> Just a little more than fifteen years since the first charter school opened in Minnesota, there are now nearly 4,000 nationwide, serving an estimated 1.1 million students. ... The laws governing charter schools differ – sometimes substantially – from state to state, of course, but some general characteristics have emerged. Charter schools receive public funding on a per-student basis, are often responsible for achieving educational outcomes defined by their government chartering entity, and are subject to at least nominal public oversight. They typically are barred from charging tuition on top of the public per-pupil allocation, but are free to pursue other forms of supplementary support from donors, foundations, or corporate sponsors. Although they must observe certain baseline regulations, such as prohibitions on discrimination and the provision of safe environments, they are exempt from many of the rules and regulations that bind regular public schools to specific standards and procedures. This hybrid status ... has made charter schools a special focus of attention and helped draw them into ideological whirlpools that raise the stakes surrounding the research into their actual form and consequences.

At their core, the central research questions in the debate are simple: Do students who attend charter schools perform better on standardized tests than they would

have performed if they had instead attended regular public schools? Would students who attend regular public schools perform better on standardized tests if they had instead attended charter schools?

The contentious research that has addressed these questions is distinguished in many respects. Not only are some of its combatants leading researchers at the nation's top universities, many of these researchers are unusually ideological (as Henig shows brilliantly in his book). This scholarly energy is amplified by the public attention that has been paid to charter schools by the national press, which is related to the support that charter schools have received from celebrity donors and from presidential aspirants. At the same time, the research that informs the debate is cutting edge in the best sense. Careful attention is paid to details of measurement, and the research designs that have been adopted are a healthy mixture of basic comparisons of achievement levels as well as daring attempts to leverage quasi-experimental variation from the ways in which charter school programs are administered.[2]

What makes pursuing these questions complex is the underlying heterogeneity of the real world. The process by which some students become enrolled in charter schools is only partly observed. It is likely that some students in charter schools are much more likely to benefit from them than others, and it is even more difficult to assess how students who never contemplated entering charter schools might fare if given the opportunity to attend them. At the same time, charter schools differ greatly from each other, such that the effect of charter schooling must surely vary because of quality differences, as well as the match between each student and the unique features of each charter school.

In the next section, we provide necessary background for our subsequent presentation of the new methodology of causal graphs by first offering a presentation of

---

[2]For very high quality examples of this research, see Abdulkadiroglu, Angrist, and Kane et al. (2009), Angrist, Dynarski, and Kane et al. (2010), Center for Research on Educational Outcomes (2009), and Hoxby, Murarka, and Kang (2009).

the charter school effect from a path-modeling perspective. We also use this material to explain how quantitatively oriented sociology opened itself up to the critique that variability and context were too frequently ignored in attempts to estimate causal effects.[3]

An Emergent Vulnerability to Critique

In the 1980s and early 1990s, a robust critique of dominant forms of quantitative research arose in sociology (see Abbott 1988, 2001; Lieberson 1985; Lieberson and Lynn 2002; Ragin 1987). This literature objected to the overly strong causal assertions in the published literature in many areas of sociology, which these authors claimed were based on misplaced faith in the capacity of linear regression results to generate warranted causal conclusions from the analysis of survey data. A great deal of this critique of research practice was on target, and it arose in response to the naivety of what we labeled elsewhere "the age of regression" (see Morgan and Winship 2007).

We will not review these critiques in this chapter, but for this section we will use the models at the heart of these critiques – simple path diagrams and their underlying linear regression equations – as a point of departure. In the remainder of the chapter that follows this section, we will explain why we feel that this robust critique of quantitatively oriented causal analysis has now been weakened by improved practice that draws on a virtuous combination of causal graphs with nonparametric foundations and causal effects with potential outcome definitions.

To understand the graphical appeal of traditional path models, consider the

---

[3]Before carrying on, we should note that our title "Bringing Context and Variability Back in to Causal Analysis" is slightly misleading, since we will explain how context and variability have been brought back into causal analysis in the past fifteen years in ways that provide a solid foundation for future research. Thus, the tone of our chapter is optimistic and forward looking, not an indictment of current practice (which is often how the phrase "Bringing _ _ _ _ Back In" has been used in the long series of critical papers in sociology that followed from Homans' classic manifesto on methodological individualism, delivered as "Bringing Men Back In" in his 1964 Presidential Address to the American Sociological Association).

path diagram presented in Figure 1, and suppose that we have data on all sixth graders in a large metropolitan school district. For this path diagram, Y is a standardized test taken at the end of the sixth grade, and C indicates whether or not a student attended a charter school for the past year. The variable P represents an omnibus parental background measure that captures differences in economic standing and other basic dimensions of resources that predict school performance. The variable N is neighborhood of residence, and we assume that there are meaningful differences in the extent to which neighborhood environments are conducive to engagement with schooling.. Thus, C is the cause of primary interest, P is a baseline confounder that represents individual determinants of C that also have direct causes on the outcome Y, and N is a measure of the social context in which the effect of C on Y occurs.

The structure of the path diagram in Figure 1 implies that the proper regression specification for Y is

$$Y = a_Y + b_{C \to Y} C + b_{P \to Y} P + b_{N \to Y} N + e_Y, \qquad (1)$$

where e_Y is a regression representation of all omitted factors that determine Y. The final e_Y term in Equation 1 is suppressed in Figure 1 because it is assumed to be independent of C, P, and N. (Similarly, as we will discuss later, no other analogous error terms are depicted in Figure 1, such as e_C, e_P, or e_N.)

[INSERT FIGURE 1 ABOUT HERE]

In the literature on path models that swept through the social sciences in the 1960s and 1970s, the effects $b_{C \to Y}$, $b_{P \to Y}$, and $b_{N \to Y}$ were presumed to be linear and constant across individuals. Accordingly, the directed arrows in Figure 1 were interpreted as linear additive effects. With the example just specified, this assumption then requires that there are separable and linear additive effects of families, schools, and neighborhoods on student achievement, and furthermore, that these effects apply

with equal force to all sixth graders in the school district.[4]

How would such a path diagram have been presented and then discussed in a typical research methods class in the 1970s (assuming that the charter school research question was on the table)? Following an introduction to graphical representations of causal relationships via path models, at least one student would invariably ask the instructor:

> Can the effect of C on Y vary across P? That seems reasonable, since it would seem that the effect of a charter school would depend on family background. Parents with college degrees probably help their kids get more out of school. Actually, now that I think about it, since N captures neighborhood characteristics, don't we think that there are better schools in some neighborhoods? In fact, charter schools are more likely to be established in areas with troubled neighborhood-based schools. And neighborhoods with weaker schools also tend to have stronger deviant subcultures with gangs and such. So the effect of charter schooling probably also depends on the neighborhood in which one lives. How do we represent such variation in effects in the path model?[5]

In response, an instructor would typically explain that one can think of such effects as supplemental arrows from a variable to an arrow in the path diagram, such that the variable itself modifies the arrow. Yet, since these sorts of arrows are not formally justified in traditional path diagrams, the instructor would almost surely have then recommended a shift toward a more complex regression specification, such as

$$Y = a_Y + b_{C \to Y}C + b_{P \to Y}P + b_{CP \to Y}(C \times P) + b_{N \to Y}N + b_{CN \to Y}(C \times N) + e_Y. \quad (2)$$

In this case, the path diagram ceases to represent an underlying set of structural

---

[4]Such assumptions are entirely implausible, of course, and one doubts that any researchers would ever have endorsed them had they been asked to do so. Rather, such assumptions would have been implicitly maintained, in part because the methodological literature had not yet shown how restrictive such assumptions can be nor offered alternative models to relax them in productive ways.

[5]Were this exchange occurring in the substance of the day, a path model from the status attainment tradition would represent the substance of the exchange. The outcome variable Y would be career success, and C would be college education. All of the same interactions noted for the charter school case would then apply in this case, though based on different narratives of causation.

causal relationships and is instead best interpreted as only a simplified reflection of a more specific regression model. After all, the interaction between the effects of C and P on Y (as well as the interaction between the effects of C and N on Y) can be estimated with little trouble. One need only calculate effects of interest, for example, by plugging in values for $\hat{b}_{C \to Y} C + \hat{b}_{P \to Y} P + \hat{b}_{CP \to Y}(C \times P)$, after producing standard regression output from estimation of Equation 2. The differences then produced can be imbued with causal interpretations, assuming no other variables that are common causes of P, C, N, or Y have been mistakenly omitted from Equation 2.

We see two related outcomes of the rise and then demise of linear path models. First, when it became clear that there was no agreed upon way to represent variability and context within path diagrams, they came to seem much less useful. Researchers interested in such variability and context may have continued to draw path diagrams on yellow pads in their offices, but rarely did their drawings turn up in published articles. Path diagrams were thereby relegated to heuristic devices for laying out possible causal relationships of interest to researchers. Estimation and reporting became a word and number affair, often with too much of each.

Second, and far more troubling, many scholars apparently chose to retain a linear additive orientation, even while no longer using path diagrams. For some, empirical research could be fruitfully advanced by ignoring the genuine interactive nonlinearity of the real world, in pursuit of a "first approximation" pragmatism. This might have been an acceptable form of pragmatism if the approximation spirit had carried over to model interpretation. Too frequently it did not, and many causal assertions can be found in the literature based on linear additive models that are overly reductionist.

This incautious literature then opened up quantitative research to the claims of critics that too many practitioners had fallen prey to the belief that linear regression modeling reveals strong causal laws in which variability and context play minor roles.

The most cogent presentation of this criticism is Abbott's oft-cited "Transcending General Linear Reality" (see Abbott 1988). Although its straw-man style is irksome to methodologists who knew of these problems all along, and who urged better practice, it was a reasonable critique of much practice at the time. In the next section, we explain how the literature has since advanced by scaling back the ambitiousness of the regression enterprise, while at the same time adopting a dedicated set of notation to define causal effects that had no necessary connection to regression methodology.

Moving Beyond Path Diagrams and Simplistic Linear Regression Models

Consider the general multiple regression model of the form

$$Y = a + b_1 X_1 + b_2 X_2 + \ ... \ + b_k X_k + e_Y, \tag{3}$$

where Y is an interval-scaled outcome variable and X_1 through X_k are predictor variables. Estimation of the slope parameters b_1 through b_k can be motivated as a descriptive data reduction exercise where the goal is to obtain a best-fitting linear approximation to the population-level relationship between Y and X_1 through X_k. Alternatively, and more ambitiously, the model can be estimated as a full causal model where the interest is in identifying the expected shifts in Y that would result from what-if interventions on all possible values of the variables X_1 through X_k. The path model tradition embraced the second and more ambitious of these two approaches.

The more recent literature that we will consider in this chapter has examined an intermediate case. For this model, the variable X_1 in Equation 3 is an indicator variable C, as in

$$Y = a_Y + b_C C + b_2 X_2 + \ ... \ + b_k X_k + e_Y, \tag{4}$$

and the goal of the analysis is to estimate the causal effect on Y under what-if shifts of C from 0 to 1. In this case, the variables X_2 through X_k are considered adjustment variables that are entered into the regression equation solely to aid in the effective estimation of the causal effect b_C. Accordingly, estimates of b_2 through b_k are of secondary interest and are not necessarily given a causal interpretation.

Most importantly, under this motivation it is generally presumed that individual-level causal effects may vary, such that the effect on Y of shifting C from 0 to 1 is not the same for all individuals. For our focal charter school example, suppose that the population of interest is again sixth graders in a large metropolitan school district, and the outcome is performance on a standardized test. The variable C is again an indicator variable for whether a student attended a charter school in the past year, and the causal effect of charter schooling, b_C, is presumed to vary across students, either because of their own characteristics, those of their schools, or even those of the neighborhoods in which they live.

The first step in moving toward models that give substantial attention to heterogeneous effects was the introduction of interaction-based multilevel modeling. One way to introduce heterogeneity of this form into Equation 4 is to define b_C as a random coefficient and then to model it as a function of other variables. A simple version of this strategy was already introduced in the last section, where such variability would have been estimated by forming cross-product interactions between C and other variables within X_2 through X_k, which were variables such as P and N for parental background and neighborhoods (see Equation 2). A more general approach was then developed in a new multilevel modeling tradition, in which any variable could be seen as a predictor of the variability of b_C, especially higher-level predictors in nested social structures. Although technically these variables were specified as additional variables within X_2 through X_k, the multilevel modeling framework offered considerable conceptual appeal as well as some elegant methods for modeling variance

components (see Raudenbush and Bryk 2002 and Gelman and Hill 2007).

At the same time that this multilevel modeling literature was reaching maturity, and its power for modeling heterogeneity came to be utilized in practice, a potential outcomes model of causality, which had been in development since at least the early 1970s, took center stage in observational data analysis from the 1990s onward. The most important foundational work on this potential outcomes framework was completed in statistics and econometrics (see the citations offered in Heckman 2000, Manski 1995, Rosenbaum 2002, and Rubin 2005, 2006 to their own work and that of their respective predecessors).

Consider the focal charter schools example again. The outcome of interest Y remains a standardized test score for sixth graders. Within the potential outcome framework, the outcome variable Y is given a definition that is based on potential outcomes associated with the causal effect of interest. Accordingly, y_i^1 is the potential outcome in the treatment state (a charter school) for individual i, and y_i^0 is the potential outcome in the control state (a regular public school) for individual i. The individual-level causal effect of the treatment is then defined as

$$\delta_i = y_i^1 - y_i^0, \tag{5}$$

which is the causal effect of charter schooling instead of regular schooling for each individual i.

The variables Y^1 and Y^0 are then population-level potential outcome random variables, and the average treatment effect (ATE) in the population is

$$E[\delta] = E[Y^1 - Y^0], \tag{6}$$

where $E[.]$ is the expectation operator from probability theory. The observed outcome

variable Y is defined as

$$Y = CY^1 + (1-C)Y^0. \tag{7}$$

Thus, the observed values for the variable Y are $y_i = y_i^1$ for individuals with $c_i = 1$ and $y_i = y_i^0$ for individuals with $c_i = 0$.

Although quite simple, this notational shift changed perspectives and allowed for the development of new techniques, as we will discuss later. It also made clear (to social scientists who may have forgotten) that causal effects exist independent of regression models and can be expressed without relying on regression-based language. With this notation, causal effects could be defined over any subset of the population. Two particular average causal effects of interest became common to investigate. The average treatment effect for the treated (ATT) is

$$E[\delta|C=1] = E[Y^1 - Y^0|C=1] \tag{8}$$

while the average treatment effect for the controls (ATC) is

$$E[\delta|C=0] = E[Y^1 - Y^0|C=0]. \tag{9}$$

For the charter school example, these are, respectively, the average effect of charter schooling for those who attend charter schools and for those who attend regular public schools. If there is reason to expect that these average causal effects do not equal each other, then this is sufficient *a priori* grounds to expect meaningful heterogeneity of individual-level causal effects. In this case, a simple regression model, as in Equation 4, that attempts to develop an estimate of a single charter school effect would be misleading, since it would mask the difference in the expected effect for those who typically do and do not attend charter schools.

More generally, individual-level heterogeneity exists if $\delta_i \neq E[\delta]$ for all individ-

uals i. Variability of δ_i across individuals could be produced by an observed variable or by an unobserved variable, and such variability is particularly consequential when the other variables that produce it are not independent of the cause. Heterogeneity of this type, for example, is present if the average causal effects in Equations 8 and 9 do not equal each other. Many other forms of heterogeneity are of interest as well and can be masked by standard forms of data analysis (see Angrist and Pischke 2009; Elwert and Winship 2010, and Morgan and Winship 2007, esp. Chapter 5).

The third major advancement that has allowed scholarship to move beyond simple regression models is the elaboration of a new form of graph-based causal modeling. Here, the contribution is in enabling new methodological insight and in providing new levels of clarity to researchers. Since this perspective is less familiar to social scientists, and often both misunderstood and underappreciated, we present it in considerable detail in the next section.

A New Methodology of Causal Graphs

Since the 1990s, the rationale for graphical depictions of causal relationships has been strengthened by scholars working at the margins of the social sciences. Judea Pearl's 2000 book, *Causality: Models, Reasoning, and Inference*, has been the most important contribution, although much of the work within it was developed in collaboration with others in the prior decade (see Pearl 2000, 2009). Morgan and Winship (2007, Chapter 3) provide an introduction to the directed acyclic graphs (DAGs) that Pearl and his colleagues are credited with developing, and we will present here only the essential points necessary to demonstrate how variability and context can be incorporated into current graphical methods for causal analysis.[6]

---

[6]We will not discuss the new literature in epidemiology on the distinction between an interaction and an effect modification (see VanderWeele 2009 and VanderWeele and Robins 2007). This literature uses causal graphs and associated structural equations to clarify many subtle points, which we will note briefly later when discussing contextual effects.

The causal graph in Figure 1 is not only a traditional linear additive path model, it can also be understood as a fully non-parametric causal graph. Moreover, it also has a foundation in structural equations, though these are written in quite different form than in the older path model tradition. For the two endogenous variables C and Y in Figure 1, the path model's structural equations would have been

$$
\begin{align}
C &= a_C + b_{P \to C}P + e_C, \tag{10}\\
Y &= a_Y + b_{C \to Y}C + b_{P \to Y}P + b_{N \to Y}N + e_Y. \tag{11}
\end{align}
$$

When seen as a causal graph as developed in the more recent literature, the structural equations for Figure 1 would have a more flexible form. They would be written for all variables in the graph and in unrestricted (possibly nonlinear and interactive) form as

$$
\begin{align}
P &= f_P(e_P), \tag{12}\\
C &= f_C(P, e_C), \tag{13}\\
N &= f_N(e_N), \tag{14}\\
Y &= f_Y(P, C, N, e_Y). \tag{15}
\end{align}
$$

Reading from left to right in the causal graph in Figure 1 and top to bottom in these equations, P is defined as an unspecified function, $f_P(.)$, with e_P as its sole argument. The function $f_P(.)$ is often labeled a "kernel" (which can be confusing because the word kernel is also used in other ways in statistics and mathematics). For our purposes, the input e_P represents all causes of P that are external to the causal model in the sense that they are completely independent of C, N, and Y.[7] This

---

[7]There is considerable debate over how to interpret these external causes. Their existence implies to some scholars that causality is fundamentally probabilistic. However, Pearl (2009) would maintain that the variables embedded in e_P are simply implicit causes of P. Under this interpretation, causality can still be considered a structural, deterministic relation. Freedman (2010, esp. Chapter 15) discusses some of the drawbacks for statistical inference of assuming determinism of this form. Although convincing in some sense, his critique does not alter the utility of these sorts of causal

idiosyncratic determinant of P is suppressed in the causal graph for simplicity, since it could be included by simply adding the causal relationship $e_P \to P$ (as would have been common in the path model tradition). The next three equations then represent analogous unrestricted functions with inputs that represent all causes of the variables on the left-hand sides of the equations. The last is the most elaborate, in that Y is a function of three observed variables P, C, and N, as well as e_Y, all of which transmit their effects on Y via the function $f_Y(.)$.

In this tradition, causal identification can be considered (and, indeed, is often best understood) without introducing any functional form for $f_P(.)$, $f_C(.)$, $f_N(.)$, or $f_Y(.)$. A value is produced for the outcome variable of each equation for every combination of the values in the corresponding function on the right-hand sides of these equations. For example, Y takes on a distinct value for each combination of values for $P = p$, $C = c$, and $N = n$ (typically then with the assumption that values of e_Y are drawn at random from some common distribution that is presumed to have finite moments; see Freedman [2010, Chapter 15] for discussion of alternative approaches).

An implication of this flexibility deserves particular emphasis, and preexisting knowledge of path models can hinder full recognition of its importance. *All* interactions between the effects of P, C, and N on Y are implicitly permitted by the lack of restrictions placed on $f_Y(.)$. Most importantly, this also means that the causal graph in Figure 1 is consistent with all such interactions, since the arrows merely signal which causes of Y in the graph belong in the function $f_Y(.)$. Thus, no new arrows are needed to represent interactions for more specific parameterizations where, for example, the effect of C on Y varies with the level of P or N.

As a result, even though it may feel natural to want to "see" a specific arrow present in the causal graph to represent an interaction effect that corresponds to

---

graphs for clarifying when causal effects are identified.

a cross-product term in a regression equation, one must learn to suppress such a desire. The key point, in considering an analysis that utilizes causal graphs, is to drop regression models from one's mind when thinking about identification issues. Instead, if one must use a data analytic machine to conceptualize how to perform an appropriate empirical analysis of the puzzle under consideration, one should default to simple tabular stratification. In this case, one should think of a sufficiently large sample, such that one could, for example, estimate with great precision the value of Y for every conceivable combination of values for $P = p$, $C = c$, and $N = n$. Average causal effects can then be calculated by appropriately weighting differences calculated within such a stratification of the data.

Representing Variability and Contextual Effects in Causal Graphs

Although there are tremendous advantages that accrue from the general nonparametric structure of causal graphs, it can still be hard to encode heterogeneity in causal graphs in transparent ways for social scientists. Moreover, many scholars who work with causal graphs but who are not social scientists (including those who have developed the case for their general applicability to all causal analysis) do not fully understand how social scientists think about heterogeneity, especially when produced by an interaction with an unobserved variable. To promote understanding by making the key conceptual linkages, we start with a model that is simpler even than the one in Figure 1. We will then add complexity to build an explicit model for a full pattern of heterogeneity that represents variability of effects that emerge in consequential contexts, using the focal charter schools example.

Two Separate Causal Graphs for Two Latent Classes

Consider the two causal graphs in Figure 2, and suppose that the population is

partitioned into two latent classes, each of which has its own graph in panel (a) or panel (b). Suppose again that P is a family's parental background, C is charter school attendance, and Y is a standardized test for all sixth graders in a large metropolitan school district. For these two graphs, the subscripts refer to the latent classes, which are also indicated by a latent class membership variable G that takes on values of 1 and 2.[8]

[INSERT FIGURE 2 ABOUT HERE]

Although surely a gross oversimplification, suppose nonetheless that the population is composed of sixth graders who have been raised in two types of families. Families with $G = 1$ choose schools predominantly for lifestyle reasons, such as proximity to their extended families and tastes for particular school cultures, assuming that all schools are similar in instructional impact because achievement is largely a function of individual effort. Families with $G = 2$ choose elementary schools for their children by selecting the school, subject to constraints, that they feel will maximize the achievement of their children, assuming that schools differ in quality and that their children may learn more in some schools than in others. Accordingly, they are attentive to the national press on educational policy, in which both the Bush and Obama administrations argued for increasing the number of charter schools in the country because some researchers had argued that charter schools are more effective. As a consequence, the second group of families is more likely to send their children to charter schools, such that the mean of C is higher for those families with $G = 2$ than $G = 1$.

Finally, suppose that parents with college degrees are more likely to value distinctive forms of education, and as a result are more likely to send their children to charter schools (independent of whether or not highly educated parents are more likely to be found in the latent class for whom $G = 2$, which we will discuss later).

---

[8] The lower case values x, d, and y for the two causal graphs are meant to connote that these are realized values of X, D, and Y that may differ in their distributions across the two latent classes.

17

They are also more likely to be able to support children in completing homework and otherwise making the most of the educational opportunities that are offered to their children. Accordingly, suppose that in both groups the causal effects $P \to C$ and $P \to Y$ are positive and substantial (i.e., that α_1, β_1, α_2, and β_2 in Figure 2 are positive and substantial).

The question for investigation is whether the effect of C on Y is positive for both groups, and if so, whether it is the same size for both groups. If we are willing to assume, as some of the literature suggests, that the second group of families is correct in the sense that school quality does matter for student learning, and further that charter schools are higher quality (as authors of this chapter, we neither agree nor disagree with this position; see Henig 2008), then we should expect that both δ_1 and δ_2 are more likely positive than not. And, if we believe that parents with $G = 2$ have some sense that this is correct, then not only will more of them send their children to charter schools, they will also sort their children more effectively into charter and noncharter schools. In other words, they will also be more likely to continue to enroll their children in regular public schools if they feel that their children will not benefit from the distinctive characteristics of available charter schools (e.g., if the charter schools that have openings have instructional themes that their children find distasteful). Because both of these self-selection effects are reinforcing, is it likely that $\delta_2 > \delta_1$.[9]

If this plausible scenario is true in reality, what would happen if a researcher ignored the latent classes (either by mistake or, more realistically, because the membership variable G is unobserved) and simply assumed that a single DAG prevailed? In this case, a researcher might estimate the effect of C on Y for each value of P and then average these effects over the distribution of P, yielding a population-level estimate δ. At best, this estimate would be uninformative about the underlying

---

[9] These target parameters, δ_2 and δ_1, are defined implicitly as the average effect of charter schooling for all students from families with $G = 2$ and $G = 1$, respectively.

pattern of heterogeneity that suggests that $\delta_2 > \delta_1$. At worst, this estimate would be completely wrong as an estimate of the average causal effect of C on Y. For example, if P predicts latent class membership G, and G predicts the size of the effect of C on Y, then P-stratum-specific effects mix together individual-level causal effects that vary with the conditional distribution of G within the strata of P. Combining P-stratum-specific effects by calculating an average effect across only the distribution of P does not properly weight the G-stratum-specific effects that are embedded in differential patterns within the strata of P.

In order to consider these possibilities, we need to have a model of selection into C that is informed by a model of the traits of individuals that would cause them to be found in underlying latent classes. It is most natural to pursue such a model in a single causal graph that explicitly represents the latent classes by including the variable G as a node within it.

A Single Causal Graph for Two Latent Classes

Consider Figure 3, which contains a standard triangular system where C has an effect on Y and where both C and Y share a common cause P. To this standard triangle, the latent class membership variable G is introduced as an explicit cause of C.[10] The variable G is given a hollow node, ∘, to indicate that it is unobserved. The arrow from G to C is present because there are alternative groups of families, coded by the alternative values of the unobserved variable G, that approach differently the decision of whether to send their children to charter schools. As a result, G predicts charter school attendance, C.[11]

The corresponding structural equations for the causal graph in Figure 3 are

_____

[10] See Elwert and Winship (2010) for a related usage of a latent class variable G to represent effect heterogeneity.

[11] Although we will continue to write as if G only takes on two values that identify two latent classes, this restriction is no longer necessary. G may take on as many values as there are alternative groups of families who approach differently the decision of whether to send their children to a charter school.

then

$$
\begin{aligned}
P &= f_P(e_P), & (16)\\
G &= f_G(e_G), & (17)\\
C &= f_C(P, G, e_C), & (18)\\
Y &= f_Y(P, C, e_Y). & (19)
\end{aligned}
$$

The latent class membership variable G only enters these structural equations in two places, on its own in Equation 17 and then as an input to $f_C(.)$ in Equation 18.

[INSERT FIGURE 3 ABOUT HERE]

To accept Figure 3 as a full representation of the true causal model that relates G to P, C, and Y, we must be able to assume that G shares no common causes with P, C, or Y that have been mistakenly suppressed. For our focal example, the necessary assumptions are that students who have parents with high levels of education are no more likely to know of the educational policy dialogue that claims that charter schools have advantages and also are no more likely to think that school quality has any effects on achievement when student effort and skill are held constant. We must also be willing to assume that, within values of P and C, G has no causal effect on Y, which with our example is tantamount to assuming that those who attempt to maximize the learning of their children by selecting optimal schools (a) do not manage to do so well enough so that the obtained effect is any larger on average, conditional on other factors, for their own children than for those who do not attempt to select on the causal effect of schooling and (b) do not do anything else that helps their children to benefit from the learning opportunities provided to them in school or in the home that is not already captured by the direct effect $P \rightarrow Y$. This would require that the impulse to select into charter schools based on beliefs about the size of the charter school effect for one's own child is a completely ignorable process, since it does not result in any actual selection on the variation in the causal effect nor

generate any reinforcing behavior that might complement the charter school effect. For the charter school effect, there is no literature to support such a dismissal of the power of self-selection.

Accordingly, for Figures 4(a) and 4(b), we add an arrow from G to Y to the graph presented earlier in Figure 3, which brings these two new graphs into alignment with the discussion earlier of the heterogeneity in the causal graphs in Figure 2. For Figure 4(a), which includes only this one additional arrow, the structural equations are then

$$
\begin{aligned}
P &= f_P(e_P), & (20)\\
G &= f_G(e_G), & (21)\\
C &= f_C(P, G, e_C), & (22)\\
Y &= f_Y(P, C, G, e_Y). & (23)
\end{aligned}
$$

[INSERT FIGURE 4 ABOUT HERE]

For Figure 4(b), we drop the assumption that G is independent of P. This elaborated graph now includes an arrow from P to G. As a result, $f_G(e_G)$ is no longer the appropriate function for G. Equation 21 must be replaced by

$$
G = f_G(P, e_G) \tag{24}
$$

so that family background is an explicit cause of latent class membership. It is likely that parents with high socioeconomic status are more likely to select on the possible causal effect of charter schooling, which is how the latent classes were discussed for Figure 2. Still, how these latent classes emerge is not sufficiently transparent in Figure 4. A more explicit causal model that gives structure to the causal pathway from P to G may help to clarify the self-selection dynamic, as we show next.

Self-Selection into the Latent Classes

Suppose that latent class membership G is determined by a variable that measures a family's subjective expectation of their child's likely benefit from attending a charter school instead of a regular public school. Although we could enter this variable into a causal graph with a single letter, such as S or E, for Figure 5 we use a full mnemonic representation as a variable labeled $\text{Exp}(C \rightarrow Y)$. For Figure 5(a), which is a direct analog to Figure 4(a), this subjective expectation is the sole determinant of G. The structural equations are then

$$
\begin{aligned}
P &= f_P(e_P), & (25) \\
\text{Exp}(C \rightarrow Y) &= f_{\text{Exp}}(e_{\text{Exp}}), & (26) \\
G &= f_G[\text{Exp}(C \rightarrow Y), e_G], & (27) \\
C &= f_C(P, G, e_C), & (28) \\
Y &= f_Y(P, C, G, e_Y). & (29)
\end{aligned}
$$

Note that $\text{Exp}(C \rightarrow Y)$ is determined solely by e_{Exp} in Equation 26. Thus, Figure 5(a) would be an accurate representation of the system of causal relationships if subjective expectations were either completely random or instead based solely on characteristics of families that are independent of the family background variables in P.

[INSERT FIGURE 5 ABOUT HERE]

Given what we have written in the last section about the likelihood that families with different patterns of P will end up in different latent classes represented by G, it seems clear that Figure 5(a) is not the proper representation for the research scenario we have already specified. Accordingly, in Figure 5(b), e_{Exp} is joined by an unspecified additional input I into the subjective expectation of the child-specific causal effect of C on Y, which is then presumed to be caused, in part, by family background. As a result, there is now a path from P to G through I and $\text{Exp}(C \rightarrow Y)$. The structural equations are now augmented as

$$
\begin{align}
P &= f_P(e_P), \tag{30}\\
I &= f_I(P, e_I) \tag{31}\\
\mathrm{Exp}(C \to Y) &= f_{\mathrm{Exp}}(I, e_{\mathrm{Exp}}), \tag{32}\\
G &= f_G[\mathrm{Exp}(C \to Y), e_G], \tag{33}\\
C &= f_C(P, G, e_C), \tag{34}\\
Y &= f_Y(P, C, G, e_Y). \tag{35}
\end{align}
$$

In sociology, the causal effect of P on $\mathrm{Exp}(C \to Y)$ via I follows from the position that privileged positions in social structure are occupied by advantaged families. From these positions, individuals acquire information I that allows them to recognize benefits that are available to them.[12]

By elaborating the causal graph progressively from Figure 3 through Figure 5(b), we have explicitly elaborated what is often presumed in models that incorporate self-selection. Background variables in P are related to the cause C by way of a set of latent classes in G that encode subjective evaluations of the individual-specific causal effect of C on Y. These expectations are functions of characteristics in P by way of the information I that is differentially available to families that differ on P. Yet, even though we now have an elaborate representation of self-selection, we still have not brought contextual effects into the model.

Self-Selection into the Treatment and a Complementary Context

How hard is the task of allowing for contextual effects? With causal graphs, it is considerably easier than one might expect. Consider Figure 6, which incorporates a contextual variable N into the causal graph in Figure 5(b). The variable N represents all causes of student achievement that can be conceptualized as either features of a

---

[12] In addition, it may be that there are also additional common causes of P and I, which would then require that a double-headed arrow between P and I be added to the graph. This would be reasonable if informational advantages that structure expectations for optimal school choices are determined by deeper structural factors that also confer socioeconomic advantages on parents before they arrive at the decision point of whether or not to send their children to charter schools.

student's residential neighborhood or features of a charter school's surrounding neighborhood. The component variables in N might be access to resources not measured by P or specific local cultures that may or may not promote student achievement.[13] With the addition of N, the function for Y is now $f_Y(P, C, G, N, e_Y)$. Recall, again, that N is not restricted by any functional form assumption for $f_Y(.)$. As a result, the causal effect of N can modify or interact with the partial effects of G, C, and P on Y.[14]

[INSERT FIGURE 6 ABOUT HERE]

Figure 6 also allows for even more powerful effects of self-selection. Suppose that self-selection into the latent classes in G is associated with self-selection into N as well. We see two separate and countervailing tendencies. Parents attuned to the potential benefits of charter schooling are also more likely to choose neighborhood contexts that best allow them to encourage their children to study hard in school. At the same time, after obtaining an attendance offer from a charter school, a family may also decide to move to an alternative neighborhood in the catchment area of a suboptimal regular public school, since attendance at such a school may no longer be a consideration in the family's residential decision. If either effect is present, then the function for N is equal to $f_N(G, e_N)$, and we then have seven structural equations as

---

[13]If the latter are only diffuse cultural understandings that only weakly shape local norms about the appropriateness of enacting the role of achievement-oriented student, then such variables may be difficult to observe. In this case, N might then be coded as a series of neighborhood dummy identifier variables. Analysis of these effects would then only be possible if there were sufficient numbers of students to analyze from within each neighborhood studied. Without such variation, the potential effects of N could not be separated from individual characteristics of students and their families. And, if modeled in this way, only the total effects of N would be identified, since the dummy variables for N would not contain any information on the underlying explanatory factors that structure the neighborhood effects that they identify.

[14]See VanderWeele (2009) for an incisive analysis of the difference between an interaction and an effect modification. Our interest, conceptually at least, is in instances of genuine causal interaction, although much of what we write would hold under simpler structures of only effect modification.

$$
\begin{align}
P &= f_P(e_P), \tag{36}\\
I &= f_I(P, e_I) \tag{37}\\
\mathrm{Exp}(C \to Y) &= f_{\mathrm{Exp}}(I, e_{\mathrm{Exp}}), \tag{38}\\
G &= f_G[\mathrm{Exp}(C \to Y), e_G], \tag{39}\\
C &= f_C(P, G, e_C), \tag{40}\\
N &= f_N(G, e_N), \tag{41}\\
Y &= f_Y(P, C, G, N, e_Y). \tag{42}
\end{align}
$$

With this set of structural equations, the nonparametric nature of the kernels allows for fully interactive effects. Again, the function for Y, $f_Y(P, C, G, N, e_Y)$, allows for the effects of C and N to vary within each distinct combination of values between them, as would be the case if the charter school effect varied based on the neighborhood within which students lived.[15]

Empirical Strategies for Modeling Variability and Context in Causal Analysis

There are two basic goals of writing down a causal graph: (1) to represent the set of causal relationships implied by the available state of knowledge, (2) to assess the feasibility of alternative estimation strategies. Figure 6 represents a causal graph that is a reasonable representation of the causal structure that generates the charter school effect. This is a matter of judgment, and one might contend, for example, that the claim that self-selection on the charter school effect generates movement between neighborhoods is overly complex. If so, then $G \to N$ could be removed from the graph, which would simplify the function in Equation 41 to $f_N(e_N)$ and the

---

[15]We should also note that we could enrich the causal model further by drawing from the literature that posits deeper causal narratives for the joint determination of P and N, as well as other causal pathway that link P to N. This additional detail would not change the analysis of the conditioning strategies that follows in the next section.

estimation procedures required.

Suppose that one has access to observational data, such as the National Assessment of Education Progress (NAEP) data analyzed by Lubienski and Lubienski (2003), that provide data on standardized test scores, school type, and some family background characteristics. For the sake of our exposition, suppose further that these NAEP data had even better measures of family background and neighborhood characteristics, so that we could conclude that high quality data are available for all of the variables in Figure 6 with solid nodes: Y, C, N, and P. Yet, no data are available for the variables with hollow nodes: I, $\text{Exp}(C \rightarrow Y)$, and G. The primary goal of analysis is to estimate the average causal effect of C on Y, as this effect interacts with the complementary causal effect of N on Y. Can one adjust for confounding in order to estimate these effects?

We answer this question by drawing on the identification rule developed by Pearl (2000, 2009), which he labels the "back-door criterion." Many explications of the back-door criterion are available, and we draw directly from our own presentation in Morgan and Winship (2007, Section 3.1.3), which contains additional explanation beyond what we have space to offer here.

For Pearl, a back-door path is any sequence of arrows between a causal variable and an outcome variable that includes an arrow that points to the causal variable. Pearl's back-door criterion states that if one or more back-door paths connects the causal variable to the outcome variable, the causal effect is identified by conditioning on a set of variables W if all back-door paths between the causal variable and the outcome variable are blocked after conditioning on W. Pearl proves that all back-door paths are blocked by W if each back-door path:

1. contains a chain of mediation $A \rightarrow M \rightarrow B$ where the middle variable M is in W, or

2. contains a fork of mutual dependence $A \leftarrow M \rightarrow B$ where the middle variable M is in W, or

26

3. contains an inverted fork of mutual causation $A \to M \leftarrow B$ where the middle variable M and all of M's descendants are *not* in W.[16]

Consider now how this back-door criterion applies to the causal graph in Figure 6. The relevant back-door paths are

For C :
1. $C \leftarrow P \to Y$,
2. $C \leftarrow P \to I \to \text{Exp}(C \to Y) \to G \to Y$
3. $C \leftarrow P \to I \to \text{Exp}(C \to Y) \to G \to N \to Y$,
4. $C \leftarrow G \to N \to Y$,
5. $C \leftarrow G \to Y$.

and

For N :
6. $N \leftarrow G \to Y$,
7. $N \leftarrow G \to C \to Y$,
8. $N \leftarrow G \leftarrow \text{Exp}(C \to Y) \leftarrow I \leftarrow P \to Y$,
9. $N \leftarrow G \leftarrow \text{Exp}(C \to Y) \leftarrow I \leftarrow P \to C \to Y$,
10. $N \leftarrow G \leftarrow \text{Exp}(C \to Y) \leftarrow I \leftarrow P \to C \leftarrow G \to Y$.

How many of these paths can be blocked by conditioning on the observed data? For models that estimate the effect of C on Y, the paths 1 through 4 can be blocked by conditioning on P and N. However, path 5 remains unblocked. Likewise, the paths 7 through 10 can be blocked by conditioning on P and C, but path 6 remains unblocked.

For the two unblocked paths, the same problematic arrow is present $G \to Y$,

---

[16]Pearl (2000, 2009) expresses his back-door criterion in a more extended way, using the concept of *d*-separation. This more elaborate explanation allows for the development of additional specificity that can be important. For example, the variables in the permissible conditioning set W cannot include any variables that are affected by the cause of primary interest (i.e., "descendants of C"). This point is irrelevant to our discussion here, since the only variable that descends from C in our graphs is the outcome variable Y. In other scenarios, where multiple outcomes of the cause are included in the graph, this additional restriction on the permissible set of conditioning variables in W can become very important. For a full discussion of the back-door criterion, we recommend a careful reading of Pearl (2009).

which carries forward exogenous causal determinants of I, $\text{Exp}(C \to Y)$, and G (i.e., e_I, e_{Exp}, and e_G) and which are not fully blocked by conditioning on P and N. Thus, if there is selection on the causal effect itself, independent of family background, then it enters the model through G and confounds the conditional association between C and Y. As a result, effective estimation of the effects of C and N on Y is impossible given the available data. Selection into charter schools and neighborhoods is partly on an unobserved variable (the concatenation of $I \to \text{Exp}(C \to Y) \to G$), and this confounding cannot be eliminated by conditioning on the observed data.[17]

How can analysis proceed under these circumstances? There are two main choices. First, the analyst can concede that self-selection on the causal effect is present, which may even generate neighborhood-based selection as a byproduct. In these circumstances, the presence of the causal relationship $G \to Y$ renders estimates of C and N on Y unidentified, either in interactive fashion or when averaging one over the other. In this case, analysis must then be scaled back, and we would recommend that set identification results be pursued (see Manski 1995, 2003). The new goal would be to estimate an interval within which the average causal effect must fall.

In the actual empirical literature on the charter school effect, this humble option has not been pursued by any of the main combatants in the debates. The desire to provide point estimates of causal effects has been too strong, even though it would seem clear to many outside readers that the debate persists simply because the point-estimate of the average causal effect is unidentified.

Instead, researchers have used the lottery nature of charter school enrollments in order to define the charter school effect in a different way: the effect of charter schooling among those who would self-select into it. This is a type of bounded analysis, and it is entirely appropriate. The problem is that these same scholars too

---

[17]This conclusion is hardly revelatory for readers who already know the literature on self-selection bias. Nonetheless, we would argue that there is substantial didactic and communicative value in seeing this result expressed with a causal graph to which the back-door criterion is then applied.

quickly forget the bounded nature of their conclusions, issuing overly broad claims. For example, in their study of charter schools in New York City, Hoxby, Murarka, and Kang (2009:vii) introduce their results in their Executive Summary with three bullet points:

- Lottery-based analysis of charter schools' effects on achievement is, by far, the most reliable method of evaluation. It is the only method that reliably eliminates "selection biases" which occur if students who apply to charter schools are more disadvantaged, more motivated, or different in any other way than students who do not apply.

- On average, a student who attended a charter school for all of grades kindergarten through eight would close about 86 percent of the "Scarsdale-Harlem achievement gap" in math and 66 percent of the achievement gap in English. A student who attended fewer grades would improve by a commensurately smaller amount.

- On average, a lotteried-out student who stayed in the traditional public schools for all of grades kindergarten through eight would stay on grade level but would not close the "Scarsdale-Harlem achievement gap" by much. However, the lotteried-out students' performance does improve and is better than the norm in the U.S. where, as a rule, disadvantaged students fall further behind as they age.

Nowhere in their Executive Summary is it stated that these comparisons across lotteried-in and lotteried-out students are only informative about those who self-select into charter schools. It is not conceded that the results are uninformative about the first order question: What is the expected charter school effect for a randomly chosen student from New York City?

How else can analysis proceed even though back-door conditioning is infeasible? The second choice is simply to assume away the unblocked paths that include $G \rightarrow Y$, which is tantamount to assuming that self-selection does not exist. The study by the Center for Research on Education Outcomes (CREDO 2009) is closer to this position, and they offer a complex set of conclusions based on national results where charter school students are matched to students from traditional public schools:

In our nationally pooled sample, two subgroups fare better in charters than in the traditional system: students in poverty and ELL [English Language Learner] students. ... These findings are particularly heartening for the charter advocates who target the most challenging educational populations or strive to improve education options in the most difficult communities. Charter schools that are organized around a mission to teach the most economically disadvantaged students in particular seem to have developed expertise in serving these communities. ... The flip-side of this insight should not be ignored either. Students not in poverty and students who are not English language learners on average do notably worse than the same students who remain in the traditional public school system. Additional work is needed to determine the reasons underlying this phenomenon. Perhaps these students are "off-mission" in the schools they attend.(CREDO 2009:7)

These conclusions are offered based on models that match students on observable characteristics, leaving unobserved selection on the causal effect unaccounted for. In fact, past research on private school effects on achievement suggests that self-selection may be able to account for the pattern of findings reported in the CREDO study. It is possible that students from families who are living in poverty but who make their way into charter schools are fleeing poor alternatives in their own neighborhood schools and, furthermore, have extra amounts of motivation to succeed in school. At the same time, it is likely that students from more advantaged families are more likely to be attending charter schools solely for lifestyle reasons. In fact, they may be trading off academic opportunities in high quality schools that they have found distasteful for medium quality charter schools with peer cultures that are more appealing.

When back-door conditioning does not identify the causal effect, there is one final possibility for analysis. One can attempt to eliminate the self-selection bias problem by locating a natural experiment. This research design has not been utilized in the charter schools debate to attempt to estimate the unconditional average causal effect (i.e., Equation 6), and we offer a cautionary assessment on its prospects in an appendix to this chapter. However, as we explain in the appendix, instrumental variables that encode natural experiments have strong potential to illuminate some

aspects of the estimation challenge, even though they cannot be considered a general solution for estimating all average causal effects of interest.

Important Caveats on the Utility of Causal Graphs

In the last section, we argued that one of the primary benefits of Pearl's causal graphs is that they are nonparametric. As such, they implicitly incorporate nonlinearities and interactions among the variables that are causes of any outcome. In many circumstances, this feature is an enormous advantage because one can state the causal relationships among variables without having to worry about the specification of the functional relationships among the variables. In this way, causal graphs allow for a careful consideration of identification challenges without also requiring that an analyst simultaneously grapple with specification issues.[18]

In this section, we want to temper our enthusiasm for this property of graphs, embedded within a more general discussion of types of heterogeneity. The main message is the following: Precisely because of their flexibility, causal graphs can obscure important distinctions. In some circumstances, whether a relationship is nonlinear and/or a set of variables interact in their effects on an outcome is of considerable theoretical and substantive interest.

We have already discussed a variety of models in which it has been critical to explicitly model the effects of heterogeneity in order to understand how family

---

[18]For example, a researcher can first determine whether a conditioning strategy can identify relevant causal effects by examining a causal graph. If the effects of interest are identified, then the researcher can proceed to consider the alternative specifications that might be used to estimate the effects from the data, considering whether interactions between causes should be explicitly modeled or instead averaged in some meaningful way. If, however, it is clear from the causal graph that the effects of interest cannot be identified, then the issue of how to specify any subsequent empirical models would be far different, since the goal of the analysis would then be to understand in a provisional way how the unidentified causal effects may appear in masked form under alternative fallible models. The goal would be to represent the contingencies in the data, not offer up estimates that warrant causal conclusions.

background and neighborhood characteristics might interact with the effect of enrolling in a charter school. In situations such as these, the fact that it is possible to bury nonlinearities and interactions in a causal graph can become a disadvantage, especially if the science of the problem demands that the particular nature of the heterogeneity be emphasized. For many applications, it would be unwise to move quickly from model identification to estimation without giving due consideration to how one should represent heterogeneity in the model specification that is selected.

In order to make some of these issues explicit, in this section we consider three different types of heterogeneity, which we label compositional heterogeneity, specification-dependent heterogeneity, and fundamental heterogeneity.[19] Each of these types of heterogeneity can be obscured by a causal graph, if such a graph and the estimation strategy that it suggests are not handled with due care.

Compositional Heterogeneity

In our presentation of the charter school example, we did not give any attention to an important source of heterogeneity: charter schools themselves are heterogeneous in the programs they offer to students. Indeed, charter schools are established precisely to cultivate distinctive identities, and there seems to be widespread recognition among all researchers engaged in this area that these differences matter.

In this sense, the two-category variable C in the causal graphs presented earlier is reductive. Research on charter schools must assiduously attend to variation across charter schools, as well as regular public schools that serve as comparative cases, in order to generate sufficiently meaningful results. Causal graphs are powerful precisely because they are compact representations of underlying structural equations, but they pose a danger in suppressing too much detail. If subdividing a causal variable into meaningfully different groups then yields alternative causal effects that depend on

---

[19]These are our own labels. Other scholars have offered their own categorizations of types of heterogeneity, but we have not found a typology that matches our own explanatory goals for this section of the chapter.

the contrast that is selected, then compositional heterogeneity exists and is obscured by the original causal graph.[20]

Specification-Dependent Heterogeneity

In many situations, a "natural" metric for a variable may not be clear. An obvious example is the relationship between education and income. Should we think about earnings in dollars or in the logarithm of dollars? It would make sense to think about earnings in dollars if we thought that an additional year of education increased one's earnings by the same amount independent of one's current education (or income). In fact, a vast literature primarily from economics has argued that an increase in education by one year increases one's earnings by differing amounts depending on one's education, but, roughly at least, increases earnings by the same percentage amount for individuals with differing years of education.[21]

To understand when specification-dependent heterogeneity can arise, and then how it can be eliminated by model respecification, suppose that we have a fully specified causal structural model, including parameter values for all causal variables, that is the same for everyone in the population under consideration. In this situation, there are no fundamental differences in the way that the causal process operates because the same causal model applies to all individuals.

In this case, apparent individual differences may emerge in the effect of a particular causal variable if differences between individuals in the true model are not reflected in a constrained specification chosen by an analyst. This would occur, for example, when (1) a cross-product interaction between two variables appears to fit the data, (2) the true functional relationship between the outcome variable and one

---

[20]In the philosophy of the social sciences, compositional heterogeneity exists when token-level differences exist within a type-level causal analysis. If it is the case that token-level causal accounts are to be privileged, then compositional heterogeneity is easy to solve. One must simply define causal states carefully so that their instantiations are sufficiently homogeneous.

[21]And a more recent literature has suggested that these effects vary substantially between adjacent years. As such, there is no general functional relationship between earnings and years of education, and only piecewise year-by-year comparisons make sense to consider.

of the variables in the cross-product interaction is quadratic, and (3) this variable in the cross-product interaction term was mistakenly parameterized only with a single linear term for its main effect when the empirical specification was submitted for estimation.

Moving away from this abstract description of specification-dependent heterogeneity, consider the charter school example when both compositional and specification-dependent heterogeneity are likely present together. Suppose that for an estimated model the effect of C on Y appears to differ by a students' race, where the effect of C on Y is larger for black and Hispanic students than for white and Asian students. Suppose that this pattern revealed itself during empirical analysis after interactions between C and dummy variables for race were added to the model. This result does not necessarily imply that black and Hispanic students experience their charter schools in any fundamentally different way. Instead, it may be the case that (1) on average the regular schools that black and Hispanic students would have attended are worse than those that white and Asian students would have attended and/or (2) a crucial variable with a nonlinear effect on achievement and with a distribution that differs by race, such as family income, is given only a linear parameterization in the model. Compositional heterogeneity would arise because of the former while specification-dependent heterogeneity would arise because of the latter. If either (1) or (2) were present, the estimates would suggest incorrectly that the charter school experience is more effective for black and Hispanic students.

Fundamental Heterogeneity

Suppose now that we have a causal graph and its associated structural equations, but suppose that the parameter values attached to the variables vary across individuals. Suppose further that no unappreciated compositional or specification-dependent heterogeneity is present.

Why would the parameter values differ across individuals? We will argue in

this section that this can only happen if the causal processes differ fundamentally across individuals, such that the mechanism by which the cause generates its effect differs across individuals.

Suppose again that the effect of charter schools is thought to be larger for black and Hispanic students than for white and Asian students. Suppose that this is still thought to be the case, even though compositional and specification-dependent heterogeneity have been explained away because the model now properly adjusts for quality differences across schools that students of different races would typically attend and all adjustment variables are given sufficiently flexible parameterizations. Fundamental heterogeneity would be present across race if the mechanism that generates the charter school effect differs in meaningful ways across racial groups.

A prominent strand of literature in education research suggests that stereotypes of inherent intellectual inferiority plague black and Hispanic students in regular public schools. It may be that some charter schools are effective at mitigating the effects of these stereotypes for black and Hispanic students, either by directly confronting the stereotypes as completely unfounded or instead by making students feel that they are attending unique schools that shield them from the effects of stereotypes that exist in the broader culture. As a result, for black and Hispanic students, the charter school effect emerges from two sources: (1) the mitigation of stereotype effects and (2) quality differences in instruction. For white and Asian students, the same stereotypes are irrelevant, and they therefore benefit only from the same higher quality instruction that is offered to all students. We would regard any such difference in the causal process as a form of fundamental heterogeneity, since part of the causal mechanism is distinct for subgroups of the population.

One might counter this presentation of fundamental heterogeneity with the argument that it is simply a more virulent form of specification-dependent heterogeneity. The argument would run as follows. A more complete causal graph can

be written down, with its structural equations including structural zeros, such that subgroup-distinct portions of the mechanism switch off and on depending on the individuals' characteristics at the time they are exposed to the cause. For the example just given, stereotype effects would exist nominally for white and Asian students in the causal graph, but the joint distribution of the variables in the mechanism would give them no weight in the production of the charter school effect for white and Asian students.

Technically, this argument is sound. However, much is lost, in our view, when such fundamental differences in the causal processes within a population are explained away so artificially. We would argue that separate causal graphs should be specified instead.[22]

Conclusion

In this chapter, we have explained how a new methodology of causal graphs can be used to represent variability and context in causal analysis in the social sciences. In developing this explanation, we have also provided a brief analysis of why the practice of quantitatively-oriented sociology between the 1960s and the 1980s opened itself up to the critique that its methods could not represent such features of causal systems. We have also discussed some alternative empirical strategies for estimating these effects, demonstrating the limited power that back-door conditioning offers when heterogeneity due to unobserved variables is present. Finally, we have concluded with a cautionary perspective on causal graphs, noting that their simplifying power has the potential to obscure the heterogeneity that they so easily accommodate. Still, we would maintain that their flexibility enables careful and precise consideration of

---

[22]This position is consistent with (1) arguments from philosophy that mechanistic explanations should "bottom out " at a level that is appropriate to practitioners in a field (see Machamer, Darden, and Craver 2000) and (2) Cartwright's perspective that alternative nomological machines should be used to define distinct causal relations (see Cartwright 1999, Ch. 3).

the challenges of causal effect identification, separated in helpful ways from many specification issues that are less fundamental.

Appendix

The Alternative Natural Experiment Approach

Because the natural experiment perspective may be unfamiliar to readers when expressed with the econometric language of instrumental variables, we will build toward the charter schools example from a more basic set of causal graphs. Figure A1 presents a standard causal graph where identification is achieved through an instrumental variable. For simplicity, assume that no data are available for confounders such as P, and thus the analyst is left with no feasible way to even begin to enact a back-door conditioning strategy for the effect of charter schools C on achievement Y. In this situation, a double-headed, dashed, and curved arrow can be used to signify the existence of common unobserved variables that cause both C and Y.

[INSERT FIGURE A1 ABOUT HERE]

Suppose, however, that a variable Z is observed that is a cause of C and that has no effect on Y except through its causal effect on C. This variable Z is an instrumental variable, and it can be thought of as a shock to C that is independent of the confounders represented by the double-headed arrow in $C \leftarrow\!-\!-\!-\!-\!\rightarrow Y$. If such an instrumental variable exists, the classical econometric literature demonstrates that an estimate of the causal effect of C on Y can be obtained by first calculating the association between Z and Y and then dividing by the association between Z and C.

Are there any plausible instrumental variables for charter school attendance? A typical candidate, as one might find in the economics literature, would be the distance between the treatment site and an individual's residence. The justification would be that the location of the treatment site is arbitrary but has an effect on enrollment propensity because of the implicit costs of traveling to the site. For the

charter school example, it is unclear whether such an instrument would have any chance of satisfying the relevant assumptions, and it would depend crucially on the extent to which charter schools are located in arbitrary places. One would tend to assume, in fact, that they are located nearer to students more likely to benefit from charter schooling, both because many have missions to serve disadvantaged students who are thought to benefit most from having a charter school opportunity and because families may then move to neighborhoods that are closer to the charter schools that they elect to attend. It is possible that these problems could be mitigated by conditioning out some other determinants of the location of charter schools within the district and also obtaining family residence data before students entered charter schools.

For the sake of methodological clarity in our presentation, we will use as our example a more convincing but unlikely instrumental variable (in the sense that it has never yet become available and is unlikely to become available). Suppose that in New York City conditional cash transfers are offered to families that send their children to charter schools. Suppose that this program is modeled on New York City's recent Opportunity NYC program, which was justified by the position that families should be given incentives to make decisions that promote their children's futures. Suppose that for the new hypothetical program $2500 in cash is offered each year to families for each child that they enroll in a charter school. Since charter schools do not charge tuition, families can spend the $2500 per child however they see fit.

Suppose further that, because of a budget constraint, cash transfers cannot be extended to all eligible families. For fairness, it is decided that families should be drawn at random from among all families resident in New York City with school-age children. Accordingly, a fixed number of letters is sent out notifying a set of winning families.

It is later determined that ten percent of students in charter schools received

cash transfers. A dataset is then compiled with performance data on all students in the school district, and the cash transfer offer is coded as a variable Z, which is equal to 1 for those who were offered a cash transfer and 0 for those who were not.

A quick analysis of the data shows that many families who received offers of cash transfers turned them down and chose to send their children to regular public schools. Moreover, it is then assumed that at least some of the charter school students who received cash transfers would have attended charter schools anyway, and they were simply lucky to have also received a cash transfer.

Relying on the classical literature in econometrics, this variable Z would typically be considered a valid instrumental variable. It is randomly assigned in the population, and it has a direct causal effect on C because it is an effective incentive for charter school attendance (recall that we have assumed that the data show that Z predicts C). The crucial assumption is then that the entire association between Z and Y is attributable solely to the causal pathway through C.

As we will further discuss in the conclusion to this appendix, this assumption is debatable because the subsidy is cash and, without further restrictions, could be used by families to purchase other goods that have effects on Y. Any such alternative uses of the cash transfer would open up additional causal pathways from Z to Y that are not intercepted by C. For now, however, we will provisionally accept the identification assumptions. In this case, an IV estimator would deliver an estimate of the causal effect of C on Y that would be considered valid by the standards articulated in the classical literature in econometrics.

Would this causal effect then apply to everyone in the population? A more recent literature has provided an answer to this question: No, unless one is willing to make the artificial assumption that the causal effect is constant in the population.[23]

---

[23]This clarity has been provided by Angrist, Imbens, and Rubin (1996), Heckman and Vytlacil (2005), and Manski (2003), both in these pieces an in their prior work. An explanation of this literature, written for non-economists, can be found in Morgan and Winship (2007, Chapter 7).

Instead, in this case Z will be interpretable as the average effect of charter schooling among those who enter charter schools in response to the cash transfer.[24] This effect is known as the local average treatment effect (LATE), or sometimes the complier average causal effect. Figure A2 explains how to interpret estimators of this form.

[INSERT FIGURE A2 ABOUT HERE]

Suppose that the population can be partitioned into two mutually exclusive groups, compliers and non-compliers. Compliers are those who, in theory, would send their children to charter schools if they received the cash transfer but would not send their children to charter schools if they did not receive the cash transfer. Non-compliers include two other groups, "always takers" and "never takers."[25] The former enter charter schools regardless of whether they receive the cash transfer, and the latter do not enter charter schools regardless of whether they receive the cash transfer. Figure A2(a) is the causal graph for compliers, and it is isomorphic with the causal graph in Figure A1 based on the classical instrumental variable literature. Figure A2(b), however, applies only to non-compliers, and here the graph is different because Z does not cause C for non-compliers. Thus, it should be clear from Figure A2 that the estimate for the effect of C on Y that is enabled by Z cannot apply to non-compliers without the introduction of additional assumptions that allow for extrapolation from compliers to non-compliers. The standard assumption in the classical literature is that the causal effect is constant in the population. Under this assumption, the effect is fixed, such that obtaining an estimate for any subset of the population is sufficient to identify the fixed parameter in the population. The

[24]To warrant this interpretation, one additional assumption is required, which is that Z has a monotonic effect on C. In this case, the assumption is likely satisfied, since cash transfers would not create a disincentive to enter charter schools. For the alternative IV that we discussed – distance from the charter school – monotonicity would probably also hold, though its tenability would be less transparent.

[25]This partition holds only under the assumption (see last note) that monotonicity applies so that there are no "defiers." If monotonicity does not hold, then defiers exist, and the IV no longer identifies the LATE.

more recent literature has demonstrated how unreasonable such assumptions are in real-world contexts, and in this case it would seem self-evident that the charter school effect cannot be assumed to be a constant fixed effect in the population. Certainly none of the literature that has evaluated the charter school effect is consistent with this position.

Thus, our overall conclusion is that LATEs, generated by IVs, can be very illuminating. But, they are limited causal parameters because they apply only to subsets of individuals who are exposed to the treatment of interest. Moreover, one cannot even observe who is a complier, since they are a theoretically defined group. All one can recover from the data is the proportion of the population that is composed of compliers.[26] Nonetheless, knowing the size of the LATE and the percentage of the population that it applies to does then allow for some policy guidance, such as offering a lower bound on the total expected benefit for the entire population of introducing the cash transfer program.

Given that a central theme of this chapter is the power of causal graphs to represent complex patterns of heterogeneity, we will conclude by addressing a final question: Can the clarity of Figure A2 be represented in a single causal graph, akin to the move from Figure 2 to Figure 3 in the main body of the chapter? Yes, but readers may not agree that the clarity is preserved. This may be another situation in which separate causal graphs should be maintained.

Figure A3 is a combined representation of Figure A2, which now applies to the full population. The representation of compliance-based heterogeneity is accomplished by augmenting the graph with a latent class dummy variable, *Compliance*, which signifies whether an individual is a complier.[27] *Compliance* interacts with Z

---

[26]See Morgan and Winship (2007, Equation 7.26), which explains for a similar example that the proportion of compliers in the population would be $(1-$ the proportion of cash transfer winners in traditional public schools $-$ the proportion of cash transfer losers in charter schools).

[27]An alternative and more compact graph could be used for Figure A3 (as well as Figure A4 introduced later). Since *Compliance* is unobserved, one could simply declare that it is a member of the set of variables that generate the double-headed arrow in Figure A3 (or as a member of the set

in determining C, and *Compliance* also may have its own non-zero effect on Y (if compliers have average causal effects that differ from others).

[INSERT FIGURE A3 ABOUT HERE]

Now, to make the connection to the fully elaborated Figure 6 presented earlier, consider Figure A4, which includes all of the relevant back-door paths between C and Y that are packed into $C \leftarrow\!--\!-\!\rightarrow Y$ in Figure A3. Note that for simplicity we have replaced the path $I \rightarrow \mathrm{Exp}(C \rightarrow Y) \rightarrow G$ from Figure 6 with the single variable V. This simplification is permissible under the assumption that an implicit error term e_V contains all of the information in the error terms e_I, e_{Exp}, and e_G in the causal graph in Figure 6. The cash transfer instrumental variable is then represented as a variable Z, which has a sole causal effect in the graph on C because we have assumed that the cash transfer does not have other effects on Y. We then add in the additional back-door path from C to Y through their new common cause *Compliance*.

[INSERT FIGURE A4 ABOUT HERE]

With the addition of $C \leftarrow Compliance \rightarrow Y$ to the graph, two sources of confusion may arise for some readers. First, it remains true that we cannot use back-door conditioning to estimate the effect of C on Y because of the unblockable back-door path through $C \leftarrow V \rightarrow Y$ (which includes implicitly the unobserved latent class variable G). However, it is important to remember that the similarly structured back-door path $C \leftarrow Compliance \rightarrow Y$ does not present any problems for an IV estimator because this back-door path does not generate an unblockable path from Z to Y. It only represents an additional unblocked back-door path from C to Y, which is yet another path that IV estimation proceeds in spite of. Second,

of variables in V that will be introduced later for Figure A4). We give *Compliance* its own pair of explicit causal effects on C and Y for clarity, even thought it makes the graph more complex than it needs to be.

43

nothing in the causal graph itself explains why a resulting IV estimator delivers an average causal effect that applies only to compliers. To understand this point, it is more helpful to deconstruct the graph into two separate causal graphs as in Figure A2. The single causal graph conceals the fact that there is an implicit interaction between Z and $Compliance$. The instrument Z does not cause C for noncompliers, and $Compliance$ does not cause C for those who do not receive the offer a cash transfer. Only the co-occurrence of Z and $Compliance$ switches members of the population from $C = 0$ to $C = 1$. That the single causal graph does not reveal this point clearly might be regarded as a weakness of causal graph representations of underlying structural equations in this context.

Now, to conclude the discussion of the estimation of the charter school effect, consider two final points. It is likely that Figure A4 improperly omits a likely causal effect $P \rightarrow Compliance$. The parental background variable P implicitly includes within it a variable for family income. Students from families with high incomes should be less likely to switch from regular public schools to charter schools because of a \$2500 incentive offered to their parents. Adding such a path, however, would not harm the feasibility of the IV estimator, since it does not generate an unblockable path from Z to Y. In fact, it helps to explain who compliers likely are, since it suggests that they are more likely to be lower income families. In this sense, recognizing the likely presence of that path helps to interpret the LATE that the IV identifies.

But, of course, not all additional causal effects will help to clarify the IV estimator. Suppose, for example, that Z generates an effect on N because the cash transfer is used to pay higher rent in another neighborhood for some families. As a result, a direct path from Z to N is opened up. Conditioning on the observed variable N will block the new problematic pathway $Z \rightarrow N \rightarrow Y$. But, because N is a collider on another pathway, $Z \rightarrow N \leftarrow V \rightarrow Y$, conditioning on N opens up this pathway by inducing a relationship between Z and V. Thus, conditioning

away self-selection into neighborhoods then allows self-selection on the causal effect of charter schooling to confound the IV estimate of the LATE.

Abbott, Andrew D. 1988. "Transcending General Linear Reality." *Sociological Theory* 6:169-86.

—. 2001. *Time Matters: On Theory and Method*. Chicago: University of Chicago Press.

Abdulkadiroglu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag Pathak. 2009. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." National Bureau of Economic Research, Working Paper w15549, Cambridge, Massachusetts.

Angrist, Joshua D., Susan M. Dynarski, Thomas J. Kane, Parag A. Pathak, and Christopher R. Walters. 2010. "Inputs and Impacts in Charter Schools: KIPP Lynn." *American Economic Review: Papers & Proceedings* 100:1-5.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 87:328-36.

Angrist, Joshua David and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Brand, Jennie E. and Yu Xie. 2010. "Who Benefits Most from College?   Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75:273-302.

Bryk, Anthony S., Valerie E. Lee, and Peter B. Holland. 1993. *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.

Cartwright, Nancy. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge, UK New York, NY: Cambridge University Press.

Center, for Research on Education Outcomes. 2009. "Multiple Choice: Charter School Performance in 16 States." Stanford University, Stanford, CA.

Downey, Douglas B., Paul T. von Hippel, and Beckett A. Broh. 2004. "Are Schools the Great Equalizer?   Cognitive Inequality During the Summer Months and the School Year." *American Sociological Review* 69:613-35.

Elwert, Felix and Christopher Winship. 2010. "Effect Heterogeneity and Bias in Main-Effects-Only Regression Models." Pp. 327-36 in *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, edited by R. Dechter, H. Geffner, and J. Y. Halpern. London: College Publications.

Freedman, David. 2010. *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*. New York: Cambridge University Press.

Gelman, Andrew and Jennifer Hill. 2007. *Applied Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Harding, David J., Lisa Gennetian, Christopher Winship, Lisa Sanbonmatsu, and Jeffrey Kling. 2011. ""Unpacking Neighborhood Influences on Education Outcomes: Setting the Stage for Future Research"." Pp. 277-96 in *Social Inequality and Educational Disadvantage*, edited by G. Duncan and R. Murnane. New York: Russell Sage.

Heckman, James J. 2000. "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective." *The Quarterly Journal of Economics* 115:45-97.

Heckman, James J. and Edward Vytlacil.   2005.   "Structural Equations, Treatment Effects, and Econometric Policy Evaluation."   *Econometrica* 73:33-9.

Henig, Jeffrey R. 2008. *Spin Cycle: How Research Is Used in Policy Debates: The Case of Charter Schools*. New York: Russell Sage.

Hoffer, Thomas, Andrew M. Greeley, and James S. Coleman. 1985. "Achievement Growth in Public and Catholic Schools." *Sociology of Education* 58:74-97.

Homans, George C. 1964. "Bringing Men Back In." *American Sociological Review* 29:809-18.

Hoxby, Caroline M., Sonali Murarka, and Jenny Kang. 2009. "How New York City's Charter Schools Affect Achievement." National Bureau of Economic Research, Cambridge, MA.

Jencks, Christopher S. and Susan E. Mayer. 1990. "The Social Consequences of Growing up in a Poor Neighborhood." Pp. 111-86 in *Inner-City Poverty in the United States*, edited by L. E. Lynn and M. G. H. McGeary. Washington, D.C.: National Academy Press.

Lieberson, Stanley. 1985. *Making It Count: The Improvement of Social Research and Theory*. Berkeley: University of California Press.

Lieberson, Stanley and Freda B. Lynn. 2002. "Barking up the Wrong Branch: Scientific Alternatives to the Current Model of Sociological Science." *Annual Review of Sociology* 28:1-19.

Lubienski, Sarah T. and Christopher Lubienski. 2003. "School Sector and Academic Achievement: A Multilevel Analysis of NAEP Mathematics Data." *American Educational Research Journal* 43:651-98.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. "Thinking About Mechanisms." *Philosophy of Science* 67:1-25.

Manski, Charles F. 1995. *Identification Problems in the Social Sciences*. Cambridge: Harvard University Press.

—. 2003. *Partial Identification of Probability Distributions*. New York: Springer.

Mare, Robert D. and Christopher Winship. 1984. "The Paradox of Lessening Racial Inequality and Joblessness among Black Youth: Enrollment, Enlistment, Adn Employment, 1964-1981." *American Sociological Review* 49:39-55.

Morgan, Stephen L. and Jennifer J. Todd. 2008. "A Diagnostic Routine for the Detection of Consequential Heterogeneity of Causal Effects." *Sociological Methodology* 38:231-81.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference (First Edition)*. Cambridge: Cambridge University Press.

—. 2009. *Causality: Models, Reasoning, and Inference (Second Edition)*. Cambridge: Cambridge University Press.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Raudenbush, Stephen W. and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks: Sage Publications.

Rosenbaum, Paul R. 2002. *Observational Studies*. New York: Springer.

Rubin, Donald B. 2005. "Causal Inference Using Potential Outcomes: Design, Modeling, Decisions." *Journal of the American Statistical Association* 100:322-31.

—. 2006. *Matched Sampling for Causal Effects*. New York: Cambridge University Press.

Sampson, Robert J. 2008. "Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure." *American Journal of Sociology* 114:189-231.

VanderWeele, Tyler J. 2009. "On the Distinction between Interaction and Effect Modification." *Epidemiology* 20:863-71.

VanderWeele, Tyler J. and James M. Robins. 2007. "Four Types of Effect Modification: A Classification Based on Directed Acyclic Graphs." *Epidemiology* 18:561-68.
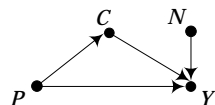
Figure 1. A traditional linear additive path diagram for the effect of parental background (P), charter schools (C), and neighborhoods (N) on test scores (Y).

(a) $G=1$ (b) $G=2$
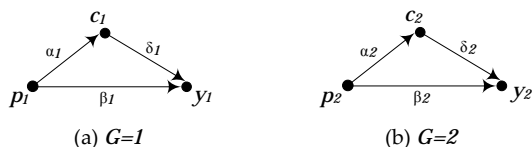
Figure 2. Separate causal graphs for two groups of individuals ($G=1$ and $G=2$) where the effects of parental background (P) and charter schools (C) on test scores (Y) may differ for the two groups.
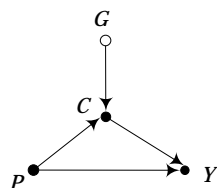

Figure 3. A causal graph where groups are represented by an unobserved latent class variable (G) in a single causal graph.

(a)   (b)
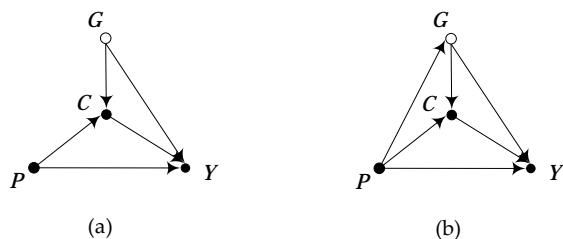
Figure 4. Two causal graphs where selection into charter schools (C) is determined by group (G) and where selection renders the effect of C on Y unidentified as long as G remains unobserved.

(a)   (b)
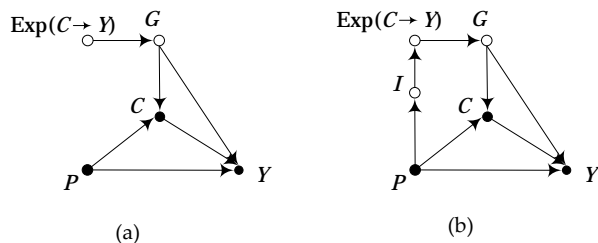
Figure 5. Two causal graphs where selection on the unobservables is given an explicit representation as self-selection on subjective expectations of variation in the causal effect of C on Y. For panel b, these expectations are determined by information (I) that is only available to families with particular parental backgrounds (P).
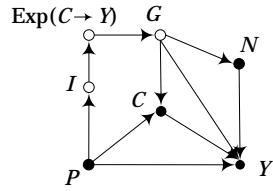
Figure 6. A causal graph where self-selection on the causal effect of charter schooling also triggers self-selection into consequential and interactive neighborhood contexts (N).
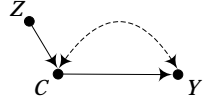

Figure A1. Instrumental variable identification of the causal effect of charter schools (C) on test scores (Y), where Z is the instrument.
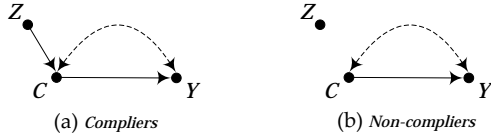

(a) *Compliers*   (b) *Non-compliers*

Figure A2. Instrumental variable identification of the causal effect of charter schools (C) on test scores (Y), where the average effect for compliers is identified in (a) but the average effect for non-compliers is not identified in (b).

Compliance

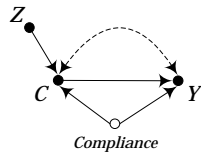Figure A3. A combined causal graph for panels (a) and (b) from Figure A2, where Z is the instrument and compliance is represented as an unobserved latent class variable (*Compliance*).
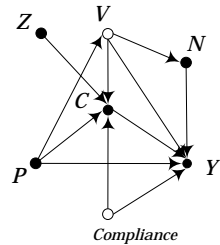

Compliance

Figure A4. Identification of the local average treatment effect using an instrument (Z) for the charter school causal graph presented earlier in Figure 6. The unobserved variable V is a composite for the causal chain that generates self-selection in Figure 6 through information access and selection on the subjective evaluation of the individual-level causal effect.