

REJOINDER: RESPONSE TO SOBEL*

James J. Heckman[†]

“The term ‘cause’ is highly unspecific. It commits us to nothing about the kind of causality involved nor about how the causes operate. Recognizing this should make us more cautious about investing in the quest for universal methods for causal inference.”
Cartwright, 1999, Chapter 5.

Sobel claims to disagree with many of the points made in my paper. He also claims that much if not all of what I say is already in the statistical treatment effect literature. He treats my Section 4 as a literature review rather than an illustration of the basic principles made in Sections 1–3 of the paper, as I intended it to be. In joint work with Edward Vytlačil, I present a comprehensive literature review (Heckman and Vytlačil 2006a,b).

The primary objective of my paper is to present a general and coherent view of causality as it applies to social science. As part of my analysis, I address the approach to causality popularized in statistics by Donald Rubin, Paul Holland and other statisticians. This is an approach to which Sobel subscribes. As my essay documents, the statistical approach suffers from many limitations and in many fundamental respects is a recapitulation of older approaches in econometrics, well understood by economists, that have been enhanced and developed further by contemporary econometricians. I am disappointed that, rather than addressing my arguments, Sobel restates misleading arguments made in the statistics literature. In

*This research was supported by NSF-SES-0241858 and NIH-R01-HDO43411. I thank Jennifer Boobar, Steve Durlauf and Hanna Lee for comments on this rejoinder.

[†]University of Chicago and the American Bar Foundation

responding to Sobel, I am in essence responding to Rubin, Holland and other statisticians whose views are reiterated by Sobel in his commentary.

1. WHAT MY PAPER IS ABOUT

My paper moves discussions of causal inference away from vague philosophical discussions about what “really” constitutes causality to a precise discussion of three prototypical policy problems.¹ In my interpretation, causal models are tools for policy analysis. Different policy problems place different demands on models and data. I articulate the econometric approach that (a) defines the problems of interest precisely; (b) describes the environments, outcomes and choices of the agents being studied precisely and (c) presents conditions on data and models under which the policy problems can be solved. The objective of my paper is not to attack statistics but rather to attack serious policy problems. Sobel attacks the explicit approach developed in econometrics and confuses clearly formulated abstract models for outcomes and selection of outcomes with assumptions made within the context of the explicit models that are maintained in particular applications of the models.

Building on my previous analysis (Heckman 2001), I reconcile the statistical treatment effect literature and the econometrics literature by noting that the wider set of questions addressed by the latter entails considering more ambitious models. Whether the particular assumptions required for identifying a parameter are satisfied is a different problem than the problem of determining conditions under which a question can in principle be answered. I draw on Marschak (1953) and later economists to note that for certain narrowly focused policy questions, it is often possible to get by with much weaker assumptions and data requirements when crafting acceptable answers.

¹Cartwright (2005) provides an illuminating discussion of alternative and often inconsistent uses of the term “causality.”

My essay is about all three policy questions, P1–P3, and not solely about P1. Sobel, however, largely focuses on P1. He briefly touches on P2 and considers a special case of an exogenous conditioning set. His discussion of problem P3 is about extrapolation from factorial experiments instead of a careful discussion of how to forecast new programs with new characteristics as discussed in my paper and in Heckman and Vytlačil (2005, 2006b). His discussion is defensive and does not grapple with the larger aims of my paper.

The careful reader of my paper, Sobel's discussion and the recent literature on causal inference in econometrics and statistics will recognize that Sobel ignores major points that are developed in the econometrics literature and are absent in the statistical treatment effect literature. These are:

1. Development of an explicit framework for outcomes, measurements and the choice of outcomes where the role of unobservables ("missing variables") in creating selection problems and justifying estimators is developed.
2. The analysis of subjective evaluations of outcomes and the use of choice data to infer them.
3. The analysis of *ex ante* and *ex post* realizations and evaluations of treatments. This analysis enables analysts to model and identify regret and anticipation by agents. Developments 2 and 3 introduce human decision making into the treatment effect literature.
4. Development of models for identifying entire distributions of treatment effects (*ex ante* and *ex post*) rather than just the traditional mean parameters focused on by statisticians. These distributions enable analysts to determine the proportion of people who benefit from treatment, something not attempted in the literature Sobel draws on.
5. Development and identification of distributional criteria allowing for analysis of alternative social welfare functions for outcome distributions comparing different treatment states.
6. Models for simultaneous causality relaxing the recursive frameworks adopted by Rubin (1978) and Holland (1986).
7. Definitions of parameters made without appeals to hypothetical experimental manipulations.

8. Demonstration of the need for invariance of parameters with respect to classes of manipulations to answer classes of questions.²

One theme developed in my paper is that major limitations hamper the statistical treatment effect literature in answering important social science questions. These limitations are not surprising since the statistical treatment effect literature is an offshoot of the experimental design literature in biostatistics. My essay shows that “technical” assumptions invoked in the statistical treatment effect literature have unappealing implications for social science.

Two cornerstone assumptions: *SUTVA* and *Strong Ignorability* (*SI*) are especially unappealing. *SUTVA* is a version of an invariance assumption developed in econometrics some 40–50 years ago and formalized in the Hurwicz (1962) paper I cite. In the form advocated by Sobel and many other statisticians, it precludes social interactions and general equilibrium effects, and so precludes the evaluation of large scale social programs. The *SI* assumption, by ruling out any role for unobservables in self selection, justifies matching by assuming away any interesting behavior of the agents being studied. While Sobel criticizes econometrics for making various assumptions, he ignores the fact that the approach that he favors makes implicit assumptions that are stronger and less tenable. The econometric approach is explicit about its assumptions.

Sobel does not acknowledge any intellectual priority for early work by economists that precedes the “Rubin model” as expounded by Holland (1986). Selection models defined over potential outcomes with explicit treatment assignment mechanisms were presented by Gronau (1974) and Heckman (1974, 1976, 1978) in the economics literature. The econometric discrete choice literature (McFadden 1974, 1981) used counterfactual utilities as did its parent literature in mathematical psychology (Thurstone 1927, 1959). Unlike the Rubin model, these models do not start with the experiment as an ideal point of departure, but they start with well-posed, clearly

²This notion is in the early Cowles Commission work. See Marschak (1953) and Koopmans, Rubin, and Leipnik (1950). It is formalized in Hurwicz (1962) as cited in my paper. Rubin’s *SUTVA* is a special case of the invariance condition formalized by Hurwicz.

articulated models for outcome and treatment choice derived from behavioral theory where the unobservables that underlie the selection and evaluation problem are made explicit.

Rubin's 1978 model of treatment choice came later and only implicitly accounts for the unobservables that drive the selection problem. His point of departure is randomization and the analysis of his 1976 and 1978 papers is a dichotomy between randomization (ignorability) and nonrandomization, not an explicit treatment of particular selection mechanisms in the nonrandomized case as developed in the econometrics literature.

Sobel dismisses the value of making clear the assumptions about model unobservables that produce selection and evaluation problems when he dismisses "structural" models. In this regard he follows Angrist, Imbens, and Rubin (1996) and Holland (1986). Sobel equates structural models (economic models) with *LISREL* type models and standard simultaneous equations models despite the greater generality of the structural models (see, e.g., Matzkin 2006).

Structural models do not "make strong assumptions." They make explicit the assumptions required to identify parameters in any particular problem. The treatment effect literature does not make fewer assumptions; it is just much less explicit about its assumptions. Like many statisticians, Sobel prefers to be implicit about many of his assumptions. This approach begs serious questions about the best way to model the severe problems that arise in making sound policy evaluations.

My essay is about:

1. Clearly defining the policy problem being addressed;
2. Asking what parameter is required to answer the problem;
3. Discussing minimal identification conditions; and
4. Analyzing the properties of various estimators.

While Sobel's discussion claims to show that there are dimensions along which the econometric literature is lacking relative to the statistical literature on treatment effects, his arguments are based on misstatements and misunderstandings of the econometrics literature that are prevalent in the statistical treatment effect literature. For example, he makes the claim, like Rubin and many other statisticians, that econometric selection

models depend on normality.³ He claims that economists, and I in particular, “adopted the Rubin model” in the 1980s. This repeats a claim made by Rubin.⁴ Sobel clearly has not read or understood the work published in econometrics in 1974–1976 which presented models of potential outcomes and treatment assignment rules long before Rubin’s 1978 paper.⁵ My 1974–1976 papers are not “informal” and they present precise discussions of potential outcomes (e.g., market and nonmarket wages) and outcome selection mechanisms. The switching regression model of Quandt (1958, 1972) describes a model of potential outcomes and develops various regime (potential outcome) selection rules. Detached readers would be advised to compare the level of formality in these papers with the relative informality of Rubin’s papers, especially his informal 1974 paper which Sobel cites. In that paper, there is no systematic discussion of treatment assignment rules whereas, by 1974, the econometric literature had systematically developed and analyzed such rules.

The early econometric work clearly separates the definition of parameters from their identification in a fashion not found in the statistics literature. Heckman and Robb (1985, 1986) present comprehensive analyses of outcome equations, selection mechanisms and unobservables using economic theory. We had no need to draw on the “Rubin Model” which was a special case of economic models that were formulated prior to Rubin’s work. A more accurate description of Rubin’s contribution is that he explicated aspects of econometric models to statisticians.

2. WHAT IS NEW IN MY PAPER AND NOT DISCUSSED BY SOBEL

Sobel does not discuss my extension of the treatment effect literature to the identification of non-recursive systems. The literature on

³Heckman (1980, 1990), Heckman and Robb (1985, 1986), Heckman and Honoré (1990), Ahn and Powell (1993), and Powell (1994), among many others, have relaxed the normality assumption made in the early 1970’s literature. See Heckman and Vytlačil (2005, 2006b) for a survey. It is far from clear that in practice normality is a poor assumption in many applications. See Heckman (2001).

⁴Rubin (2000).

⁵The Roy model (1951) is a clear predecessor as are the switching models of Quandt (1958, 1972).

causality in statistics is recursive and rules out simultaneous causality. My Section 2.5 is standard econometrics but new statistics that reveals the power of the econometric approach over “frontier” methods in statistics like directed acyclic graphs (see Pearl 2000) that are recursive. Given long-standing interest in social interactions by sociologists (see also Durlauf and Young 2001) it is unfortunate that Sobel dismisses out of hand this contribution of my essay that allows sociologists to define and identify models of social interactions that are ruled out in the Rubin (1978)–Holland (1986, 1988) approach that he espouses.

Sobel misses another main contribution of my paper: to make the literature on causality specific by addressing real problems. Abstract discussions of causality with appeals to philosophy, “closest worlds,” “regularity,” and the like sound profound but in fact are superficial since they have no operational content and do not address policy problems. Most empirical social scientists are not concerned with philosophy *per se*, but instead they want honest answers to clearly stated problems.

My essay is organized around the theme of addressing three policy evaluation questions that arise in everyday practice. Problem P1 is what statisticians focus on. Problems P2 and P3 are new problems that can be answered using the econometric approach. Sobel’s attempt to address P2—the extrapolation problem—demonstrates the *ad hoc* nature of current statistical approaches and the power of the econometric approach. By keeping the unobservables implicit, he disguises an implicit exogeneity assumption for the conditioning variables in the approach he advocates. Heckman and Vytlačil (2005, 2006b) consider extrapolation under more general conditions.

An alternative approach discussed in Heckman and Vytlačil (2005, 2006b) is to model the dependence between the conditioning variables and the unobservables, but since Sobel (like Rubin and Holland) does not like to make unobservables explicit, this route is denied him. He implicitly agrees that the literature in statistics does not address P3 and his discussion of it confuses problem P2 with problem P3.

He offers a discussion of factorial experiments as a substitute for a clear discussion of P3. Since he will not make explicit statements about unobservables, he begs a central question addressed in the econometrics literature of how to predict the effects of new programs never previously observed. His version of P3 is a simple extrapolation

exercise and hence a version of P2. He does not offer a general treatment of P3 as I do in the paper and in Heckman and Vytlačil (2005, 2006b).

Our approach goes well beyond the interpolation and extrapolation advocated by Sobel to consider predictions of programs with features never previously observed.⁶ Making predictions about new programs and policies is an essential task of social science. If these predictions are not made in a cautious, principled, explicit way, plenty of “causal inference experts” stand waiting to fill the vacuum and provide less credible estimates. By discussing hard problems clearly, analysts can pinpoint limits to knowledge that raise the standards of evidence.

In failing to present a serious discussion of problem P3, which is a central focus of my essay, Sobel echoes the conventional statistical approach that ignores this policy evaluation problem. Had he carefully considered my analysis of this problem in this paper and in my work with Vytlačil, his claims about close “agreement” between the econometric approach and the statistical approach would vanish.

Sobel misses another major theme of my essay and the entire econometric evaluation literature: that treatment effects should be defined relative to the problem being analyzed (Marschak 1953; Heckman 2001; Heckman and Vytlačil 2001b, 2006a,b). The econometric literature develops the point that the choice of an estimator cannot be separated from the choice of the question being addressed by the investigator and the *a priori* assumptions made by the investigator.

Thus, although *ACE*, *ITT* or *TT* may be traditional parameters, they do not address many specific policy questions. My work with Vytlačil (2001b, 2005, 2006b) discusses specific policy questions and devises estimators that address them. Manski (2000, 2004), in his interesting work, derives treatment assignment rules for particular loss functions.

3. IS CAUSALITY IN THE MIND?

Sobel, evidently influenced by my exchange with Tukey (in Wainer 1986, reprinted 2000), sharply attacks me for claiming that “causality

⁶His approach is based on arbitrary functional form assumptions whereas we present a general analysis guided by theory.

is in the mind” and that causal knowledge is provisional. Then, throughout the rest of his discussion, he demonstrates the validity of my point by offering a series of unsupported opinions and assertions about what is “reasonable” and what is not. His opinions and value judgements are his expressions of intuitive models in his mind that he never formalizes or makes explicit, but introduces casually using rhetorical devices. Sobel is typical of many statisticians who keep crucial assumptions implicit.

My claim is not intended as a defense of solipsism, post-modernist relativism or the notion that analysts are free to make up any crazy model they like. Instead, I am saying that all scientific activity is predicated on assumptions.

A clearly formulated causal model should (a) define the rules or theories that generate the counterfactuals being studied, including specification of the variables known to the agents being studied as well as the properties of the unobservables of the model where the unobservables are not known to the analyst but may be partly known by the agent; (b) define how a particular counterfactual (or potential outcome) is chosen; (c) make clear the assumptions used to identify the model (or to address the policy questions being considered); and (d) justify the properties of estimators under the maintained assumptions and under alternative assumptions. These are tasks 1 (corresponding to a and b), 2 (corresponding to c) and 3 (corresponding to d) in my Table 1. These ingredients are the hallmark of the selection model as analyzed in Heckman (1974, 1976, 1979), Gronau (1974), Roy (1951), Willis and Rosen (1979) and numerous other papers in the econometrics literature. Understanding the relationship between the unobservables generating choice of treatment and the unobservables generating outcomes is the key to understanding the properties of various evaluation estimators, a point first made in Heckman and Robb (1985, 1986, reprinted 2000).

One will look in vain in the papers of Neyman, Cox, Kempthorne, Rubin or Holland for the specification of precise treatment assignment rules that have been the hallmark of econometric selection models since 1974. Sobel’s claim that Rubin (1978) or any other statistician has systematically developed treatment assignment mechanisms is false. Rubin (1976, 1978) contrasts randomization with nonrandomization and does not develop the structure

of nonrandomized selection rules in the fashion pioneered and developed in econometrics.⁷

The act of defining a model is a purely mental activity. It may draw on preexisting theory (which is itself derived from earlier mental acts), interpretations of data (which involve a mental act using models for the phenomenon being studied and models of statistical inference) and the rules of logic. There is no purely empirical process for discovering or defining causality. All causal knowledge is conditional on maintained assumptions.⁸

Sobel (p. 106) dismisses the conditional nature of causality, writing that “I do not believe that most scientists (or philosophers) would subscribe to this view and were they to do so they would presumably have little further interest in causality.” This claim runs contrary to a large body of thought in philosophy associated with Kant and Hume, among others.

Indeed, “causality” is not a central issue in fields with well-formulated models where it usually emerges as an automatic by-product and not as the main feature of a scientific investigation. Moreover, intuitive notions about causality have been dropped in pursuit of a rigorous physical theory. As I note in my essay with Abbring (2006), Richard Feynman in his work on quantum electrodynamics allowed the future to cause the past in pursuit of a scientifically rigorous model even though it violated “common sense” causal principles. The less clearly developed is a field of inquiry, the more likely is it to rely on vague notions like causality rather than explicitly formulated models.

Most scholars intend to describe the real world with their models. Certainly, in addressing P1–P3, I am referring to real world problems. However, any empirical or theoretical analysis rests on assumptions. The clearer analysts are about these assumptions and the more they are able to test them, the more clearly stated are the sources of agreement or disagreement among analysts. Such clarity determines the next steps in the scientific process of constructing

⁷When Sobel writes that $TT = ACE = TUT$ when treatment decisions do not depend on potential outcomes, he presents a garbled version of a precise result established in my 1974 and 1976 papers.

⁸Indeed all knowledge is conditional in this sense. For example, much of modern mathematics is predicated on the Axiom of Choice.

better models and better data to narrow down the zones of disagreement among analysts.

Although Sobel objects to my assertion about the importance of maintained assumptions and *a priori* beliefs in the causality enterprise, he proceeds, like many statisticians, to pass judgements about what is and is not good practice. Unlike the econometric approach, Sobel adopts the statistical approach which often hides key assumptions by invoking slogans instead of science.

Each of Sobel's judgements is based on maintained assumptions and beliefs "in his mind." They involve his implicit assumptions and value judgements. Statisticians like him who are not explicit do not convey their private thoughts in an objective, publicly interpretable way. Like many statisticians working in the field of causal inference, he is not clear about many of his crucial implicit assumptions. Implicit assumptions are entailed in writing down the arguments of causal relationship (1) (his notation) and characterizing its properties.

Sobel writes "the focus of the statistical literature is primarily on obtaining the best possible estimate of the causal parameter of interest" without defining "best possible," the basis for the choice of the "parameter of interest" or the question being addressed.⁹ In his defense of matching, he appeals to conditional independence assumption (11) (in the notation of his paper) as "straightforward" and that it "readily lends itself to use by empirical investigators." Some statisticians often use the phrase "it works well in practice" without defining "works well." Another common slogan used to justify matching is that "my clients understand it." In these and numerous other instances, Sobel and a large statistical community implicitly appeal to a variety of conventions rather than presenting explicit rigorous models and assumptions. The credo "let sleeping dogs lie" is good for sales, but it is bad for science. Instead of invoking slogans as a solution to problems, the structural approach emphasizes understanding the underlying mechanisms producing outcomes and selection rules.

As an application of the scientific approach, my discussion of matching develops the point that (i) there is no rigorous basis for picking the set or sets of conditioning variables that make the method

⁹The "best possible estimate" is defined precisely in Bayesian and Wald decision theories. See, e.g., the discussion in Manski (2000).

“work.” Heckman and Navarro (2004) show how the conventional model selection rules for picking the conditioning variables W in a set of data can produce badly biased estimates of the average causal effect and other parameters. (ii) The method assumes that the analyst has as much relevant information as the agent being studied (see, e.g., Heckman and Navarro 2004, for a precise definition of “relevant” information). If the agent knows more than the analyst and acts on it, matching breaks down. (iii) The analyst assumes that people at the margin of being attracted into a program are the same (have same outcomes on average) as average participants. Matching is just non-parametric regression analysis. It is more careful than Ordinary Least Squares (*OLS*) in accounting for empirical support problems but it assumes that the conditioning variables that the analyst has at his disposal fortuitously solve selection problems.¹⁰

To take another point, like many statisticians, Sobel resolutely defends randomization. The Rubin (1978)–Holland (1986) papers take as their benchmark randomized trials where treatments are selected by a hypothetical randomization. As I point out in my essay, even under ideal conditions, unaided randomization cannot answer some very basic questions such as what fraction of a population benefits from a program.¹¹ And in practice, contamination and cross over effects make randomization a far from sure-fire solution even for constructing *ATE* or *ACE* (see the evidence on disruption bias and contamination bias arising in randomized trials that is presented in Heckman, LaLonde, and Smith 1999; Heckman, Hohmann, Smith, and Khoo 2000). Sobel makes a series of implicit assumptions about what questions should be answered, the effects of randomization on participants and the like.

Sobel also disagrees with my claim that statisticians conflate the three tasks shown in my Table 1. The analysis of Holland (1986, 1988) is a good illustration of my point. It also illustrates the central

¹⁰In one of his many attributional errors, he echoes Imbens (2004) and credits Barnow, Cain, and Goldberger (1980) with the phrase “selection on observables” to describe matching. The term originates in Heckman and Robb (1985) and is not to be found in Barnow, Cain, and Goldberger.

¹¹See Carneiro, Hansen, and Heckman (2001, 2003), where this parameter is identified using choice data and/or supplementary proxy measures. See also Cunha and Heckman (2006a,b) and Cunha, Heckman, and Navarro (2005, 2006).

role of the randomized trial to the Holland-Rubin analysis. After explicating what he calls the “Rubin model,” Holland gives a very revealing illustration of how the first two tasks of Table 1 are conflated by one leading figure in the statistical treatment effect literature. Holland claims that there can be no causal effect of gender on earnings. Why? Because we cannot randomly assign gender. This confused statement conflates the act of definition of the causal effect (a purely mental act) with empirical difficulties in estimating it (Steps 1 and 2 in my Table 1). This type of reasoning is prevalent in statistics.¹²

As another example of the same point, Rubin (1978, p. 39) denies that it is possible to define a causal effect of sex on intelligence because a randomization cannot *in principle* be performed.¹³ In this and many other passages in the statistics literature, a causal effect is defined by a randomization. Issues of definition and identification are confused. A recent paper shows that this fallacy is alive and well in statistics. A paper by Berk, Li, and Hickman (2005) makes the same error as Rubin and Holland. Sobel is correct in saying that population treatment parameters can be defined abstractly. However, that point was not made in the statistical treatment effect literature. It is made in econometrics.¹⁴

I agree with Sobel that the act of definition is logically separate from the acts of identification and inference. That is a main point of my paper. We both agree that a purely mental act can define a causal effect of gender. That is a separate task from identifying it. What is odd is that he states his agreement with my position and that of the econometrics literature as a disagreement. And he fails to accurately

¹²Parentetically, my title “Scientific Causality” was motivated by Holland’s contrast between models of science that attempt to probe deeply and understand the “causes of effects” and the statistical treatment effect literature. Understanding the causes of effects is an essential activity for prediction and forecasting—problems P2 and P3 in my paper.

¹³“Without treatment definitions that specify actions to be performed on experimental units, we cannot unambiguously discuss causal effects of treatments.” (Rubin 1978, p. 39).

¹⁴The *LATE* parameter of Imbens and Angrist (1994) is defined by an instrument and conflates task 1 and 2 (definition and identification). Heckman and Vytlačil (2001b, 2005, 2006b) define the *LATE* parameter abstractly and separate issues of definition of parameters from issues of identification. Imbens and Angrist (1994) use instrumental variables as surrogates for randomization.

represent a pervasive point of view among statisticians that gives rise to the myth that causality can only be determined by randomization, and that glorifies randomization as the “gold standard” of causal inference.¹⁵

4. THE ROY MODEL, THE SWITCHING MODEL AND THE RUBIN MODEL

Sobel repeats an assertion made by Rubin: that I, and other economists, “started using the Rubin model in the 1980s.”¹⁶ Sobel has clearly not studied the econometrics literature with any care. The “Rubin model” is in fact a version of an econometric model developed by Roy (1951). It is also a version of the switching regression model of Quandt (1958, 1972). That model contains both a framework for potential outcomes (Y_0 , Y_1) and also a choice of treatment rule.¹⁷ There was no explicit discussion of the treatment assignment rule in any of the Rubin papers that Sobel cites or in the statistics literature until very recently.¹⁸

Heckman and Honoré (1990) present a comprehensive analysis of the Roy model. Heckman (1990) and Heckman and Smith (1998) extend it (see also Heckman 2001, and Heckman and Vytlacil 2006a,b). Unlike the statisticians, Pearl (2000) is forthright about his own debt to the economics literature in the distinction between “fixing” and “conditioning,” which is central to his work on causality. See Haavelmo (1943) for the source of Pearl’s “do” operator.¹⁹

¹⁵As noted in my essay, and in Heckman (1992), self selection provides information on agent-subjective evaluations of programs.

¹⁶See Rubin (2000).

¹⁷One cannot find any explicit analysis of treatment selection rules in the statistical literature (Neyman 1923; Rubin 1978; Holland 1986; Rubin 1986) other than the randomized-nonrandomized dichotomy previously discussed.

¹⁸Sobel cites Rosenbaum (2002) for use of such rules. As previously noted, Rubin does develop the dichotomy “randomized vs. nonrandomized.” He does not go deeper, nor does he consider how the form of the treatment assignment rule affects the choice of an appropriate estimator. That point is developed in Heckman and Robb (1985, 1986).

¹⁹Lewis (1963) is an early pioneering analysis of counterfactuals in economics that also considers the problems raised by self selection and general equilibrium effects.

The simplest form of the Roy model has two potential outcomes and a decision rule (treatment assignment rule). In its simplest version, the treatment indicator variable is $D = \mathbf{1}(Y_1 \geq Y_0)$, where $\mathbf{1}(\cdot) = 1$ if the argument is true and is zero otherwise. Thus a doctor might assign treatment on the basis of which therapy has the best outcome. A student may decide to go to college vs. stopping at high school based on which option has the highest income. The Roy model is a version of the competing risks model of biostatistics.²⁰ This model of potential outcomes and treatment selection predates Cox and Rubin, as does the Thurstone (1927) model of counterfactual utilities of choices developed in mathematical psychology.

More general versions of this model developed in econometrics allow agents to be partially informed about (Y_1, Y_0) when they make their decisions and to allow for more general costs. In the generalized Roy model, $D = \mathbf{1}(E(g(Y_1, Y_0, C)|\mathcal{I}) > 0)$ where \mathcal{I} is the agent's information set, C is the cost of moving from "0" to "1" where "0" is the initial state, and g is a general preference function for the agent making the treatment decision. In the original Roy model $C = 0$, $\mathcal{I} = (Y_1, Y_0)$ and $g = (Y_1 - Y_0)$. The general form of this model allows analysts to distinguish objective from subjective evaluations of treatments and *ex ante* and *ex post* versions of both. See Carneiro, Hansen, and Heckman (2001, 2003), Cunha and Heckman (2006a), Cunha, Heckman, and Navarro (2005, 2006), Heckman and Vytlacil (2006b) and Heckman and Navarro (2006) for more general analyses.

In my 1974 paper, Y_1 is the market wage of a woman. Y_0 is her nonmarket wage (her value in home production). Her decision rule is to work ($D = 1$) if the market wage is greater than the reservation wage $D = \mathbf{1}(Y_1 \geq Y_0)$.²¹ Otherwise, she does not work. I also develop a model for hours of work. Willis and Rosen (1979) use this model when Y_1 is college earnings and Y_0 is high school earnings. They allow for costs C . $D = 1$ (a person goes to college) if $Y_1 - Y_0 - C > 0$ ($D = \mathbf{1}(Y_1 - Y_0 - C \geq 0)$).²² There is a huge literature starting in economics long before the "Rubin model" became popularized in

²⁰See Heckman (1987) where this link is established. Versions of the competing risks model go back to early Twentieth Century work by Danish actuaries.

²¹See Gronau (1974) for a closely related model.

²²They assume perfect certainty. See Cunha and Heckman (2006a) for a version of this model with uncertainty as well as additional features.

statistics and it is this literature that influences econometric analyses of causal models.

The “Rubin model” is thus a version of this classical econometric model without explicit specification of the decision model for choice of treatment. Sobel claims that statisticians such as Rosenbaum (2002) are now using the framework developed by economists some 30 years before. If true, this is a welcome acquisition from economics by statistics.

As previously noted, Sobel also claims that Rubin explicitly discusses treatment assignment rules. In fact, Rubin relies on the crutch of randomization to define his models and only vaguely describes other assignment rules as “not randomized.”²³ One needs to look to Gronau (1974) and Heckman (1974, 1976, 1979, 1978) for explicit development of selection models with explicit treatment assignment and selection rules. Had Rubin understood the general selection model, he would not have advocated matching or balancing to overcome nonrandomized assignment as he does in his 1974 and 1978 papers. Heckman and Navarro (2004), Heckman and Vytlačil (2006b), and Heckman, Urzua, and Vytlačil (2006) show the bias that arises from using matching when a general selection rule characterizes treatment choice.

5. COMPARISON OF ESTIMATORS

My analysis in Section 4 was only intended to illustrate the basic point that each evaluation estimator makes assumptions. Sobel misinterprets this section as my attempt to write an exhaustive survey instead of my attempt to illustrate some points from the earlier part of the paper. In Section 4, I focus attention on certain mean treatment effect parameters because of their familiarity and simplicity. Heckman, LaLonde, and Smith (1999) and Heckman and Vytlačil (2006b) present comprehensive surveys of the econometric approach.

Sobel confines his discussion to a few mean treatment parameters, ignoring the range of parameters introduced in the earlier

²³Rubin (1978) discusses the distribution of treatment assignment rules in a general way but never develops their properties in the systemic, formal way it is developed in economics.

sections of my essay. His discussion is selective and he seizes on small points to make objections to my paper. He misses key developments in the econometrics literature that show that in models with heterogeneous responses, *IV* and selection models are closely related (Heckman and Vytlacil 2005; Heckman, Urzua, and Vytlacil 2006).

I use separability in my analysis of selection models in Section 4 only to simplify the exposition. Matzkin (2006) presents a comprehensive discussion of nonparametric identifiability in nonseparable selection (and other) models.

His contrast between matching and control functions on this issue is specious and ignores an entire recent semiparametric literature in econometrics (see Heckman and Vytlacil 2006a; Matzkin 2006). Selection models do not require normality, separability or standard exclusion restrictions in order to be identified.

He takes out of context my claim that control functions are more general than matching. That claim is made under a series of assumptions about the separable selection model and was not intended as a general characterization.

The recent semiparametric literature by Heckman (1980, 1990), Powell (1994) and Carneiro, Hansen, and Heckman (2003) does not rely on normality or functional form assumptions. On this point Sobel inaccurately characterizes the econometrics literature. Even in the absence of distributional assumptions, no exclusions are needed to identify the Roy model, contrary to his claims. In his notation, a model with $Z = X$ can be identified using curvature restrictions without any exclusion of variables. See Heckman and Honoré (1990).²⁴

The key idea underlying the control function approach introduced in Heckman (1980) and in Heckman and Robb (1985, 1986) is to model the relationship between the unobservables in the treatment choice equation and the unobservables in the outcome equations rather than to assume they are independent given a specified set of variables as is done in the matching literature. Sobel inaccurately compares matching and selection estimators in terms of the number of assumptions invoked by each method and not in terms of their strong implications.

²⁴Sobel relies on a flawed survey by Vella (1998) which does not accurately portray the econometric selection literature.

Matching assumes that, on average, the marginal person and average person with the same observed conditioning variables respond the same to treatment ($TT = ATE = TUT$). It assumes that the analyst knows the right conditioning set and uses it. Selection models allow for variables that produce conditional independence invoked in matching to be unobserved by the analyst (see Carneiro, Hansen, and Heckman 2003). Sobel's analysis of IV also ignores the entire body of recent econometric work which establishes what instrumental variables estimate in the general nonseparable case (see Heckman and Vytlacil 1999, 2001a, 2006b). I now turn to that work.

6. THE UNIFYING ROLE OF THE MARGINAL TREATMENT EFFECT

Sobel has evidently not read my 2001 Nobel Lecture or my work with Vytlacil (1999, 2001a,b, 2005, 2006a,b). Had he done so he would not claim that "sociologists will usually be more interested in Treatment on the Treated (TT) or ACE (Average Causal Effect) than the Marginal Treatment Effect (MTE)." In rereading my essay, I now realize it was a mistake for me not to discuss my work with Vytlacil in my paper.

Vytlacil and I establish that the marginal treatment effect (MTE) is a device that unifies the evaluation literature. From knowledge of the MTE , analysts can interpret what IV estimates as well as the commonly used treatment effects, OLS and matching estimators as a different weighted average of the marginal treatment effect. Under the assumptions clearly stated in our papers, we establish that all treatment effects and all estimands (probability limits of IV , matching, OLS , control function estimators) can be expressed as weighted averages of the MTE with known weights, i.e., weights that can be estimated from the sample data. Letting $MTE(x, u)$ be the MTE for a given value of $X = x$ (observables) and $U = u$ (unobservables), we may write the estimand or treatment effect j given x , $\Delta^j(x)$ as

$$\Delta^j(x) = \int_b^a MTE(x, u)\omega^j(x, u)du \quad (1)$$

where $\omega^j(x, u)$ is a weight for estimand or effect j that can be empirically determined and the limits (a, b) are known in any application. We can generate *LATE* as a special case of this formula. When the model has limited support (regions where *MTE* can be identified), the estimator automatically adjusts for it.²⁵

Bounds for the treatment parameters are presented in Heckman and Vytlačil (2006b). Different instruments produce different weights and these weights are generally not the weights required to define the standard treatment effects. Our approach is far more general than the piecemeal type of analysis of what *IV* estimates of the sort presented by Sobel in his comments on the statistical literature. Each of his special cases drops out from our general analysis. The *MTE* approach presents a nonparametric control function analysis where the propensity score plays a conceptually distinct role from the role it plays in matching models (Heckman and Vytlačil 2006b). Our analysis is not to be found in the statistics literature.

Sobel is clearly a fan of the *LATE* approach. Therefore, he has to be a fan of *MTE*. The Imbens-Angrist (1994) *LATE* parameter is a discrete version of the Björklund-Moffitt (1987) marginal gain parameter introduced into the evaluation literature in a selection model framework. The Björklund-Moffitt parameter is the mean gain to participants induced into the program by an instrument. They identify the parameter in a selection framework. Imbens-Angrist show how *IV* can approximate it. Heckman and Vytlačil (1999) show how local instrumental variables (*LIV*) identify it. Heckman, Urzua, and Vytlačil (2006) and Heckman and Vytlačil (2006b) show that *IV* and selection models are closely related. *IV* and its extension Local *IV* (*LIV*) estimate the slopes of the models estimated by selection models in levels.

As pointed out in Heckman, Urzua, and Vytlačil (2006) and Heckman and Vytlačil (2005), the “monotonicity” assumptions made in the *LATE* literature are not innocuous. If, in response to a change in an instrument, some people go into treatment and others drop out, instrumental variables do not identify any treatment effect but they do identify a weighted average of two way flows (Heckman and

²⁵Software for estimating *MTE* and generating all of the treatment parameters is available from Heckman, Urzua, Vytlačil (2006). See the website <http://jenni.uchicago.edu/underiv>.

Vytlacil 2005, 2006b; Heckman, Urzua, and Vytlacil 2006). The recent *IV* literature is asymmetric. Outcomes are permitted to be heterogeneous among persons in a general way. Choices of treatment are not permitted to be heterogeneous in a general way.

7. POLICY RELEVANT TREATMENT PARAMETERS

The Policy Relevant Treatment Effect (*PRTE*) introduced in Heckman and Vytlacil (2001a) and elaborated in Heckman and Vytlacil (2005, 2006a,b) is a good example of the benefits of the econometric approach. It is defined by stating a policy problem—estimating the effect of a policy on mean outcomes—and showing that this treatment effect can be generated as a weighted average of the marginal treatment effect with known weights using formula (1) presented in the preceding section. Standard *IV* and matching estimators do not, in general, identify this parameter.

Policy problems dictate the identification and estimation strategy in our approach. As shown in Heckman (2001), Heckman and Vytlacil (2001a, 2005, 2006b), Heckman, Urzua, and Vytlacil (2006) and Carneiro, Heckman, and Vytlacil (2005), the weights on *MTE* required to form the *PRTE* parameter are generally not the same as the weights for *OLS*, matching or *IV*, although an *IV* estimator can be devised to identify the *PRTE*.

Heckman and Vytlacil (2005) develop an algorithm for defining causal effects that answer specific policy problems from a general list of possible problems rather than relying exclusively on the standard set of causal effects discussed by Sobel in his Section 3 that answer only a few narrowly selected policy problems.²⁶ Sobel ignores parameters like the *PRTE* and fails to recognize that the standard treatment estimators do not identify this parameter. Heckman and Vytlacil develop estimators for specific well-posed policy problems

²⁶In his defense of *ACE*, Sobel makes a familiar error. In defending *ACE* as estimating the effect of a policy with universal coverage compared to the effect with no coverage, he fails to account for the effects of large scale programs on potential outcomes—what economists call “general equilibrium” effects. Heckman, Lochner, and Taber (1998a,b) show that these are empirically important in the analysis of education policies. These effects violate the *SUTVA* assumption of Holland (1986) or the invariance assumption of Hurwicz (1962).

rather than hope that a favored estimator just happens to hit the selected target. This is a large advance over the existing literature in statistics. Just compare Sobel's discussion of IV with our own.

8. ESTIMATING THE PROPORTION OF PEOPLE WHO BENEFIT FROM A PROGRAM

Sobel's discussion of the benefits of randomization illustrates all of the problems with the *ad hoc* statistical approach he favors. Randomized trials cannot identify $\Pr(Y_1 > Y_0)$. In a large sample, this is the proportion of the population that benefits from a program.²⁷ See Heckman (1992). This is because randomized trials produce Y_1 or Y_0 but not both for each person. The parameter $\Pr(Y_1 > Y_0)$ is not even contemplated in the Neyman (1923)–Rubin (1978) setup. Using the Roy model (Heckman and Honoré 1989) or more general models (Carneiro, Hansen, and Heckman 2001, 2003; Cunha, Heckman, and Navarro 2005, 2006; Cunha and Heckman 2006a,b; Heckman, Lochner, and Todd 2006) it is possible to estimate this proportion. Modeling the unobservables and their relationship with the treatment selection rule and any related measurement equations plays an important role in their analysis. The statistical treatment effect literature is silent on this crucial parameter. Modeling the dependence among the unobservables in choice, outcome and auxiliary measurement equations, is the key to identifying this proportion.

Sobel says that “much stronger assumptions” are required to estimate this parameter. In any specific case, this claim is not true. The assumptions required to justify randomization (no randomization bias; no contamination or crossover effects; see Heckman, LaLonde, and Smith 1999) are *different* and not weaker or stronger than the assumptions used to identify the Roy model and its extensions. Indeed when randomization breaks down, Roy models and their generalizations can exploit the attrition and self selection information to identify $\Pr(Y_1 > Y_0)$. See Heckman (1992) and Heckman and Vytlacil (2006a,b).

²⁷I keep the conditioning on covariates implicit. I assume a heterogeneous response model.

9. SENSITIVITY ANALYSIS

If the purpose of my essay had been a comprehensive review of econometric evaluation estimators, I would have discussed bounds and sensitivity analysis. See Heckman and Vytlacil (2006b) for such a discussion using the *MTE* as an organizing device for an entire literature. Vijverberg (1993) is a good reference for sensitivity analysis in a Roy model. Peterson (1976) is an early example of bounds for the competing risks model which is a version of the Roy model.²⁸

10. POLICY FORECASTING

Sobel explores a few special cases of my analysis of problem P2. This section of his discussion abounds with personal opinions like “intuitively easy” and “think reasonably.” These assertions are his excuse for begging the general problems considered in my paper. If he does not like unobservables, he can use standard change of variable arguments to substitute out unobservables from my equations and recast my argument into observables to solve identification problems.

My argument uses structural models directly because they are interpretable in terms of theory and they explicitly recognize missing variables (“unobservables”). They also provide the machinery for integrating information on auxiliary measures into analyses to help overcome problems with missing variables.

His analysis makes implicit exogeneity assumptions. See Heckman and Vytlacil (2005, 2006b) for a general analysis of the extrapolation problem based on the *MTE* which explicitly discusses the role of exogeneity assumptions in policy forecasts.

11. THE ROLE OF UNOBSERVABLES

Sobel, like many statisticians, says he does not like unobservables. Neither do I. I wish all important variables were observed. But it is the unobservables in the outcome equations and the outcomes in the treatment choice

²⁸See Heckman and Honoré (1989, 1990).

equation that give rise to selection and evaluation problems. One can be implicit about them and their properties, as are Sobel and most statisticians, or explicit, as is common in the econometric approach.

Implicit in all statistical analyses are unobservables that generate outcomes given the observables. Matching assumes that all unobservables in outcome equations are random with respect to the unobservables in the treatment choice equation given the matching variables. The *LATE* estimator that Sobel implicitly espouses assumes a latent unobserved variable with a special structure.²⁹ By being clear and objective, the econometric approach allows analysts to pinpoint their differences in assumptions about the unobservables and explore the role of assumptions in producing any differences in conclusions.

12. TOWARDS CONVERGENCE

I applaud Sobel's desire to see the emergence of a literature that combines the best features of the econometric and statistical treatment effect literatures. For that synthesis to occur, statisticians like Sobel should carefully read the econometrics literature and its genuine pioneering contributions and more carefully note the publication dates of key ideas before dismissing 30 years of econometric research as the application of an incompletely formulated statistical model.

REFERENCES

- Abbring, J. H., and J. J. Heckman. 2006. "Dynamic Policy Analysis." In *The Econometrics of Panel Data*, edited by L. Matyas and P. Sevestre. Kluwer Academic Publishers. Forthcoming.
- Ahn, H., and J. Powell. 1993. "Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism." *Journal of Econometrics* 58(1-2):3-29.
- Angrist, J. D., G. W. Imbens, and D. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444-55.

²⁹Vytlacil (2002) explicitly develops the latent variable model implicitly assumed by Imbens and Angrist (1994) and Angrist, Imbens, and Rubin (1996).

- Barnow, B. S., G. G. Cain, and A. S. Goldberger. 1980. "Issues in the Analysis of Selectivity Bias." Pp. 42–59 in *Evaluation Studies*, vol. 5, edited by E. Stromsdorfer and G. Farkas. Beverly Hill, CA: Sage Publications.
- Berk, R., A. Li, and L. Hickman. 2005. "Statistical Difficulties in Determining the Role of Race in Capital Cases: A Re-Analysis of Data from the State of Maryland." *Journal of Quantitative Criminology* 21(4):365–90.
- Björklund, A., and R. Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection." *Review of Economics and Statistics* 69(1):42–49.
- Carneiro, P., K. Hansen, and J. J. Heckman. 2001. "Removing the Veil of Ignorance in Assessing the Distributional Impacts of Social Policies." *Swedish Economic Policy Review* 8(2):273–301.
- . 2003. "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." 2001 Lawrence R. Klein Lecture. *International Economic Review* 44(2):361–422.
- Carneiro, P., J. J. Heckman, and E. J. Vytlačil. 2005. "Understanding What Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education." Presented at Harvard University, 2001. Under review.
- Cartwright, N. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge, UK: Cambridge University Press.
- . 2005. "Causation: One Word Many Things." Unpublished manuscript, London School of Economics, Centre for Philosophy of Natural and Social Science.
- Cunha, F., and J. J. Heckman. 2006a. "The Evolution of Earnings Risk in the US Economy." Presented at the 9th World Congress of the Econometric Society, London.
- . 2006b. "A Framework for the Analysis of Inequality." *Journal of Macroeconomics*, forthcoming.
- Cunha, F., J. J. Heckman, and S. Navarro. 2005. "Separating Uncertainty from Heterogeneity in Life Cycle Earnings." 2004 Hicks Lecture. *Oxford Economic Papers* 57(2):191–261.
- . 2006. "Counterfactual Analysis of Inequality and Social Mobility." In *Mobility and Inequality: Frontiers of Research from Sociology and Economics*, ch. 4, edited by S. L. Morgan, D. B. Grusky, and G. S. Fields. Palo Alto, CA: Stanford University Press.
- Durlauf, S. N., and H. P. Young. 2001. *Social Dynamics*. Cambridge, MA: MIT Press.
- Gronau, R. 1974. "Wage Comparisons—A Selectivity Bias." *Journal of Political Economy* 82(6):1119–43.
- Haavelmo, T. 1943. "The Statistical Implications of a System of Simultaneous Equations." *Econometrica* 11(1):1–12.
- Heckman, J. J. 1974. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42(4):679–94.

- . 1976. “The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models.” *Annals of Economic and Social Measurement* 5(4):475–92.
- . 1978. “Dummy Endogenous Variables in a Simultaneous Equation System.” *Econometrica* 46(4):931–59.
- . 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47(1):153–62.
- . 1980. “Addendum to Sample Selection Bias as a Specification Error.” In *Evaluation Studies Review Annual*, vol. 5, edited by E. Stromsdorfer and G. Farkas. Beverly Hills, CA: Sage Publications.
- . 1987. “Selection Bias and Self-Selection.” Pp. 287–97 in *The New Palgrave: A Dictionary of Economics*, edited by J. Eatwell, M. Milgate, and P. Newman. London: Palgrave Macmillan Press.
- . 1990. “Varieties of Selection Bias.” *American Economic Review* 80(2):313–18.
- . 1992. “Randomization and Social Policy Evaluation.” Pp. 201–30 in *Evaluating Welfare and Training Programs*, edited by C. Manski and I. Garfinkel. Cambridge, MA: Harvard University Press.
- . 2001. “Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture.” *Journal of Political Economy* 109(4):673–748.
- Heckman, J. J., N. Hohmann, J. Smith, and M. Khoo. 2000. “Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment.” *Quarterly Journal of Economics* 115(2):651–94.
- Heckman, J. J., and B. E. Honoré. 1989. “The Identifiability of the Competing Risks Model.” *Biometrika* 76(2):325–30.
- . 1990. “The Empirical Content of the Roy Model.” *Econometrica* 58(5):1121–49.
- Heckman, J. J., R. J. LaLonde, and J. A. Smith. 1999. “The Economics and Econometrics of Active Labor Market Programs.” Pp. 1865–2097 in *Handbook of Labor Economics*, vol. 3A, edited by O. Ashenfelter and D. Card. New York: North-Holland.
- Heckman, J. J., L. J. Lochner, and C. Taber. 1998a. “Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents.” *Review of Economic Dynamics* 1(1):1–58.
- . 1998b. “General-Equilibrium Treatment Effects: A Study of Tuition Policy.” *American Economic Review* 88(2):381–86.
- Heckman, J. J., L. J. Lochner, and P. E. Todd. 2006. “Earnings Equations and Rates of Return: The Mincer Equation and Beyond.” In *Handbook of the Economics of Education*, edited by E. A. Hanushek and F. Welch. Amsterdam: North-Holland. Forthcoming.
- Heckman, J. J., and S. Navarro. 2004. “Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models.” *Review of Economics and Statistics* 86(1):30–57.

- . 2006. “Dynamic Discrete Choice and Dynamic Treatment Effects.” *Journal of Econometrics*. Forthcoming.
- Heckman, J. J., and R. Robb. 1985. “Alternative Methods for Evaluating the Impact of Interventions.” Pp. 156–245 in *Longitudinal Analysis of Labor Market Data*, vol. 10, edited by J. Heckman and B. Singer. New York: Cambridge University Press.
- . 1986. “Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes.” Pp. 63–107 in *Drawing Inferences from Self-Selected Samples*, edited by H. Wainer. New York: Springer-Verlag. Reprinted in 2000, Mahwah, NJ: Lawrence Erlbaum.
- Heckman, J. J., and J. A. Smith. 1998. “Evaluating the Welfare State.” Pp. 241–318 in *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, edited by S. Strom. New York: Cambridge University Press.
- Heckman, J. J., S. Urzua, and E. J. Vytlacil. 2006. “Understanding Instrumental Variables in Models with Essential Heterogeneity.” *Review of Economics and Statistics*. Forthcoming.
- Heckman, J. J., and E. J. Vytlacil. 1999. “Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects.” *Proceedings of the National Academy of Sciences* 96:4730–34.
- . 2001a. “Local Instrumental Variables.” Pp. 1–46 in *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, edited by C. Hsiao, K. Morimune, and J. L. Powell. New York: Cambridge University Press.
- . 2001b. “Policy-Relevant Treatment Effects.” *American Economic Review* 91(2):107–111.
- . 2005. “Structural Equations, Treatment Effects, and Econometric Policy Evaluation.” *Econometrica* 73(3):669–738.
- . 2006a. “Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation.” In *Handbook of Econometrics*, vol. 6, edited by J. Heckman and E. Leamer. Amsterdam: Elsevier, forthcoming.
- . 2006b. “Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Economic Estimators to Evaluate Social Programs and to Forecast Their Effects in New Environments.” In *Handbook of Econometrics*, vol. 6, edited by J. Heckman and E. Leamer, Amsterdam: Elsevier, forthcoming.
- Holland, P. W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81(396):945–60.
- . 1988. “Causal Inference, Path Analysis, and Recursive Structural Equation Models.” Pp. 449–84 in *Sociological Methodology*, edited by C. Clogg and G. Arminger. Washington, DC: American Sociological Association.

- Hurwicz, L. 1962. "On the Structural Form of Interdependent Systems." Pp. 232–39 in *Logic, Methodology and Philosophy of Science*, edited by E. Nagel, P. Suppes, and A. Tarski. Stanford, CA: Stanford University Press.
- Imbens, G. W. 2004. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." *Review of Economics and Statistics* 86(1):4–29.
- Imbens, G. W., and J. D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2):467–75.
- Koopmans, T. C., H. Rubin, and R. B. Leipnik. 1950. "Measuring the Equation Systems of Dynamic Economics." Pp. 53–237 in *Statistical Inference in Dynamic Economic Models*, no. 10, ch. 2 in Cowles Commission Monograph, edited by T. C. Koopmans. New York: John Wiley & Sons.
- Lewis, H. G. 1963. *Unionism and Relative Wages in the United States: An Empirical Inquiry*. Chicago: University of Chicago Press.
- Manski, C. F. 2000. "Identification Problems and Decisions under Ambiguity: Empirical Treatment of Response and Normative Analysis of Treatment Choice." *Journal of Econometrics* 95(2):415–42.
- . 2004. "Statistical Treatment Rules for Heterogeneous Populations." *Econometrica* 72(4):1221–46.
- Marschak, J. 1953. "Economic Measurements for Policy and Prediction." Pp. 1–26 in *Studies in Econometric Method*, edited by W. Hood and T. Koopmans. New York: Wiley.
- Matzkin, R. L. 2006. "Nonparametric Identification." In *Handbook of Econometrics*, vol. 6, edited by J. Heckman and E. Leamer. Amsterdam: Elsevier.
- McFadden, D. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by P. Zarembka. New York: Academic Press.
- . 1981. "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. Manski and D. McFadden. Cambridge, MA: MIT Press.
- Neyman, J. 1923. "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles." *Roczniki Nauk Rolniczych* 10:1–51. In Polish; edited and translated version of Section 9 by D. M. Dabrowska and T. P. Speed. *Statistical Science* 5:465–72.
- Pearl, J. 2000. *Causality*. Cambridge, UK: Cambridge University Press.
- Peterson, A. V. 1976. "Bounds for a Joint Distribution Function with Fixed Sub-distribution Functions: Application to Competing Risks." *Proceedings of the National Academy of Sciences of the United States of America* 73(1):11–13.
- Powell, J. L. 1994. "Estimation of Semiparametric Models." Pp. 2443–2521 in *Handbook of Econometrics*, vol. 4, edited by R. Engle and D. McFadden. Amsterdam: Elsevier.
- Quandt, R. E. 1958. "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes." *Journal of the American Statistical Association* 53(284):873–80.
- . 1972. "A New Approach to Estimating Switching Regressions." *Journal of the American Statistical Association* 67(338):306–10.

- Rosenbaum, P. R. 2002. *Observational Studies* (2d ed.), Series in Statistics. New York: Springer-Verlag.
- Roy, A. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3(2):135–46.
- Rubin, D. B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5):688–701.
- . 1976. "Inference and Missing Data." *Biometrika* 63(3):581–92.
- . 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *Annals of Statistics* 6(1):34–58.
- . 1986. "Statistics and Causal Inference: Comment: Which Ifs Have Causal Answers." *Journal of the American Statistical Association* 81(396):961–62.
- . 2000. "Comment on 'Causal Inference Without Counterfactuals.'" *Journal of the American Statistical Association* 95(450):435–38.
- Thurstone, L. 1927. "A Law of Comparative Judgement." *Psychological Review* 34:273–86.
- . 1959. *The Measurement of Values*. Chicago, IL: University of Chicago Press.
- Vella, F. 1998. "Estimating Models with Sample Selection Bias: A Survey." *Journal of Human Resources* 33(1):127–69.
- Vijverberg, W. P. M. 1993. "Measuring the Unidentified Parameter of the Extended Roy Model of Selectivity." *Journal of Econometrics* 57(1–3):69–89.
- Vytlacil, E. J. 2002. "Independence, Monotonicity, and Latent Index Models: An Equivalence Result." *Econometrica* 70(1):331–41.
- Wainer, H., ed. 1986. *Drawing Inferences from Self-Selected Samples*. Papers from a conference sponsored by Educational Testing Service. Originally published by Springer-Verlag. Reprinted in 2000, Mahwah, NJ: Lawrence Erlbaum Associates.
- Willis, R. J., and S. Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87(5, part 2):S7–S36.