Working paper.

TECHNICAL REPORT
R-420
November 2013

Reflections on Heckman and Pinto's "Causal Analysis After Haavelmo"

Judea Pearl
University of California, Los Angeles
Computer Science Department
Los Angeles, CA, 90095-1596, USA
(310) 825-3243
judea@cs.ucla.edu

Abstract

This paper reflects on a recent article by Heckman and Pinto (2013) in which they discuss a formal system, called *do*-calculus, that operationalizes Haavelmo's conception of policy intervention. They replace the *do*-operator with an equivalent operator called "fix," highlight the capabilities of "fix," discover limitations in "do," and inform readers that those limitations disappear in "the Haavelmo approach." I examine the logic of HP's paper, its factual basis, and its impact on econometric research and education.

1 Introduction

A forthcoming special issue of Econometric Theory, dedicated to Haavelmo's centennial, will contain two papers on causation. The first is "Trygve Haavelmo and the emergence of Causal Calculus" (Pearl, 2013) and the second is "Causal Analysis After Haavelmo" by Heckman and Pinto (HP).

The HP paper is devoted almost entirely to the causal inference framework that I have summarized in (Pearl, 2013) and, in particular, to causal models that can be represented by Directed Acyclic Graphs (DAGs) or Bayesian Networks (Pearl, 1985; Verma and Pearl, 1988) and to the *do*-operator that acts on, and helps draw causal and counterfactual inferences from such models (Pearl, 1993, 1994, 2009; Spirtes et al., 1993; Strotz and Wold, 1960). This note reflects on the way HP present the *do*-operator, and highlights key features of the *do*-calculus that are not described in HP's paper.

2 Summary of "Causal Analysis After Haavelmo"

In a nutshell, what HP's paper does is: (1) replaces the *do*-operator with a a logically equivalent operator called "fix,"[1] (2) unveils the power and capabilities of "fix" while exposing "limitations" of "do," and (3) argues that it is "fix," not "do," which captures the original (yet implicit) intent of Haavelmo. I am pleased of course that Heckman and Pinto took the time to learn the machinery of the *do*-calculus, be it in $do(x)$, $fix(x)$, $set(x)$, $exogenized(x)$, or $randomized(x)$ dressing, and to lay it out before economists so that they too can benefit from its power.

Though we differ on the significance of the difference between the "do" and the "fix" operators, the important thing is that HP call economists' attention to two facts that are practically unknown in the mainstream econometric literature:

1. Identification of causal parameters in the entire class of recursive nonparametric economic models is now a SOLVED PROBLEM, and this include counterfactual parameters related to "effect of treatment on the treated" (ETT), mediation, attribution, external validity, heterogeneity, selection bias, missing data, and more. By "nonparametric" I mean a structural equation model in which no restriction is imposed on the form of the equations or on the distribution of the disturbances (which may be correlated).

2. The age-old confusion between regression and structural parameters (Pearl, 2009, pp. 368–374) can finally come to an end with the help of the notational distinction between "do/fix" vs. "see."

Practically, this means that economics students should now be able to solve the eight toy problems I posed in Pearl, 2013 (see Appendix A, Section A.2). Likewise, students can liberate themselves from the textbook confusion regarding the interpretation of structural parameters, as documented in Chen and Pearl, 2013.

To me, HP's paper reflects Heckman's way of acknowledging the need to translate Haavelmo's ideas into tools of inference, and his determination to satisfy this need by rigorous mathematical means. I am glad that he chose to do so in the style of *do*-calculus, namely, a calculus based on a hypothetical modification of the economic model, often called "surgery," in which variables are exogenized by local reconfiguring of selected equations.[2] Manifestly, Heckman and Pinto do recognize the power and capabilities of "do-calculus", but feel that economists will be more receptive to new tools once the tools are domesticated and treated as home-grown. I concur.

Unfortunately, in the process of domestication, some of the major capabilities of the *do*-calculus were lost while others were presented as "limitations." In par-

[1]Whereas the "do"- operator simulates a hypothetical policy intervention that keeps a variable constant, as in Haavelmo's example of Government deciding "to keep income, r_i, at a given level," the "fix" operator subjects a variable to exogenous variations, as in classical randomized experiments. Clearly, any conclusion obtained in one system is a valid conclusion in the other, as can be seen from the fact that "fix" obeys the axioms of *do*-calculus or from the fact that results of policy decisions can be predicted from controlled randomized experiments (formally proven in *do*-calculus).

[2]In the past, Heckman has resisted the idea of model modification in favor of "external variations" (Heckman and Vytlacil, 2007; Pearl, 2009, pp. 374–380).

ticular, the fact that the *do*-calculus is merely one among several tools of inference that emerges in the framework of Structural Causal Models (SCM) (see Section 2 of (Pearl, 2013)) has escaped HP's description, together with the fact that extensions to simultaneous causation, parametric restrictions, counterfactual reasoning, mediation, heterogeneity, and transportability follow naturally from the SCM framework, and have led to remarkable results.

More unfortunate perhaps is the fact that HP do not address the practical problems posed in Pearl, 2013 (duplicated in Appendix A), which demonstrate tangible capabilities that economists could acquire from the SCM framework. Consequently, the remedy proposed by HP does not equip economists with tools to solve these problems and, in this respect, it falls short of fully utilizing Haavelmo's ideas.

3 Reservations on "Causation After Haavelmo"

My main reservation to HP's presentation of the do/fix calculus is that it does not go all the way to unveil its powers. Specifically, the following two points were sidelined.

1. By not discussing the concept of "completeness" (which *do*-calculus enjoys) HP deny readers one of the major benefits of causal analysis. Completeness means that, if the calculus fails to answer a research question (say whether a causal effect is identifiable) then no such answer exists; i.e., no other method and no other "approach" or "framework" can produce such answer without strengthening the assumptions. The importance of knowing when "no solution exists" is crucial in this line of research, where investigators are often uncertain whether observed discrepencies are due to theoretical impediments, bad design, wrong assumptions, or inadequate framework. (The completeness of *do*-calculus was proven independently by Huang and Valtorta (2006); Shpitser and Pearl (2006).)

2. By delegating the handling of conditional independencies entirely to the mercy of the graphoid axioms (Dawid, 1979; Pearl and Paz, 1986; Pearl, 1988, pp. 82–115), rather than graph separation, HP are preventing econometric students from solving the eight toy problems I posed (Appendix A), as well as many practical problems they face daily (e.g., finding a good IV in a given model). While the graphoid axioms are good for confirming a derivation (of one independence from others), they are not very helpful in FINDING such derivation or in deciding whether one exists. DAGs, on the other hand, make all valid independencies explicit, thus saving us the labor of searching for a valid derivation.[3]

Fortunately, these deficiencies are correctable and I am confident that, as soon as the basic powers of the do/fix calculus come to the attention of econometric students, they will discover for themselves the added capabilities, and will apply them in all aspects of econometric research. (Including, or course, the solution of the toy problems posed in (Pearl, 2013, see Appendix A).)

[3] A good analogy is the search saved by alphabetically-sorted lists vis-a-vis unsorted lists.

4 On the "limitations" of *do*-calculus

HP spend inordinate amount of effort seeking "limitations" in the *do*-operator, in the *do*-calculus, and presumably other methods of representing interventions (e.g., Strotz and Wold, 1960; Spirtes et al., 1993) that preceded HP's interpretation of Haavelmo's papers. These fall into several categories.

4.1 "Fix" vs. "do"

The semantical difference between "fix" and "do" is so infinitesimal that it does not warrant the use of two different labels. As we explained above, there is no conclusion that can be inferred from "fix" that cannot also be inferred from "do," and vice versa. From its birth (Pearl, 1993; Spirtes et al., 1993; Strotz and Wold, 1960) the *do*-operator was used to legitimize results derived from randomized experiments and, conversely, the ideal randomized experiment was used to explain the *do*-operator. Indeed, Fisher's gold standard of controlled experiments consists of two components: *randomization* and *external intervention* (Pearl, 2009, p. 418). In (Pearl, 2013), for example, I introduce $a = \frac{\partial}{\partial x} E(Y|do(x))$ as referring to "a controlled experiment" in which an agent (e.g., Government) is controlling x and observing y.[4]

What is clear in this context is that Haavelmo was more concerned with the idea of "holding X constant at x" than with "randomize X and condition on $X = x$" which is what "fix" instructs us to imagine. In other words, he sought to simulate the actual implementation of a pending policy rather than the Fisherian experiment from which we can learn about the policy.[5]

As I wrote in my book (2009, p. 377): "...most policy evaluation tasks are concerned with *new* external manipulations which exercise direct control over endogenous variables. Take for example a manufacturer deciding whether to double the current price of a given product after years of letting the price track the cost, i.e., *price* = f(*cost*). Such decision amounts to removing the equation *price* = f(*cost*) in the model at hand (i.e., the one responsible for the available data), and replacing it with a constant equal to the new price. This removal emulates faithfully the decision under evaluation, and attempts to circumvent it by appealing to 'external variables' are artificial and hardly helpful." ... "It is also interesting to note that the method used in Haavelmo (1943) to define causal effects is mathematically equivalent to surgery, not to external variation."

[4]It is also interesting to note the operation of "keeping X constant" leads to a formal definition of unit-level counterfactuals, via $Y_x(u) = Y_{M_x}(u)$, (see Appendix A, Definition 1) whereas $P(y|do(x))$ as well as its clone $P(y|\mathit{fix}(x))$ are limited to population-level relations.

[5]I labored to find anything resembling randomized "fix" in Haavelmo's papers, but all I could find was "do," as in "where g_i is Government expenditure, so adjusted as to keep r constant, whatever be b and u" (Haavelmo, 1943, p. 12). Nor could I find an operator resembling "fix" in the econometric literature since (Strotz and Wold, 1960) including the writings of Heckman. Therefore, labeling "fix" "the Haavelmo's approach," however tenaciously (HP, 2013, pp. 2, 5–6, 10, 12, 27, 33, 38), may possibly be a well meaning attempt to endow "fix" with a halo of tradition, or to empower economists with a sense of ownership, but it is historically inaccurate.

HP argue that replacing $P(y|do(X = x))$ with $P_H(y|X = x)$ avoids the use of extra-statistical notation and gives one the comfort of staying within traditional statistics. The comfort however is illusionary and short-lived; it disappears upon realizing that the construction of P_H itself is an extra-statistical operation, for it requires extra-statistical information (e.g., the structure of the causal graph). This craving for orthodox statistical notation is endemic of a long cultural habit to translate the phrase "holding X constant" into probabilistic conditionalization. The habit stems from the absence of probabilistic notation for "holding X constant," which has forced generations of statisticians to use a surrogate in the form of "conditioning on X"; the only surrogate in their disposal. This unfortunate yet persistent habit is responsible for a century of blunders and confusions; from "probabilistic causality" (Pearl, 2011b; Suppes, 1970) to "evidential decision theory (Jeffrey, 1965; Pearl, 2009, pp. 108–109) and Simpson's paradox (Pearl, 2009, pp. 173–180); from Fisher's error in handling mediation (Fisher, 1935; Rubin, 2005) to "Principal Stratification" mishandling of mediation (Pearl, 2011a; Rubin, 2004); from misinterpretations of structural equations (Freedman, 1987; Hendry, 1995; Holland, 1995; Pearl, 2009, pp. 135–138; Sobel, 2008; Wermuth, 1992) to the structural-regressional confusion in econometric textbooks today (Chen and Pearl, 2013).

In light of this rather embarrassing record, I would argue that traditionalists' addiction to conditioning ought to be cured, not appeased. This is especially true in economics, where structural models provide transparent and unambiguous definition of "holding X constant" in the form of $do(X = x)$, rendering surrogates unnecessary.[6]

4.2 Nonparametric models: victory or "limitations"

One of the "major limitations" that HP discover in DAGs is: "A DAG does not generate or characterize any restrictions of functional forms or parametric specifications" (HP, p. 30). They further argue that "the non-identification of the instrumental variable model poses a major limitation for the identification literature that relies exclusively on DAGs."

The word "limitation" is ordinarily attached to a method that fails to perform a task that lies is within its scope. We do not, for instance, describe multiplication to be "limited" due to its inability to perform addition. DAGs were chosen specifically to represent unrestricted functions and cannot therefore be considered "limited" for not generating restrictions that they were designed to relax.

In the early days of causal inference, a community of researchers agreed on the need to minimize modeling assumptions and to explore models that do not impose any restrictions of functional forms or parametric specification. They called such models "nonparametric" (NP); a label that forbids restrictions such as linearity, separability, additivity, monotonicity, effect-homogeneity, non-interaction etc. and set out to explore their limits.

[6]Appeasement attempts using exogenous decision variables, not unlike those invoked by HP, are described in (Dawid, 2002; Pearl, 2009, p. 71). The structural definition of $P_M(y|do(x))$ is $P_{M_x}(y)$ where M_x is a mutilated version of M from which the equation for x is "wiped out," (see Appendix A, Definition 1).

Following this agenda, researchers have labored to unveil the logic behind NP models, and to understand why some permit identification while others do not. In 2006, this labor culminated in a success: A calculus, together with effective algorithms were found (Shpitser and Pearl, 2006) that give us precise and complete answers to the motivating question, "Can we tell, given an ARBITRARY recursive structural model, with latent variables, whether it permits nonparametric identification or not?"

HP portray this success as a failure, noting that "a DAG does not generate or characterize any restrictions of functional forms or parametric specifications" (HP, 2013, p. 30). Avoiding such restrictions is a challenge not a limitation, in much the same way that we regard the Greeks attempts to construct geometrical figures using only a straight edge and a compass to be a challenge, not a limitation.

Greek geometry does not prevent us from constructing fancier tools, beyond straight edge and compass, if we choose to. On the contrary, geometry actually helps us build those tools properly, in much the same way as nonparametric analysts help us harness parametric assumptions properly. If we wish to incorporate additional sources of identifying information, and invoke assumptions such as separability, monotonicity, linearity, effect-homogeneity, non-interaction etc., a powerful logic is available for us to do so. I am speaking here about the logic of structural counterfactuals (Galles and Pearl, 1998; Halpern, 1998; Pearl, 2009, pp. 203–207) that emanates from Principle 1 of causal analysis (Appendix A, Definition 1). Epidemiologists, bio-statisticians, and social scientists are gaining quite a bit of mileage from harnessing additional assumptions through this logic, and there is no reason that economists should stay behind. The HP paper could do more to close this gap.

4.3 The case of the "Generalized Roy model"

HP mentioned the "Generalized Roy model" as a nonrecursive example which *do*-calculus classifies as non-identifiable, allegedly refuting Pearl's claim of solving "all" recursive models. For readers not familiar with the name, the "Generalized Roy model" is a version of the IV (or non-compliance) model treated in Angrist et al., 1996; Balke and Pearl, 1995, 1997; Pearl, 2009, Ch. 8.

The nonparametric version of Roy model reads:[7]

$$\begin{aligned} X &= f(Z, U) \\ Y &= g(U, X) \\ Z \perp\!\!\!\perp U &\quad (Z \text{ is independent of } U) \end{aligned} \tag{1}$$

where $f(\cdot)$ and $g(\cdot)$ are arbitrary functions and U is a vector of unmeasured, yet arbitrarily distributed variables. It is well known that, in this model, the Average Causal Effect (ACE) of X on Y is not identified, except in special cases (Pearl, 1995b), or in unidentified sub-populations (e.g., compliers), when additional restrictions are placed on the functions f and g. An explicit proof is given by the tight bounds derived by Balke (Pearl, 2009, p. 267).

[7] Testable implications of (1) are given in Pearl (1995a) and Richardson and Robins (2010).

According to HP, however, the Roy mode is "nonparametric" and "identifiable," presumably because assumptions such as "monotonicity" or "separability," which are needed for identification in this model, manage to restrict the functions f and g without invoking any parameters.

Does this invalidate the *do*-calculus classification of the Roy model as "non-identifiable nonparametrically"? Not at all. The calculus does exactly what it was designed to do; to decide if identification is feasible when we *do not* allow any restrictions on f and g, except the identities of their arguments (i.e., exclusion restrictions).

Does this render the *do*-calculus criterion too narrow or uninteresting? Let us examine the records. First, it is due to the *do*-calculus that researchers can determine today (through the back-door criterion) what variables need be measured, controlled, or adjusted before identification is possible in fully nonparametric models, with no functional or distributional restrictions.[8]

Second, it is due to the do-calculus that researchers have discovered a class of fully nonparametric models that permit identification by means other than adjustment (or "matching"); the front-door model is one simple example of this class (Pearl, 2009, p. 92).[9] Thirdly, it is the completeness of the *do*-calculus that tells us when functional restrictions are necessary for identification, i.e., that no method whatsoever can identify a causal parameter without such restrictions. Finally, the *do*-calculus is not as helpless as it is portrayed in HP's paper. While the calculus itself merely proclaims models that require functional restrictions as "non-identifiable," the methodological framework in which the calculus is embedded (i.e., SCM and its logic) does not stop at that. It explores, for example, if another variable can be observed which resides between X and Y, or U and Y, to enable identification. It also seeks to impose further assumptions that would produce identification (linearity and monotonicity are examples), as does the standard economic literature, but it does so in a systematic way, because it has the logic of counterfactuals and causal graphs for guidance, and these can tell us when linearity or other assumptions may or may not help. (See Pearl, 2009, Chapters 5 and 9 for use of linearity and monotonicity in identifying counterfactual queries.)

[8] HP call this set of variables "matching variables" (p. 22), which are defined by the conditional independence given in their Lemma L-1. Like the phantom condition of conditional ignorability (Rosenbaum and Rubin, 1983), Lemma L-1 is valid, but does not tell researchers how to decide whether the independence holds in any given model, (see Pearl, 2009, p. 352). Students of *do*-calculus make this decision in seconds, using graph separation, by merely glancing at the economic model (see Appendix A). HP chose to replace graph separation by the Local Markov Condition (LMC) and the graphoid axioms, presumably to prove that things can be done "using conventional matching methods," without modern tools of causal inference. The result is another generation of economists who are unable to identify matching variables in a given economic model. The litmus test remains Section A.2.3 of Appendix A; I challenge HP to demonstrate "conventional matching methods" on these toy problems.

[9] HP labor hard to prove the validity of the front-door formula using the Local Markov Condition (LMC) and the graphoid axioms (HP, pp. 26-29), but do not inform readers how to recognize identifiable effects generally and directly from any economic model. Such recognition requires *do*-calculus and the graphical algorithms that it entails (Shpitser and Pearl, 2006; Tian and Pearl, 2002).

4.4 Why economists do not use the *do*-calculus:

It is tempting to speculate that the scanty use *do*-calculus in economics reflects economists' perception that the class of models handled by the calculus is either narrow or uninteresting.

I take issue with this theory. The main reason, in my opinion, is that economists are still scared of graphs, and this educational deficiency prevents them, not only from using *do*-calculus, but also from doing simple routine tasks of estimation, even in linear models, such as deciding if a system of equations has testable implications (see Appendix A, Example 2.1) or deciding which regression coefficient will remain unaltered if we add another regressor to an equation (see example 2.7). These tasks have little to do with causality or identification; they are invoked frequently in the econometric literature (under the rubric of "misspecification" and "robustness") and, yet, only a handful of economists have the skill and tools to manage them. This educational impairment is the main factor that prevents economists from appreciating much of the recent progress in causal inference. It could have been addressed and rectified by HP's paper.

A second reason is more mystical, and stems from the habit of parametric thinking, often unnecessarily.

How does one know that parametric assumptions are needed if one is not prepared to conduct a nonparametric analysis first, to find out if identification is possible without making any functional restrictions? Let us take the front door model as an example, which in equational form reads as:

$$\begin{aligned} Y &= f(U_1, Z) \\ Z &= g(U_2, X) \\ X &= h(U_1) \text{ and } U_1 \text{ independent of } U_2 \end{aligned} \tag{2}$$

Faced with such a model, an unseasoned economist might be tempted to conclude (by analogy with the Roy model) that ATE is not identifiable nonparametrically, and that some restrictions on f, g, and h are needed for identification. Fortunately, however, causal effects in this model *are* identifiable nonparametrically. Yet this toy problem was not even considered in the econometric literature before 2012 (Chalak and White, 2012), because letting f, g, and h be arbitrary sounds truly scary. This model turned out not only to be identifiable, but to have non-trivial applications in econometric and social science (Chalak and White, 2012; Knight and Winship, 2013; Morgan and Winship, 2007). Next to this one, there are hundreds of fully nonparametric models awaiting to be discovered. These are problems that economists might be tempted, habitually, to analyze by imposing functional or distributional restrictions, yet they are identifiable nonparametrically, and can be recognized as such by glancing at the graph.

For this goal, the completeness of the *do*-calculus sheds a much needed light; it gives us the license to give up and start searching for plausible parametric assumptions.

5 Conclusions

HP's paper is a puzzle. From the fact that HP went to a great length studying the do-calculus, replacing it with a clone called "fix", demonstrating the workings of "fix" on a number of laborious examples and presenting "fix" (not "do") as the legitimate heir of "the Haavelmo approach", one would assume that HP would invite economists to use the new tool of inference as long as they speak "fix" and not "do", and as long as they believe that "fix" is a homegrown product of "the Haavelmo approach." But then the paper presents readers with a slew of "limitations" that apply equally to "fix" and "do" (recall, the two are logically equivalent) and promises readers that "Haavelmo's approach naturally generalizes to remove those limitations" (e.g., simultaneous causation, parametric restrictions, and more). One begins to wonder then what HP's readers are encouraged to do. Should they be content with the traditional literature in which they can find neither "do" nor "fix" nor any other mathematical symbol denoting intervention, or should they cross traditional boundaries and examine first-hand how other communities are benefitting from Haavelmo's ideas?

The main victim of HP's paper is the "fix-operator"; first anointed to demonstrate what "the Haavelmo approach" can do, then indicted with "major limitations" that only "the Haavelmo approach" can undo. What then is the role of the "fix-operator" in economics research? I hope the history of economic thought unravels this puzzle.

I will end this section with comments that I found in a blog run by Kevin Bryan, a PhD student in economics, Kellogg College, Northwestern University (Bryan, 2012).

This is what Kevin writes on SCM and causal calculus:

> "What's cool about SCM and causal calculus more generally is that you can answer a bunch of questions without assuming anything about the functional form of relationships between variables; all you need are the causal arrows. Take a model of observed variables plus unobserved exogenous variables. Assume the latter to be independent. The model might be that X is a function of Y, W, and an unobserved variable U_1, Y is a function of V, W, and U_2, V is a function of U_3 and W is a function of U_4. You can draw a graph of causal arrows relating any of these concepts. With that graph in hand, you can answer a huge number of questions of interest to the econometrician. For instance: what are the testable implications of the model if only X and W are measured? Which variables can be used together to get an unbiased estimate of the effect of any one variable on another? Which variables must be measured if we wish to measure the direct effect of any variable on any other? There are many more, with answers found in Pearl's 2009 textbook. Pearl also comes down pretty harshly on experimentalists of the Angrist type. He notes correctly that experimental potential-outcome studies also rely on a ton of underlying assumptions concerning external validity, in particular and at heart structural models just involve stating those assumptions clearly."

I quote Kevin for two reasons. First, the short paragraph above explains in simple econometric vocabulary what SCM can do, a task that HP's paper had difficulties conveying.

Second, Kevin restored my faith in the future of econometrics; he and students like him will not settle for partial descriptions of merits turned "limitations" but will insist on finding out for themselves what Haavelmo's ideas were and what they can do for economics. These students will be the true beneficiaries of *do*-calculus, and I am grateful to Heckman and Pinto for stimulating their curiosity with the marvels of causal analysis.

Acknowledgement

I thank Rosa Matzkin for commenting on an earlier version of this paper, Elias Bareinboim for defending the logic of *do*-calculus in a somewhat unsupportive environment, and all email discussants who remained silent throughout the discussion, but whose presence nevertheless contributed to the clarity of the ideas presented above.

This research was supported in parts by grants from NSF #IIS-1249822 and #IIS-1302448 and ONR #N00014-13-1-0153 and #N00014-10-1-0933.

Appendix A: Causal Calculus, Tools, and Frills (Based on Section 3 of (Pearl, 2013))

By "causal calculus" I mean mathematical machinery for performing causal inference tasks using Structural Causal Models (SCM).

These include:

1. Tools of reading and explicating the causal assumptions embodied in structural models as well as the set of assumptions that support each individual causal claim.

2. Methods of identifying the testable implications (if any) of the assumptions encoded in the model, and ways of testing, not the model in its entirety, but the testable implications of the assumptions behind each causal claim.

3. Methods of deciding, prior to taking any data, what measurements ought to be taken, whether one set of measurements is as good as to another, and which adjustments need to be made so as to render our estimates of the target quantities unbiased.

4. Methods for devising critical statistical tests by which two competing theories can be distinguished.

5. Methods of deciding mathematically if the causal relationships of interest are estimable from non-experimental data and, if not, what additional assumptions, measurements or experiments would render them estimable.

6. Methods of recognizing and generating equivalent models.

7. Methods of locating instrumental variables for any relationship in a model, or turning variables into instruments when none exists.

8. Methods of evaluating "causes of effects" and predicting effects of choices that differ from the ones actually made, as well as the effects of dynamic policies which respond to time-varying observations.

9. Solutions to the "Mediation Problem," which seeks to estimate the degree to which specific mechanisms contribute to the transmission of a given effect, in models containing both continuous and categorical variables, linear as well as nonlinear interactions (Pearl, 2001, 2012b).

10. Techniques coping with the problem of "external validity" (Campbell and Stanley, 1963), including formal methods of deciding if a causal relation estimated in one population can be transported to another, potentially different population, in which experimental conditions are different (Pearl and Bareinboim, 2011).

A full description of these techniques is given in (Pearl, 2009) as well as in recent survey papers (Pearl, 2010a,b). Here I will demonstrate by examples how some of the simple tasks listed above are handled in the nonparametric framework of a SCM.

A.1 Two models for discussion

Consider a nonparametric structural model defined over a set of endogenous variables $\{Y, X, Z_1, Z_2, Z_3, W_1, W_2, W_3\}$, and unobserved exogenous variables $\{U, U', U_1, U_2, U_3, U'_1, U'_2, U'_2, U'_3\}$. The equations are assumed to be structured as follows:

Model 1

$$\begin{array}{rclcrcl} Y &=& f(W_3, Z_3, W_2, U) & \qquad & X &=& g(W_1, Z_3, U') \\ W_3 &=& g_3(X, U'_3) & & W_1 &=& g_1(Z_1, U'_1) \\ Z_3 &=& f_3(Z_1, Z_2, U_3) & & Z_1 &=& f_1(U_1) \\ W_2 &=& g_2(Z_2, U'_2) & & Z_2 &=& f_2(U_2) \end{array}$$

$f, g, f_1, f_2, f_3, g_1, g_2, g_3$ *are arbitrary, unknown functions, and all exogenous variables are assumed mutually independent but otherwise arbitrarily distributed.*

For the purpose of illustration, we will avoid assigning any economic meaning to the variables and functions involved, thus focusing on the formal aspects of such models rather than their substance. The model conveys two types of theoretical (or causal) assumptions:

1. Exclusion restrictions, depicted by the absence of certain variables from the arguments of certain functions, and

2. Causal Markov conditions, depicted by the absence of common U-terms in any two functions, and the assumption of mutual independence among the U's.

Given the qualitative nature of these assumptions, the algebraic representation is superfluous and can be replaced, without loss of information, with the diagram depicted in Fig. 1.[10] To anchor the discussion in familiar grounds, we also present

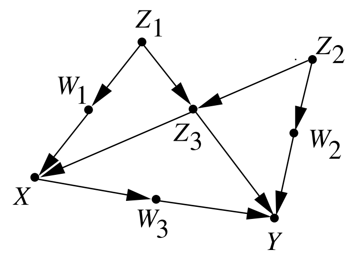

Figure 1: A graphical representation of Model 1. Error terms are assumed mutually independent and not shown explicitly.

the linear version of Model 1:

Model 2 *(Linear version of Model 1)*

$$\begin{array}{rclcrcl}
Y & = & aW_3 + bZ_3 + cW_2 + U & \qquad & X & = & t_1W_1 + t_2Z_3 + U' \\
W_3 & = & c_3X + U_3' & & W_1 & = & a_1'Z_1 + U_1' \\
Z_3 & = & a_3Z_1 + b_3Z_2 + U_3 & & Z_1 & = & U_1 \\
W_2 & = & c_2Z_2 + U_2' & & Z_2 & = & U_2
\end{array}$$

All $U's$ are assumed to be uncorrelated.

While the orthogonality assumption renders these equations regressional, we can easily illustrate non-regressional models by assuming that some of the endogenous variables are not measurable.

A.2 Illustrating typical question-answering tasks

Given the model defined above, the following are typical questions that an economist may wish to ask.

A.2.1 Testable implications (misspecification tests)

a. What are the testable implications of the assumptions embedded in Model 1?

b. Assume that only variables X, Y, Z_3, and W_3 are measured, are there any testable implications?

c. The same, but assuming only variables X, Y, and Z_3 are measured,

d. The same, assuming all but Z_3 are measured.

[10]This is entirely optional; readers comfortable with algebraic representations are invited to stay in their comfort zone.

e. Assume that an alternative model, competing with Model 1, has the same structure, with the $Z_3 \rightarrow X$ arrow reversed. What statistical test would distinguish between the two models?

f. What regression coefficient in Model 2 would reflect the test devised in (e)?

A.2.2 Equivalent models

a. Which arrows in Fig. 1 can be reversed without being detected by any statistical test?

b. Is there an equivalent model (statistically indistinguishable) in which Z_3 is a mediator between X and Y (i.e., the arrow $X \leftarrow Z_3$ is reversed)?

A.2.3 Identification

a. Suppose we wish to estimate the average causal effect of X on Y

$$ACE = P(Y = y|do(X = 1)) - P(Y = y|do(X = 0)).$$

Which subsets of variables need to be adjusted to obtain an unbiased estimate of ACE?

b. Is there a single variable that, if measured, would allow an unbiased estimate of ACE?

c. Assume we have a choice between measuring $\{Z_3, Z_1\}$ or $\{Z_3, Z_2\}$, which would be preferred?

A.2.4 Instrumental variables

a. Is there an instrumental variable for the $Z_3 \rightarrow Y$ relationship?
If so, what would be the IV estimand for parameter b in Model 2?

b. Is there an instrument for the $X \rightarrow Y$ relationship?
If so, what would be the IV estimand for the product $c_3 c$ in Model 2?

A.2.5 Mediation

a. What variables must be measured if we wish to estimate the direct effect of Z_3 on Y?

b. What variables must be measured if we wish to estimate the indirect effect of Z_3 on Y, mediated by X?

c. What is the estimand of the indirect effect in (b), assuming that all variables are binary?

A.2.6 Sampling selection bias[11]

Suppose our aim is to estimate the conditional expectation $E(Y|X = x)$, and samples are preferentially selected to the dataset depending on a set V_S of variables,

a. Let $V_S = \{W_1, W_2\}$, what set, T, of variables need be measured to correct for selection bias? (Assuming we can estimate $P(T = t)$ from external sources e.g., census data.)

b. In general, for which sets, V_S, would selection bias be correctable.

c. Repeat (a) and (b) assuming that our aim is to estimate the causal effect of X on Y.

A.2.7 Linear digressions

Consider the linear version of our model (Model 2)

Question 1: Name three testable implications of this model

Question 2: Suppose X, Y, and W_3 are the only variables that can be observed. Which parameters can be identified from the data?

Question 3: If we regress Z_1 on all other variables in the model, which regression coefficient will be zero?

Question 4: If we regress Z_1 on all the other variables in the model and then remove Z_3 from the regressor set, which coefficient will not change?

Question 5: ("Robustness" – a more general version of Question 4.) Model 2 implies that certain regression coefficients will remain invariant when an additional variable is added as a regressor. Identify five such coefficients with their added regressors.[12]

A.2.8 Counterfactual reasoning

a. Find a set S of endogenous variables such that X would be independent of the counterfactual Y_x conditioned on S.

b. Determine if X is independent of the counterfactual Y_x conditioned on all the other endogenous variables.

---

[11] This section illustrates nonparametric extensions of Heckman's approach to selection bias (Heckman, 1979). A complete theory can be found in Bareinboim and Pearl (2012) and Pearl (2012c).

[12] According to White and Lu (2010) "A common exercise in empirical studies is a 'robustness check,' where the researcher examines how certain 'core' regression coefficient estimates behave when the regression specification is modified by adding or removing regressors." "of the 98 papers published in *The American Economic Review* during 2009, 76 involve some data analysis. Of these, 23 perform a robustness check along the lines just described, using a variety of estimators." Since this practice is conducted to help diagnose misspecification, the answer to Question 5 is essential for discerning whether an altered coefficient indicates misspecification or not.

c. Determine if X is independent of the counterfactual $W_{3,x}$ conditioned on all the other endogenous variables.

d. Determine if the counterfactual relationship $P(Y_x|X = x')$ is identifiable, assuming that only X, Y, and W_3 are observed.

A.3 Solutions

The problems posed in Section A.2 read like homework problems in Economics 101 class. They should be! Because they are fundamental, easily solvable, and absolutely necessary for even the most elementary exercises in nonparametric analysis. Readers should be pleased to know that with the graphical techniques available today, these questions can generally be answered by a quick glance at the graph of Fig. 1 (see, for example, Greenland and Pearl (2011), Kyono (2010), or Pearl (2010a,b, 2012a)).

More elaborate problems, like those involving transportability or counterfactual queries may require the inferential machinery of *do*-calculus or counterfactual logic. Still, such problems have been mathematized, and are no longer at the mercy of unaided intuition, as they are presented for example in Campbell and Stanley (1963).

It should also be noted that, with the exception of our linear digression (A.2.7) into Model 2, all queries were addressed to a purely nonparametric model and, despite the fact that the form of our equations and the distribution of the U's are totally arbitrary, we were able to extract answers to policy-relevant questions in a form that is estimable from the data available.

For example, the answer to the first identification question (a) is: The set $\{W_1, Z_3\}$ is sufficient for adjustment and the resulting estimand is:

$$P(Y = y|do(X = x)) = \sum_{w_1,z_3} P(Y = y|X = x, Z_3 = z_3, W_1 = w_1)P(Z_3 = z_3, W_1 = w_1).$$

This can be derived algebraically using the rules of *do*-calculus or seen directly from the graph, using the back-door criterion (Pearl, 1993). When a policy question is not identifiable, graphical methods can detect it and exit with failure. Put in econometric vocabulary, these results mean that the identification problem in nonparametric triangular simultaneous equations models is now solved. Given any such model, an effective algorithm exists that decides if the causal effect of any subset of variables on another is identifiable and, if so, the algorithm delivers the correct estimand (Shpitser and Pearl, 2008).

The nonparametric nature of these exercises represents the ultimate realization of what Heckman calls the Marschak's Maxim (Heckman, 2010), referring to an observation made by Jacob Marschak (1953) that many policy questions do not require the estimation of each and every parameter in the system – a combination of parameters is all that is necessary and, moreover, it is often possible to identify the desired combination without identifying the individual components. The exercises presented above show that Marschak Maxim goes even further – the desired quantity can often be identified without ever specifying the functional or distributional forms of these economic models.

A.4 What kept the Cowles Commission at bay?

A natural question to ask is why these recent developments have escaped the attention of Marschak and the Cowles Commission who, around 1950, already adopted Haavelmo interpretation of structural models, and have formulated mathematically many of the key concepts and underlying theories that render structural models useful for policy making, including theories of identification, structural invariance and structural estimation. What then prevented them from making the next logical move and tackle nonparametric models such as those exemplified in Section A.2?

I believe the answer lies in two ingredients that where not available to Cowles Commission's researchers and which are necessary for solving nonparametric problems. (These had to wait for the 1980–90's to be developed.) I will summarize these ingredients as "principles" since the entire set of tools needed for solving these problems emanate from these two:

Principle 1: "The law of structural counterfactuals."

Principle 2: "The law of structural independence."

The first principle is described in Definition 1:

Definition 1 (unit-level counterfactuals**)** (Pearl, 2000, p. 98)
Let M be a fully specified structural model and X and Y two arbitrary sets of variables in M. Let M_x be a modified version of M, with the equation(s) of X replaced by $X = x$. Denote the solution for Y in the modified model by the symbol $Y_{M_x}(u)$, where u stands for the values that the exogenous variables take for any given individual (or unit) in the population. The counterfactual $Y_x(u)$ (Read: "The value of Y in unit u, had X been x") is define by

$$Y_x(u) \triangleq Y_{M_x}(u). \tag{3}$$

Principle 2 instructs us how to detect conditional independencies from the structure of the model, i.e., the graph. This principle states that, regardless of the functional form of the equations in a recursive model M, and regardless of the distribution of the exogenous variables U, if the disturbances are mutually independent, the distribution $P(v)$ of the endogenous variables must obey certain conditional independence relations, stated roughly as follows:

> Whenever sets X and Y of nodes in the graph are "separated" by a set Z, X is independent of Y given Z in the probability.[13]

This powerful theorem, called d-separation (Pearl, 2000, pp. 16–18; Verma and Pearl, 1990) constitutes the link between causal relationships encoded in the model and the observed data. It serves as the basis for all graphical models and is used for causal discovery algorithms (Pearl and Verma, 1991; Spirtes et al., 1993) as well as deciding identification and testing misspecification.

[13] The "separation" criterion requires that all paths between X and Y be intercepted by Z, with special handling of paths containing head-to-head arrows (Pearl, 1993; Pearl, 2000, pp. 16–18). In linear models, Principle 2 is valid for non-recursive models as well.

References

Angrist, J., Imbens, G. and Rubin, D. (1996). Identification of causal effects using instrumental variables (with comments). *Journal of the American Statistical Association* **91** 444–472.

Balke, A. and Pearl, J. (1995). Universal formulas for treatment effect from noncompliance data. In *Lifetime Data: Models in Reliability and Survival Analysis* (N. Jewell, A. Kimber, M.-L. Lee and G. Whitmore, eds.). Kluwer Academic Publishers, Dordrecht, 39–43.

Balke, A. and Pearl, J. (1997). Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association* **92** 1172–1176.

Bareinboim, E. and Pearl, J. (2012). Controlling selection bias in causal inference. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*. La Palma, Canary Islands.

Bryan, K. (2012). A Fine Theorem, 'Trygve Haavelmo and the emergence of causal calculus, J. Pearl (2011)'. URL: http://afinetheorem.wordpress.com/2012/02/16/trygve-haavelmo-and-the-emergence-of-causal-calculus-j-pearl-2011/.

Campbell, D. and Stanley, J. (1963). *Experimental and Quasi-Experimental Designs for Research*. Wadsworth Publishing, Chicago.

Chalak, K. and White, H. (2012). An extended class of instrumental variables for the estimation of causal effects. *Canadian Journal of Economics* **44** 1–31.

Chen, B. and Pearl, J. (2013). Regression and causation: A critical examination of six econometrics textbooks. *Real-World Economics Reviews* **65** 2–20.

Dawid, A. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society, Series B* **41** 1–31.

Dawid, A. (2002). Influence diagrams for causal modelling and inference. *International Statistical Review* **70** 161–189.

Fisher, R. (1935). *The Design of Experiments*. Oliver and Boyd, Edinburgh.

Freedman, D. (1987). As others see us: A case study in path analysis (with discussion). *Journal of Educational Statistics* **12** 101–223.

Galles, D. and Pearl, J. (1998). An axiomatic characterization of causal counterfactuals. *Foundation of Science* **3** 151–182.

Greenland, S. and Pearl, J. (2011). Adjustments and their consequences – collapsibility analysis using graphical models. *International Statistical Review* **79** 401–426.

Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica* **11** 1–12. Reprinted in D.F. Hendry and M.S. Morgan (Eds.), *The Foundations of Econometric Analysis*, Cambridge University Press, 477–490, 1995.

Halpern, J. (1998). Axiomatizing causal reasoning. In *Uncertainty in Artificial Intelligence* (G. Cooper and S. Moral, eds.). Morgan Kaufmann, San Francisco, CA, 202–210. Also, *Journal of Artificial Intelligence Research* 12:3, 17–37, 2000.

Heckman, J. (2010). Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature* **48** 356–398.

Heckman, J. and Pinto, R. (2013). Causal analysis after Haavelmo. Tech. Rep. NBER Technical Working Paper 19453, National Bureau of Economic Research, MA.

Heckman, J. and Vytlacil, E. (2007). *Handbook of Econometrics*, vol. 6B, chap. Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation. Elsevier B.V., Amsterdam, 4779–4874.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica* **47** 153–161.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford University Press, New York.

Holland, P. (1995). Some reflections on Freedman's critiques. *Foundations of Science* **1** 50–57.

Huang, Y. and Valtorta, M. (2006). Pearl's calculus of intervention is complete. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence* (R. Dechter and T. Richardson, eds.). AUAI Press, Corvallis, OR, 217–224.

Jeffrey, R. (1965). *The Logic of Decisions*. McGraw-Hill, New York.

Knight, C. and Winship, C. (2013). The causal implications of mechanistic thinking: Identification using directed acyclic graphs (DAGs). In *Handbook of Causal Analysis for Social Research* (S. Morgan, ed.). Springer, 275–299.

Kyono, T. (2010). Commentator: A front-end user-interface module for graphical and structural equation modeling. Tech. Rep. R-364, <http://ftp.cs.ucla.edu/pub/stat_ser/r364.pdf>, Master Thesis, Department of Computer Science, University of California, Los Angeles, CA.

Marschak, J. (1953). Economic measurements for policy and prediction. In *Studies in Econometric Method* (W. C. Hood and T. Koopmans, eds.). Cowles Commission Monograph 10, Wiley and Sons, Inc., 1–26.

Morgan, S. and Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research)*. Cambridge University Press, New York, NY.

Pearl, J. (1985). Bayesian networks: A model of self-activated memory for evidential reasoning. In *Proceedings, Cognitive Science Society*. Irvine, CA.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, CA.

Pearl, J. (1993). Comment: Graphical models, causality, and intervention. *Statistical Science* **8** 266–269.

Pearl, J. (1994). From Bayesian networks to causal networks. In *Bayesian Networks and Probabilistic Reasoning* (A. Gammerman, ed.). Alfred Walter Ltd., London, 1–31.

Pearl, J. (1995a). Causal diagrams for empirical research. *Biometrika* **82** 669–710.

Pearl, J. (1995b). Causal inference from indirect experiments. *Artificial Intelligence in Medicine* **7** 561–582.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York. 2nd edition, 2009.

Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann, San Francisco, CA, 411–420.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge University Press, New York.

Pearl, J. (2010a). The foundations of causal inference. *Sociological Methodology* **40** 75–149.

Pearl, J. (2010b). An introduction to causal inference. *The International Journal of Biostatistics* **6** DOI: 10.2202/1557–4679.1203. Available at: <http://ftp.cs.ucla.edu/pub/stat_ser/r354–corrected–reprint.pdf>.

Pearl, J. (2011a). Principal stratification – a goal or a tool? *The International Journal of Biostatistics* **7**. Article 20, DOI: 10.2202/1557-4679.1322. Available at: <http://ftp.cs.ucla.edu/pub/stat_ser/r382.pdf>.

Pearl, J. (2011b). The structural theory of causation. In *Causality in the Sciences* (P. M. Illari, F. Russo and J. Williamson, eds.). Clarendon Press, Oxford, 697–727.

Pearl, J. (2012a). The causal foundations of structural equation modeling. In *Handbook of Structural Equation Modeling* (R. Hoyle, ed.). Guilford Press, New York, 68–91.

Pearl, J. (2012b). The causal mediation formula – a guide to the assessment of pathways and mechanisms. *Prevention Science* **13** 426–436, DOI: 10.1007/s11121–011–0270–1.

Pearl, J. (2012c). A solution to a class of selection-bias problems. Tech. Rep. R-405, <http://ftp.cs.ucla.edu/pub/stat_ser/r405.pdf>, Department of Computer Science, University of California, Los Angeles, CA.

Pearl, J. (2013). Trygve Haavelmo and the emergence of causal calculus. Tech. Rep. R-391, <http://ftp.cs.ucla.edu/pub/stat_ser/r391.pdf>, Department of Computer Science, University of California, Los Angeles, CA. Forthcoming, *Econometric Theory*, special issue on Haavelmo Centennial.

Pearl, J. and Bareinboim, E. (2011). Transportability of causal and statistical relations: A formal approach. In *Proceedings of the Twenty-Fifth Conference on Artificial Intelligence (AAAI-11)*. Menlo Park, CA. Available at: <http://ftp.cs.ucla.edu/pub/stat_ser/r372a.pdf>.

Pearl, J. and Paz, A. (1986). On the logic of representing dependencies by graphs. In *Proceedings 1986 Canadian AI Conference*. Montreal, Ontario, Canada.

Pearl, J. and Verma, T. (1991). A theory of inferred causation. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference* (J. Allen, R. Fikes and E. Sandewall, eds.). Morgan Kaufmann, San Mateo, CA, 441–452.

Richardson, T. and Robins, J. (2010). Analysis of the binary instrumental variable model. In *Heuristics, Probability, and Causality* (R. Dechter, H. Geffner and J. Halpern, eds.). College Publications, London, 415–440.

Rosenbaum, P. and Rubin, D. (1983). The central role of propensity score in observational studies for causal effects. *Biometrika* **70** 41–55.

Rubin, D. (2004). Direct and indirect causal effects via potential outcomes. *Scandinavian Journal of Statistics* **31** 161–170.

Rubin, D. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association* **100** 322–331.

Shpitser, I. and Pearl, J. (2006). Identification of conditional interventional distributions. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence* (R. Dechter and T. Richardson, eds.). AUAI Press, Corvallis, OR, 437–444.

Shpitser, I. and Pearl, J. (2008). Complete identification methods for the causal hierarchy. *Journal of Machine Learning Research* **9** 1941–1979.

Sobel, M. (2008). Identification of causal parameters in randomized studies with mediating variables. *Journal of Educational and Behavioral Statistics* **33** 230–231.

Spirtes, P., Glymour, C. and Scheines, R. (1993). *Causation, Prediction, and Search*. Springer-Verlag, New York.

Strotz, R. and Wold, H. (1960). Recursive versus nonrecursive systems: An attempt at synthesis. *Econometrica* **28** 417–427.

Suppes, P. (1970). *A Probabilistic Theory of Causality.* North-Holland Publishing Co., Amsterdam.

Tian, J. and Pearl, J. (2002). A general identification condition for causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*. AAAI Press/The MIT Press, Menlo Park, CA, 567–573.

Verma, T. and Pearl, J. (1988). Causal networks: Semantics and expressiveness. In *Proceedings of the Fourth Workshop on Uncertainty in Artificial Intelligence*. Mountain View, CA. Also in R. Shachter, T.S. Levitt, and L.N. Kanal (Eds.), *Uncertainty in AI 4*, Elesevier Science Publishers, 69–76, 1990.

Verma, T. and Pearl, J. (1990). Equivalence and synthesis of causal models. In *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence*. Cambridge, MA. Also in P. Bonissone, M. Henrion, L.N. Kanal and J.F. Lemmer (Eds.), *Uncertainty in Artificial Intelligence 6*, Elsevier Science Publishers, B.V., 255–268, 1991.

Wermuth, N. (1992). On block-recursive regression equations. *Brazilian Journal of Probability and Statistics* (with discussion) **6** 1–56.

White, H. and Lu, X. (2010). Robustness checks and robustness tests in applied economics. Tech. rep., Department of Economics University of California, San Diego, CA. Discussion paper.