A Distinction between Causal Effects in Structural and Rubin Causal Models

Dionissi Aliprantis*

April 28, 2017

Abstract: Unspecified mediators play different roles in the outcome equations of Structural Causal Models (SCMs) and Rubin Causal Models (RCMs), also known, respectively, as structural equation modeling and the potential outcome framework. Direct effects are defined for potential outcomes relative to a set of specified mediators, but direct effects in structural equations are defined relative to all possible mediators at a given scale of observation. An implication is that potential outcomes can be generated by many Directed Acyclic Graphs (DAGs), while a structural outcome equation can be generated by only one DAG.

Keywords: Causal Effect, Structural Equation, Potential Outcome

*Federal Reserve Bank of Cleveland, Research Department, PO Box 6387, Cleveland, OH 44101-1387, USA. Phone: +1(216)579-3021. E-mail: *dionissi.aliprantis@clev.frb.org*.

I thank Francisca G.-C. Richter for helpful comments. The opinions expressed are those of the author and do not represent views of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

1 Introduction

Structural Causal Models (SCMs) and Rubin Causal Models (RCMs), also known, respectively, as structural equation modeling and the potential outcome framework, are often viewed as analogues (Pearl (2014b), Pearl (2012)). One formal result is that the structural equation and potential outcome approaches to modeling selection are equivalent: A latent index model of selection is equivalent to selection under the assumptions of independence and monotonicity (Vytlacil (2002)). It has also been shown that SCMs are formally equivalent to RCMs in the following sense: There exists a SCM generating the hypothetical contingency tables represented by any RCM (Pearl (2009b) p 244, Halpern (2000), Galles and Pearl (1998)).

This paper discusses a key distinction between SCMs and RCMs: The models make different assumptions about the role of unspecified mediators in the outcome equation.[1] In an SCM all mediators are explicitly specified, so that changing a variable outside of the model is assumed to not change the outcome variable. Potential outcomes make no claims about how changing a variable other than treatment or a specified mediator will or will not change the outcome variable.

The fact that potential outcomes do not represent an "all causes" model (Heckman and Vytlacil (2007)) can be seen as both a strength and a weakness. The strength is that potential outcomes are able to characterize Data Generating Processes (DGPs) without knowing all mediators (Freedman (1987), Imai et al. (2010)), but the weakness is that they require assumptions similar to those invoked in the structural equation approach when used to make out of sample forecasts (Heckman and Vytlacil (2005)).

Distinguishing between direct effects in Structural and Rubin Causal Models illustrates that any given RCM has multiple Directed Acyclic Graphs. In other words, there exist many SCMs generating the contingency tables represented by any RCM. This highlights the tradeoffs between working with causal effects from either type of model. While causal effects from RCMs may be more easily identified, they can be less useful than causal effects from SCMs for predicting changes in the outcome variable when the mediators in a DGP change (Aliprantis (2015)). This gap between RCMs and SCMs offers clarity on a small part of the spectrum between effects discovered atheoretically by computers and those found within a conceptual model (Keane (2010), Weinberger (2017)).

2 Unspecified Mediators and Structural Equations

The model at time t, $\mathcal{M}_t$, is a set of relationships between variables specified by the econometrician, and the DGP, $\mathcal{D}_t$, is the set of relationships between variables specified by nature. One definition of structural equation might be inspired by Angrist et al. (1996a) (p 450)

Definition 1 $Y_i = Y_i(D)$ is a structural equation *if there exists some variable* Z such that
$Z \cap Y = Z \cap D = \emptyset \quad$ and $\quad Y_i(d, z') = Y_i(d, z) \quad$ for all $\quad z' \neq z$.

Another definition might be inspired by Pearl (2009b) (Definition 5.4.1):

[1] Mediators are variable that affect the outcome variable that are themselves affected by changes in treatment.

Definition 2 $Y_i = Y_i(D, \epsilon)$ is a structural equation *if for any variable* M *observable at the same scale as* Y *and* D, $M \cap Y = M \cap D = M \cap \epsilon = \emptyset$ implies that $Y_i(D, m', \epsilon) = Y_i(D, m, \epsilon)$ for all $m' \neq m$.

To investigate the implications of adopting one of these definitions rather than the other, consider the following class of DGPs, which represent a typical mediation system without confounders (Pearl (2014a)).[2] Suppose that at time $t \in \mathbb{N}$ data are generated by a DGP in which the outcome variable (Y_{ti}) for each individual i is causally determined by two observed variables, treatment (D_{ti}) and an observed mediator (M_{ti}), as well as unmeasured factors (U^Y_{ti}). The unmeasured covariates can be broken down into those factors that are unobserved (E_{ti}) and those that are unobservable (ϵ_{ti}) at the given scale of measurement. In order to focus on unobserved mediators and outcome equations, I consider DGPs without unobserved confounders.

If Z is an instrument mimicking an observable intervention manipulating treatment and U represents unmeasured variables, then the DGP $\mathcal{D}_t$ is characterized by the following structural equations:[3]

$$Z_{ti} \leftleftarrows f^Z_t(U^Z_{ti}) \tag{1}$$
$$D_{ti} \leftleftarrows f^D_t(Z_{ti}, U^D_{ti}) \tag{2}$$
$$M_{ti} \leftleftarrows f^M_t(D_{ti}, U^M_{ti}) \tag{3}$$
$$Y_{ti} \leftleftarrows f^Y_t(D_{ti}, M_{ti}, U^Y_{ti}) \tag{4}$$

The $\leftleftarrows$ notation is used to indicate that these are structural equations under Definition 2. Under both definitions, a structural equation contains information about counterfactual manipulations to the right hand side variables, and not only information about what is passively observed in the data. This represents an asymmetric relationship between the variables on the left and right hand sides of the equation.[4] For example, structural Equation 4 provides information about the counterfactual values of Y_{ti} if we were to control the variables on the *rhs*, but makes no claim about how any of the variables on the *rhs* would behave if we were to control the *lhs* variable Y_{ti} (Pearl (2009b), p 160). A standard equation makes no distinction between interventions counterfactually manipulating right hand side variables and information about what is passively observed in the data.[5]

Under Definition 2, a further implication of the $\leftleftarrows$ notation is that given control over the specified observable (but not necessarily observed) variables on the *rhs* of the equation, changes made to additional observable variables would provide no further change to the outcome variable.

[2] These DGPs are specified to represent simple systems amenable to potential outcomes (Pearl et al. (2014)).

[3] In Pearl (2009b)'s Definition 7.1.1 of a structural causal model, the triple $< U, V, F >$ is defined here as $U \equiv (U^Z_{ti}, U^D_{ti}, U^M_{ti}, U^Y_{ti})$, $V \equiv (Z_{ti}, D_{ti}, M_{ti}, Y_{ti})$, $F \equiv (f^Z_{ti}, f^D_{ti}, f^M_{ti}, f^Y_{ti})$. In this specification Z_{ti} randomizes D_{ti} so as to satisfy the standard ignorability assumption (Imbens (2014), White and Lu (2011)), and Z_{ti} separately satisfies the relevant exclusion restriction from the outcome equation (Angrist et al. (1996b)).

[4] Some examples of the asymmetry of "directions of influence" (Strotz and Wold (1960)) include the fact that symptoms do not cause disease (Pearl (2009a)), a child's height does not cause her father's height (Goldberger (1991)), and rainfall determines crop yields but not the reverse.

[5] Chalak and White (2012) study these distinctions in terms of settings and responses, and White and Chalak (2009) allow for causal effects in systems of symmetric equations.

This is what distinguishes Definition 2: All variables at the given scale of observation not included on the right hand side satisfy an exclusion restriction (Pearl (2009b), Definition 5.4.1). Definition 1 requires only that an exclusion restriction must hold for *one* additional variable not included on the *rhs* (the one manipulating treatment), not *all* variables not included as arguments of the outcome function (Angrist et al. (1996a)).

3 Defining Causal Effects

3.1 Defining Causal Effects as Changes from Interventions to a DGP

One definition of causal effects is as a quantitative characterization of the change in the outcome variable that would result from an intervention to the DGP. Such interventions to the DGP can be characterized by how they would, or would not, impact mediators, especially unmeasured variables. In order to be precise about which features of the DGP are changed, and which are not, under specific interventions, I use Pearl (2009b)'s *do*-operator throughout the remainder of the analysis.

Direct effects characterize the change in the outcome variable from a specific type of intervention to the DGP. Specifically, the controlled direct effect of D_t on Y_t, $\Delta_{ti}^{CDE}(d', d)$, is the change in Y_t that would result from an intervention setting D_t from d to d' while setting all other variables entering as arguments in f^Y to fixed values:

$$\begin{aligned}\boldsymbol{\Delta}_{\mathbf{t}}^{\mathbf{CDE}}(\mathbf{d}', \mathbf{d}) &\equiv \mathbb{E}[f_t^Y(d', m, U_{ti}^Y)] - \mathbb{E}[f_t^Y(d, m, U_{ti}^Y)] \\ &= \mathbb{E}[Y_{ti}|do(D_{ti} = d', M_{ti} = m)] - \mathbb{E}[Y_{ti}|do(D_{ti} = d, M_{ti} = m)]\end{aligned}$$

Following Pearl (2014a), this definition is made at the population level, with individual-level effects given by the expressions under the expectation. Expectations are taken over U_{ti}^Y for the $\Delta_{ti}^{CDE}(d', d)$.

Direct effects are often defined relative to a reference set of variables that are set to fixed values by an intervention, and the values to which the reference variables are set (Spirtes et al. (2001)). Under such a definition, direct effects need not be invariant to interventions changing variables outside that reference set. This direct effect, however, is invariant to changes to any variables outside the parents of Y_{ti}, as the reference set in this definition is all other variables at a given scale of measurement. This is because the definition of structural equation adopted in this paper indicates that *all* variables at the given level of observation not included on the *rhs* of Equation 4 satisfy an exclusion restriction. In other words, because we are examining the DGP and not a model of it, mediators outside the reference set can only be found at a finer scale of observation (A process which, as noted by Holland (1988), always appears to be possible.).

A second useful definition of causal effect is the total effect:

$$\begin{aligned}\boldsymbol{\Delta}_{\mathbf{t}}^{\mathbf{TE}}(\mathbf{d}', \mathbf{d}) &\equiv \mathbb{E}\big[\, f_t^Y\big(\ d',\quad f_t^M(d', U_{ti}^M),\quad U_{ti}^Y\big)\,\big] \;-\; \mathbb{E}\big[\, f_t^Y\big(\ d,\quad f_t^M(d, U_{ti}^M),\quad U_{ti}^Y\big)\,\big] \\ &= \mathbb{E}[Y_{ti}|do(D_{ti} = d')] - \mathbb{E}[Y_{ti}|do(D_{ti} = d)].\end{aligned}$$

As with the $\Delta_{ti}^{CDE}(d', d)$, this definition is also made at the population level, with individual-level effects given by the expressions under the expectation. Expectations are taken over U_{ti}^{M} and U_{ti}^{Y} for the $\Delta_{ti}^{TE}(d', d)$.

Defining the vector $S \equiv (D, M)$ and re-writing Equation 4 more compactly as $Y_{ti} \xleftarrow{} Y_{ti}(S_{ti})$, both direct and total effects can be written in terms of the econometric or graphical definitions given in Heckman (2008) and Pearl (2009b) (Definition 3.2.1).

3.2 Defining Causal Effects as Changes from Interventions to a Model

The Rubin Causal Model (Rubin (2005), Angrist et al. (1996a)) defines causal effects in terms of the counterfactual outcome variable that would be observed under interventions to treatment. These counterfactual outcomes are also known as potential outcomes,

$$Y_{ti}(D_{ti}),$$

where $Y_{ti}(d)$ is the outcome of individual i at time t if treatment were set to $D_{ti} = d$ by an intervention setting D_{ti} but affecting none of the mediators of the total effect of D_{ti} on Y_{ti} (ie, none of the other parents of Y_{ti}). Although potential outcomes are generated by the DGP, they are defined as features of a model $\mathcal{M}_t$ that can describe many DGPs. That is, the average causal effect in the Rubin Causal Model is defined as

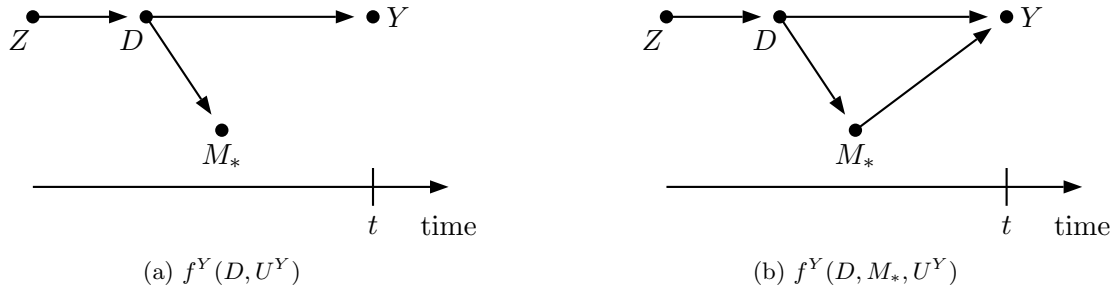
$$\mathbf{\Delta_t^{RCM}}(\mathbf{d}', \mathbf{d}) \equiv \mathbb{E}[Y_{ti}(d') - Y_{ti}(d)].$$

The expectation in $\Delta_t^{RCM}(d', d)$ is taken over individuals in the given population, allowing for any number of underlying functional forms and distributions. In contrast, the expectation in $\Delta_t^{TE}(d', d)$ is taken with respect to the single set of functional forms and distributions specified by the DGP. This causal effect can be generalized to incorporate mediators, including multiple mediators (VanderWeele and Vansteelandt (2014)).

4 Interventions to a DGP versus Interventions to a Model

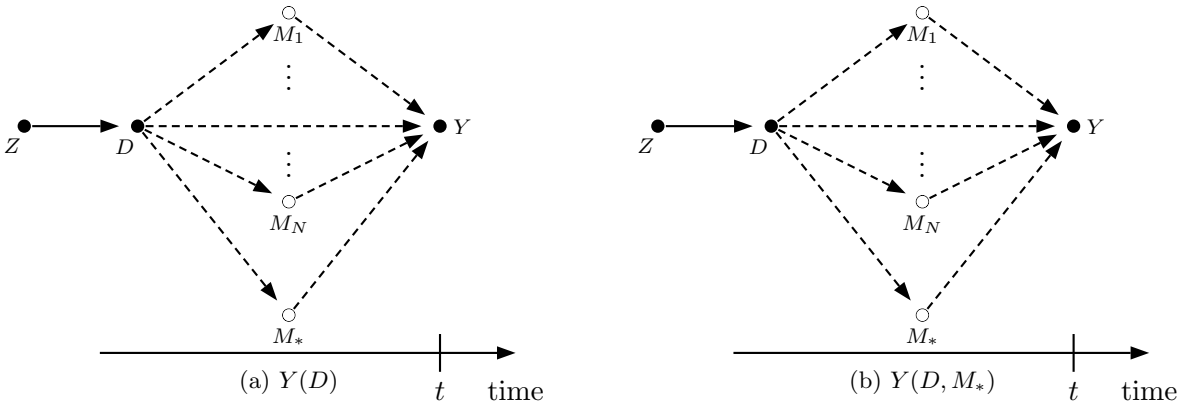
Let $\{M_*, M_1, M_2, \ldots, M_N\}$ be a set of unspecified potential mediators. Figure 1 shows that only the variable M_* specified in the structural outcome equation can appear as parents in the DAG of a SCM.[6] As a result, SCM can only be represented by one DAG. In contrast, any of the variables $\{M_*, M_1, M_2, \ldots, M_N\}$ not specified in the potential outcome equation can appear as parents in the DAG of a RCM (Figure 2). As a result, many DAGs, and therefore many DGPs, could generate the same potential outcomes.

[6]This Figure follows the convention from Pearl (2009b) of communicating that a variable is observed by drawing a solid line to its descendants, and communicating that a variable is unobserved by drawing a dashed line to its descendants.

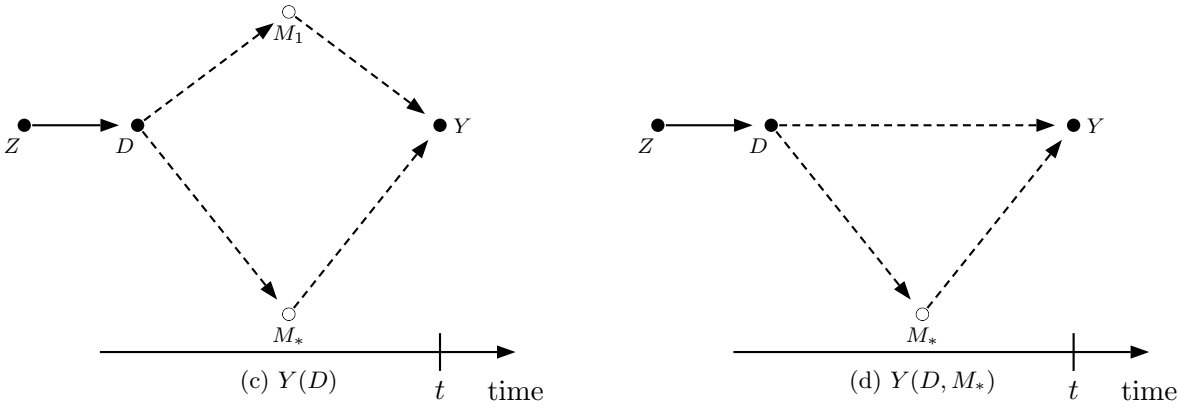


(a) $f^Y(D, U^Y)$ (b) $f^Y(D, M_*, U^Y)$

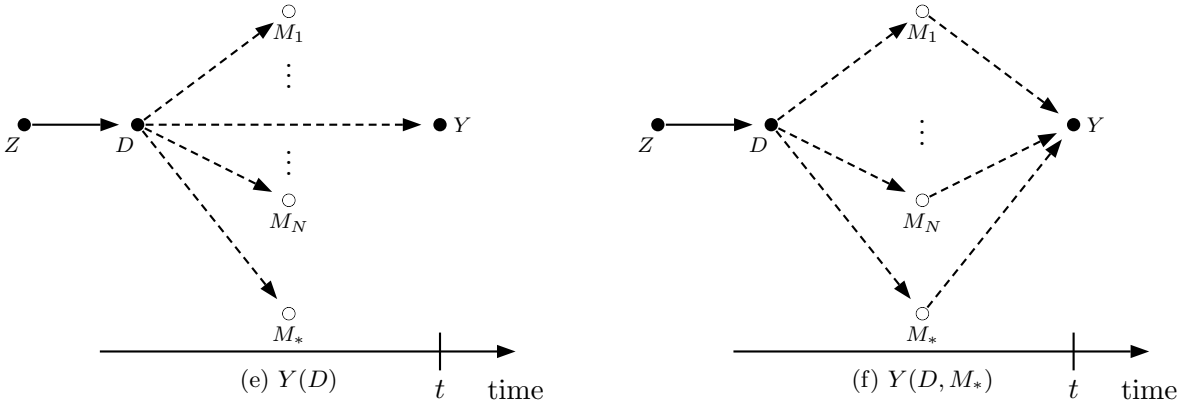
Figure 1: Directed Acyclic Graphs of DGPs Represented by Structural Outcome Equations



(a) $Y(D)$ (b) $Y(D, M_*)$



(c) $Y(D)$ (d) $Y(D, M_*)$



(e) $Y(D)$ (f) $Y(D, M_*)$

Figure 2: Directed Acyclic Graphs of DGPs Represented by Potential Outcomes

References

Aliprantis, D. (2015). Covariates and causal effects: The problem of context. *Federal Reserve Bank of Cleveland Working Paper 13-10R*.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996a). Identification of causal effects using Instrumental Variables. *Journal of the American Statistical Association 91*(434), 444–455.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996b). Rejoinder. *Journal of the American Statistical Association 91*(434), 465–468.

Chalak, K. and H. White (2012). Causality, conditional independence, and graphical separation in settable systems. *Neural Computation* (24), 1611–1668.

Freedman, D. A. (1987). As others see us: A case study in path analysis. *Journal of Educational Statistics 12*(2), 101–128.

Galles, D. and J. Pearl (1998). An axiomatic characterization of causal counterfactuals. *Foundations of Science 3*(1), 151–182.

Goldberger, A. S. (1991). *A Course in Econometrics*. Harvard University Press.

Halpern, J. Y. (2000). Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research 12*(1), 317–337.

Heckman, J. J. (2008). Econometric causality. *International Statistical Review 76*(1), 1–27.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Heckman, J. J. and E. Vytlacil (2007). Econometric evaluation of social programs, Part I: Causal models, structural models and econometric policy evaluation. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6B, Chapter 70, pp. 4779 – 4874. Elsevier.

Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. *Sociological Methodology 18*(1), 449–484.

Imai, K., L. Keele, and D. Tingley (2010). A general approach to causal mediation analysis. *Psychological Methods 15*(4), 309–334.

Imbens, G. W. (2014). Instrumental Variables: An econometrician's perspective. *Statistical Science 29*(3), 323–358.

Keane, M. P. (2010). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics 156*(1), 3–20.

Pearl, J. (2009a). Causal inference in statistics: An overview. *Statistics Surveys 3*, 96–146.

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

Pearl, J. (2012). The causal foundations of Structural Equation Modeling. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling*, Chapter 5, pp. 68–91. New York: Guilford Press.

Pearl, J. (2014a). The deductive approach to causal inference. *Journal of Causal Inference 2*(2), 115–129.

Pearl, J. (2014b). Interpretation and identification in causal mediation analysis. *Psychological Methods 19*(4), 459–481.

Pearl, J., G. Imbens, B. Chen, and E. Bareinboim (2014, October-November). Are economists smarter than epidemiologists? (Comments on Imbens' recent paper). *UCLA Causality Blog*. http://www.mii.ucla.edu/causality/?p=1241.

Rubin, D. B. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association 100*(469), 322–331.

Spirtes, P., C. Glymour, and R. Scheines (2001). *Causation, Prediction, and Search* (2nd ed.). Bradford.

Strotz, R. H. and H. O. Wold (1960). Recursive vs. nonrecursive systems: An attempt at synthesis. *Econometrica 28*(2), 417–427.

VanderWeele, T. and S. Vansteelandt (2014). Mediation analysis with multiple mediators. *Epidemiologic Methods 2*(1), 95–115.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica 70*(1), 331–341.

Weinberger, D. (2017, April 18). Alien knowledge: When machines justify knowledge. *Backchannel*.

White, H. and K. Chalak (2009). Settable systems: An extension of Pearl's Causal Model with optimization, equilibrium, and learning. *Journal of Machine Learning Research* (10), 1759–1799.

White, H. and X. Lu (2011). Causal diagrams for treatment effect estimation with application to efficient covariate selection. *Review of Economics and Statistics 93*(4), 1453–1459.