Causal, Casual and Curious

Judea Pearl*

Is Scientific Knowledge Useful for Policy Analysis? A Peculiar Theorem Says: No

Abstract: Conventional wisdom dictates that the more we know about a problem domain the easier it is to predict the effects of policies in that domain. Strangely, this wisdom is not sanctioned by formal analysis, when the notions of "knowledge" and "policy" are given concrete definitions in the context of nonparametric causal analysis. This note describes this peculiarity and speculates on its implications.

Keywords: policy evaluation, transportability, causal effects, identification, instrumental variables

*Corresponding author: Judea Pearl, Department of Computer Science, University of California – Los Angeles, Los Angeles, CA, 90095-1596, USA, E-mail: judea@cs.ucla.edu

1 Introduction

In her book, *Hunting Causes and Using Them* [1], Nancy Cartwright expresses several objections to the $do(x)$ operator and the "surgery" semantics on which it is based (pp. 72 and 201). One of her objections concerned the fact that the *do*-operator represents an ideal, atomic intervention, different from the one implementable by most policies under evaluation. According to Cartwright, for policy evaluation "we generally want to know what would happen were the policy really set in place," and "the policy may affect a host of changes in other variables in the system, some envisaged and some not."

In my answer to Cartwright [2, p. 363], I stressed two points. First, the *do*-calculus enables us to evaluate the effect of compound interventions as well, as long as they are described in the model and are not left to guesswork. Second, I claimed that "in many studies our goal is not to predict the effect of the crude, non-atomic intervention that we are about to implement but, rather, to evaluate an ideal, atomic policy that cannot be implemented given the available tools, but that represents nevertheless scientific knowledge that is pivotal for our understanding of the domain."

The example I used was as follows: Smoking cannot be stopped by any legal or educational means available to us today; cigarette advertising can. That does not stop researchers from aiming to estimate "the effect of smoking on cancer," and doing so from experiments in which they vary the instrument – cigarette advertisement – not smoking. The reason they would be interested in the atomic intervention $P(cancer|do(smoking))$ rather than (or in addition to) $P(cancer|do(advertising))$ is that the former represents a stable biological characteristic of the population, uncontaminated by social factors that affect susceptibility to advertisement, thus rendering it transportable across cultures and environments. With the help of this stable characteristic, one can assess the effects of a wide variety of practical policies, each employing a different smoking-reduction instrument. For example, if careful scientific investigations reveal that smoking has no effect on cancer, we can comfortably conclude that increasing cigarette taxes will not decrease cancer rates and that it is futile for schools to invest resources in anti-smoking educational programs.

This note takes another look at this argument, in light of recent results in transportability theory (Bareinboim and Pearl [3], hereafter BP).

2 A theorem and its implications

The question investigated in BP was whether one can infer the causal effect of X on Y by randomizing a surrogate variable Z, which is more easily controllable than X. This problem was addressed earlier in Pearl [2, pp. 88–89] where a sufficient condition was derived for a variable Z to act as an experimental surrogate for X. BP have obtained a condition that is both necessary and sufficient for surrogacy, which reads as follows:

Theorem 1 *(BP [3])*,
The causal effect $P(y|do(x))$ can be inferred from experiments on Z if and only if:

1. $P(y|do(x))$ can be inferred from observational studies alone, or
2(i). All directed paths from Z to Y go through X, and
2(ii). $P(y|do(x), do(z))$ can be inferred from observational studies.

Remark: Condition 2(i), in effect, turns Z into an instrumental variable, when randomized.

If X stands for a treatment, then Z plays the role of an "intent-to-treat" variable in noncompliance situations. Condition 2(i) ensures that Z has no side effects on Y; i.e. it acts as an instrumental variable when randomized. Condition 2(ii) ensures a nonparametric identification of treatment effects, using Z as an instrument [4–6].

Figure 1(a) and (b) illustrates models where both 2(i) and 2(ii) are satisfied, while in Figure 1(c) 2(i) fails, because a directed path exists from Z to Y. For example, if Z represents cigarette tax and X represents smoking, then we can infer the causal effect of smoking on cancer, $P(y|do(x))$, by experimenting with tax rates; 2(i) is satisfied because taxes do not directly affect cancer, and 2(ii) is satisfied because, in Figure 1(a) and (b), $P(y|do(x))$ is identifiable in the models that result from intervening on Z (i.e. deleting all arrows pointing to Z.)
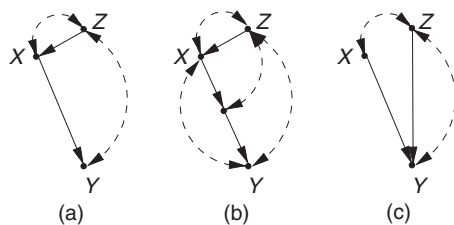

Figure 1 Models (a) and (b) satisfy the conditions of Theorem 1, thus permitting the identification of $P(y|do(x))$ from experiments conducted on Z. Model (c) does not permit this identification because of the arrow from Z to Y

We now return to the question of whether scientific knowledge can be useful in evaluating practical policies. We ask: Suppose $do(Z = z)$ represents a specific implementation of a policy that intends to enact $do(X = x)$, ostensibly because the latter is not directly implementable. Would knowledge of $P(y|do(x))$ help us evaluate $P(y|do(z))$, the policy that is "really set in place"?

Formally, the problem amounts to reversing the role of X and Z in Theorem 1 and yields:

Theorem 2 *The causal effect $P(y|do(z))$ can be inferred from observational studies and knowledge of $P(y|do(x))$ if and only if*:

1. $P(y|do(z))$ can be inferred from observational studies alone, or
2(i). All directed paths from X to Y go through Z, and
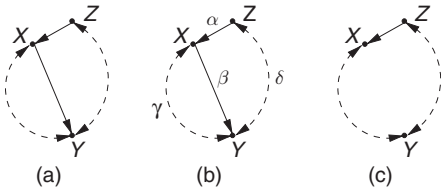*2(ii). $P(y|do(x), do(z))$ is identifiable in observational studies.*

Figure 2 Model (a) does not satisfy Conditions 1 and 2(i) of Theorem 2, thus prohibiting the identification of $P(y|do(z))$ from knowledge of $P(y|do(x))$. Model (b), which is a linear version of (a), permits this identification. Model (c) trivially permits this identification due to the missing arrow from X to Y

This is a surprising result, saying in effect that knowing how X affects Y (i.e. $P(y|do(x))$) is useless for estimating the effect of a policy $do(Z = z)$ that is intended to utilize the effect of X on Y. Put differently, knowing how effective a treatment is does not tell us how effective any policy is, which is intended to administer that treatment in practice. This can be seen by noting that 2(i) cannot be satisfied unless Z contains descendants of X, and this will never be the case when Z is chosen so as to influence Y through X. Therefore, the causal effect $P(y|do(z))$ can be inferred from knowledge of $P(y|do(x))$ if and only if it can be inferred from observational studies alone, as in Condition 1.

To see the ramification of this impossibility result, consider again the smoking-cancer example, depicted in Figure 2. Here Z represents cigarette tax, X represents smoking, and Y represents cancer. Our aim is to estimate the effect of policy $do(Z = z)$ (setting the level of cigarette taxes) on cancer. The dashed curved line between Z and Y represents confounding factors, for example, factors that render communities that impose high cigarette taxes more diet-conscience, hence, less cancer prone. In model 2(a), neither $P(y|do(z))$ nor $P(y|do(x))$ is identifiable from observational data (as can be seen from the graphical criteria of Shpitser and Pearl [7]), and the question we ask is whether knowledge of $P(y|do(x))$ can help us identify $P(y|do(z))$. Theorem 2 answers this question in the negative, since Z does not block the directed path from X to Y, thus violating Condition 2(i).

3 Discussion

This result is peculiar, for it implies that policies such as imposing cigarette taxes cannot be informed by knowing the extent to which smoking causes cancer. It reflects an idiosyncratic property of nonparametric analysis in which knowledge of causal effects (such as $P(y|do(x))$ is insufficient to turn other causal effects identifiable. In other words, the requirement of nonparametric identification (of $P(y|do(z))$) is so stringent that the information provided by other causal effects (e.g. $P(y|do(x))$) is too weak to make a difference.

Things are different in parametric systems, as can be seen from Figure 2(b), which represents a linear version of Figure 2(a), with parameters α and β. Here, the causal effect of Z on Y is $\alpha\beta$ that is not identifiable. However, if the causal effect (β) of X on Y is given, $\alpha\beta$ is identifiable because α can easily be estimated by regression ($\alpha = cov(Z,X)/var(Z)$).

Another exception to this impossibility result is the case where X has zero effect on Y, namely, $P(y|do(x)) = P(y)$. In this case, 2(i) is satisfied by default, since there is no directed path from X to Y, as shown in Figure 2(c) and the conclusion $P(y|do(z)) = P(y)$ follows. Indeed, if smoking has no effect on cancer it would be futile to attempt a reduction in cancer cases by increasing tax on cigarettes.

This observation mitigates substantially our initial disappointment with formal analysis. It implies that, whereas knowledge of $P(y|do(x))$ does not yield a *point-estimate* of $P(y|do(z))$ it provides, nevertheless, an *interval estimate* that vanishes when X is known to have no effect on Y at the population level, i.e. $P(y|do(x)) = P(y)$. It would be interesting to find out, in general, how quantitative knowledge of non-zero effects helps reduce uncertainties about practical policies.

Finally, another exception to Theorem 2 occurs when a policy $do(Z = 1)$ can enforce a treatment $do(X = 1)$ deterministically, i.e. with no exceptions. For example, if the policy $do(Z = 1)$ stands for inoculating every individual in the population, then the implication $Z = 1 \Rightarrow X = 1$ renders $P(y|do(X = 1))$ identifiable whenever Condition 2(i) of Theorem 1 holds, that is, when Z has no side effects on Y (see Pearl [2, p. 358]).

References

1. Cartwright N. Hunting causes and using them: approaches in philosophy and economics. New York, NY: Cambridge University Press, 2007.
2. Pearl J. Causality: models, reasoning, and inference, 2nd ed. New York: Cambridge University Press, 2009.
3. Bareinboim E, Pearl J. Causal inference by surrogate experiments: z-identifiability. In: de Freitas N, Murphy K, editors. Proceedings of the twenty-eighth conference on uncertainty in artificial intelligence. Corvallis, OR: AUAI Press, 2012:113–20.
4. Angrist J, Imbens G, Rubin D. Identification of causal effects using instrumental variables (with comments). J Am Stat Assoc 1996;91:444–72.
5. Balke A, Pearl J. Universal formulas for treatment effect from noncompliance data. In: Jewell N, Kimber A, Lee M-L, Whitmore G, editors. Lifetime data: models in reliability and survival analysis. Dordrecht: Kluwer Academic Publishers, 1995:39–43.
6. Balke A, Pearl J. Bounds on treatment effects from studies with imperfect compliance. J Am Stat Assoc 1997;92:1172–6.
7. Shpitser I, Pearl J. Complete identification methods for the causal hierarchy. J Mach Learn Res 2008;9:1941–79.