TECHNICAL REPORT
R-493
August 2020

A Crash Course in Good and Bad Controls

Carlos Cinelli* Andrew Forney† Judea Pearl ‡

August 28, 2020

Abstract

Many students, especially in econometrics, express frustration with the way a problem known as "bad control" is evaded, if not mishandled, in the traditional literature. The problem arises when the addition of a variable to a regression equation produces an unintended discrepancy between the regression coefficient and the effect that the coefficient is expected to represent. Avoiding such discrepancies presents a challenge not only to practitioners of econometrics, but to all analysts in the data intensive sciences. This note describes graphical tools for understanding, visualizing, and resolving the problem through a series of illustrative examples. We have found that the examples presented here can serve as a powerful instructional device to supplement formal discussions of the problem. By making this "crash course" accessible to instructors and practitioners, we hope to avail these tools to a broader community of scientists concerned with the causal interpretation of regression models.

Introduction

Students trained in the traditional econometrics pedagogy have likely encountered the problem of "bad controls" (Angrist and Pischke, 2009, 2014). The problem arises when an analyst needs to decide whether or not the addition of a variable to a regression equation helps getting estimates closer to the parameter of interest. Analysts have long known that some variables, when added to the regression equation, can produce unintended discrepancies between the regression coefficient and the effect that the coefficient is expected to represent. Such variables have become known as "bad controls," to be distinguished from "good controls" (also known as "confounders" or "deconfounders") which are variables that must be added to the regression equation to eliminate what

*Department of Statistics, University of California, Los Angeles. Email: carloscinelli@ucla.edu
†Department of Computer Science, Loyola Marymount University, Los Angeles. Email: Andrew.Forney@lmu.edu
‡Department of Computer Science, University of California, Los Angeles. Email: judea@cs.ucla.edu

came to be known as "omitted variable bias"(Angrist and Pischke, 2009; Steiner and Kim, 2016; Cinelli and Hazlett, 2020).

The problem of "bad controls" however, is usually evaded, if not mishandled in the standard econometrics literature. While most of the widely adopted textbooks discuss the problem of omitting "relevant" variables, they do not provide guidance on deciding which variables are relevant, nor which variables, if included in the regression, could induce, or worsen existing biases (see Chen and Pearl (2013) for a recent appraisal of econometrics textbooks). Researchers exposed only to this literature may get the impression that adding "more controls" to a regression model is always better. The few exceptions that do discuss the problem of "bad controls" (see Angrist and Pischke (2009, 2014); Imbens and Rubin (2015)) unfortunately cover only a narrow aspect of the problem. Typical is the discussion found in Angrist and Pischke (2009, p.64)

> Some variables are bad controls and should not be included in a regression model, even when their inclusion might be expected to change the short regression coefficients. Bad controls are variables that are themselves outcome variables in the notional experiment at hand. That is, bad controls might just as well be dependent variables too. Good controls are variables that we can think of having been fixed at the time the regressor of interest was determined.

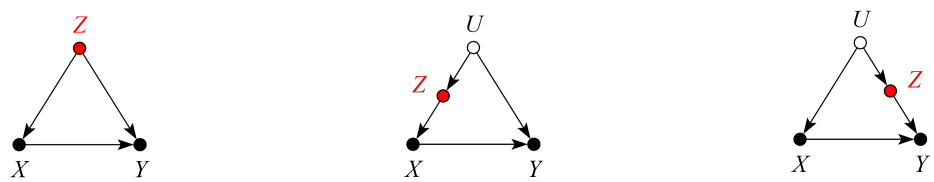
Here, "good controls" are defined as variables that are thought to be unaffected by the treatment, whereas "bad controls" are variables that could be in principle affected by the treatment. Similar discussion can be found in Rosenbaum (2002) and Rubin (2009), for qualifying a variable for inclusion in propensity score analysis. Although an improvement over an absence of discussion, these conditions are neither necessary nor sufficient for deciding whether a variable is a good control.

Recent advances in graphical models have produced simple criteria to distinguish "good" from "bad" controls; these range from necessary and sufficient conditions for deciding which set of variables should be adjusted for to identify the causal effect of interest (e.g, the backdoor criterion and adjustment criterion (Pearl, 1995; Shpitser et al., 2012)), to, among a set of valid adjustment sets, deciding which ones would yield more precise estimates (Hahn, 2004; White and Lu, 2011; Henckel et al., 2019; Rotnitzky and Smucler, 2019; Witte et al., 2020). The purpose of this note is to provide practicing analysts a concise, simple, and *visual* summary of these criteria through illustrative examples.

Here we will assume that readers are familiar with the basic notions of causal inference, directed acyclic graphs (DAGs), and in particular "path-blocking" (or d-separation) as well as back-door paths. For a gentle introduction, we refer readers to Pearl (2009a, Sec. 11.1.2). In the following set of models, the target of the analysis is the *average causal effect* (ACE) of a treatment X on an outcome Y, which stands for the expected increase of Y per unit of a *controlled* increase in X. Observed variables will be designated by black dots and unobserved variables by white empty circles. Variable Z, highlighted in red, will represent the variable whose inclusion in the regression equation

is to be decided, with "good control" standing for bias reduction, "bad control" standing for bias increase, and "neutral control" when the addition of Z neither increases nor decreases the asymptotic bias. For this last case, we will also make brief remarks about how Z could affect the precision of the ACE estimate.

Models 1, 2 and 3 – Good Controls

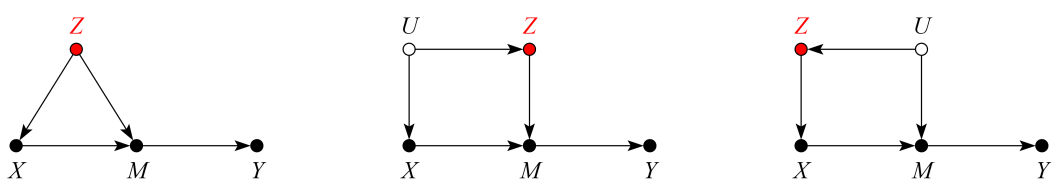


(a) Model 1 (b) Model 2 (c) Model 3

Figure 1: Models 1, 2 , and 3.

In Model 1, Z stands for a common cause of both X and Y. Once we control for Z, we block the back-door path from X to Y, producing an unbiased estimate of the ACE. In Models 2 and 3, Z is not a common cause of both X and Y, and therefore, not a traditional "confounder" as in Model 1. Nevertheless, controlling for Z blocks the back-door path from X to Y due to the unobserved confounder U, and again, produces an unbiased estimate of the ACE.

Models 4, 5 and 6 – Good Controls



(a) Model 4 (b) Model 5 (c) Model 6

Figure 2: Models 4, 5 and 6.

When thinking about possible threats of confounding, modelers need to keep in mind that common causes of X and any mediator (between X and Y) also confound the effect of X on Y. Therefore, Models 4, 5 and 6 are analogous to Models 1, 2 and 3—controlling for Z blocks the backdoor path from X to Y and produces an unbiased estimate of the ACE.

Model 7 – Bad Control

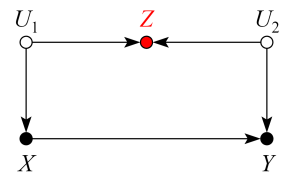


Figure 3: Model 7

We now encounter our first "bad control." Here Z is correlated with the treatment and the outcome and it is also a "pre-treatment" variable. Traditional econometrics textbooks would deem Z a "good control" (Angrist and Pischke, 2009, 2014; Imbens and Rubin, 2015). The backdoor criterion, however, reveals that Z is a "bad control." Controlling for Z will induce bias by opening the backdoor path $X \leftarrow U1 \rightarrow Z \leftarrow U2 \rightarrow Y$, thus spoiling a previously unbiased estimate of the ACE. This structure is known as the "M-bias", and has spurred several discussions. Readers can find further discussion in Pearl (2009a, p. 186), and Shrier (2009); Pearl (2009c,b); Sjölander (2009); Rubin (2009); Ding and Miratrix (2015); Pearl (2015).

Model 8 – Neutral Control (possibly good for precision)

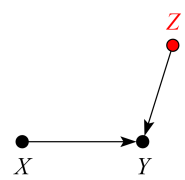


Figure 4: Model 8

In Model 8, Z is not a confounder nor does it block any backdoor paths. Likewise, controlling for Z does not open any backdoor paths from X to Y. Thus, in terms of asymptotic bias, Z is a "neutral control." Analysis shows, however, that controlling for Z reduces the variation of the outcome variable Y, and helps improving the precision of the ACE estimate in finite samples (Hahn, 2004; White and Lu, 2011; Henckel et al., 2019; Rotnitzky and Smucler, 2019).

Model 9 – Neutral Control (possibly bad for precision)

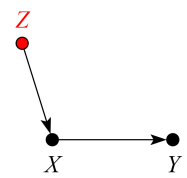


Figure 5: Model 9

Similar to the previous case, in Model 9 Z is "neutral" in terms of bias reduction. However, controlling for Z will reduce the variation of the treatment variable X and so may hurt the precision of the estimate of the ACE in finite samples (Henckel et al., 2019, Corollary 3.4). As a general rule of thumb, parents of X which are not necessary for identification are harmful for the asymptotic variance of the estimator; on the other hand, parents of Y which do not spoil identification are beneficial. See Henckel et al. (2019) for recent developments in graphical criteria for efficient estimation via adjustment in linear models. Remarkably, these conditions also have been show to hold for non-parametric models (Rotnitzky and Smucler, 2019).

Model 10 – Bad Control

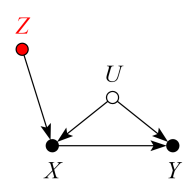
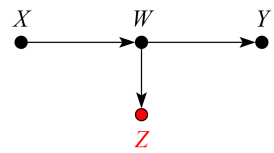


Figure 6: Model 10

We now encounter our second "pre-treatment" "bad control," due to a phenomenon called "bias amplification" (Bhattacharya and Vogt, 2007; Wooldridge, 2009; Pearl, 2011, 2010, 2013; Middleton et al., 2016; Steiner and Kim, 2016). Naive control for Z in this model will not only fail to deconfound the effect of X on Y, but, in linear models, will amplify any existing bias.

Models 11 and 12 – Bad Controls

(a) Model 11



(b) Model 12

Figure 7: Models 11 and 12

If our target quantity is the ACE, we want to leave all channels through which the causal effect flows "untouched." In Model 11, Z is a mediator of the causal effect of X on Y. Controlling for Z will block the very effect we want to estimate, thus biasing our estimates. In Model 12, although Z is not itself a mediator of the causal effect of X on Y, controlling for Z is equivalent to partially controlling for the mediator M, and will thus bias our estimates. Models 11 and 12 violate the backdoor criterion (Pearl, 2009a), which excludes controls that are descendants of the treatment along paths to the outcome.

Model 13 – Neutral Control (possibly good for precision)

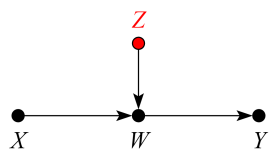


Figure 8: Model 13

At first look, Model 13 might seem similar to Model 12, and one may think that adjusting for Z would bias the effect estimate, by restricting variations of the mediator M. However, the key difference here is that Z is a cause, not an effect, of the mediator (and, consequently, also a cause of Y). Thus, Model 13 is analogous to Model 8, and so controlling for Z will be neutral in terms of bias and may increase precision of the ACE estimate in finite samples. Readers can find further discussion of this case in Pearl (2013).

Models 14 and 15 – Neutral Controls (possibly helpful in the case of selection bias)

(a) Model 14

(b) Model 15

Figure 9: Models 14 and 15

Contrary to econometrics folklore, not all "post-treatment" variables are inherently bad controls. In Models 14 and 15 controlling for Z does not open any confounding paths between X and Y. Thus, Z is neutral in terms of bias. However, controlling for Z does reduce the variation of the treatment variable X and so may hurt the precision of the ACE estimate in finite samples. Additionally, in Model 15, suppose one has only samples with $W = 1$ recorded (a case of selection bias). In this case, controlling for Z can help obtaining the W-specific effect of X on Y, by blocking the colliding path due to W.

Model 16 – Bad Control

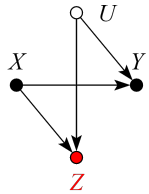


Figure 10: Model 16

Contrary to Models 14 and 15, here controlling for Z is no longer harmless. Adjusting for Z opens the backdoor path $X \to Z \leftarrow U \to Y$ and so biases the ACE.

Model 17 – Bad Control

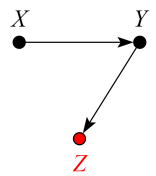


Figure 11: Model 17

In our last example, Z is not a mediator, and one might surmise that, as in Model 14, controlling for Z is harmless. However, controlling for the effects of the outcome Y will induce bias in the estimate of the ACE, making Z a "bad control." A visual explanation of this phenomenon using "virtual colliders" can be found in (Pearl, 2009a, Sec. 11.3). Model 17 is usually known as a "case-control bias" or "selection bias." Finally, although controlling for Z will generally bias numerical estimates of the ACE, it does have an exception when X has no causal effect on Y. In this scenario, X is still d-separated from Y even after conditioning on Z. Thus, adjusting for Z is valid for testing whether the effect of X on Y is *zero*.

Final remarks

In this note, we demonstrated through illustrative examples how simple graphical criteria can be used to decide when a variable should (or should not) be included in a regression equation—and thus whether it can be deemed a "good" or "bad" control. Many of these examples act as cautionary notes against prevailing practices in traditional econometrics: for instance, Models 7 to 10 reveal that one should be wary of the propensity score "mantra" of conditioning on all "pre-treatment predictors of the treatment assignment;" whereas Models 14 and 15 demonstrate that not all "post-treatment" variables are "bad-controls", and some may even help with identification.

In all cases, structural knowledge is indispensable for deciding whether a variable is a good or bad control, and graphical models provide a natural language for articulating such knowledge, as well as efficient tools for examining its logical ramifications. We have found that an example-based approach to "bad controls," such as the one presented here, can serve as a powerful instructional device to supplement formal discussions of the problem. By making this "crash course" accessible to instructors and practitioners, we hope to avail these tools to a broader community of scientists concerned with the causal interpretation of regression models.

References

Angrist, J. and Pischke, J.-S. (2009). *Mostly harmless econometrics: an empiricists guide*. Princeton: Princeton University Press.

Angrist, J. D. and Pischke, J.-S. (2014). *Mastering 'metrics: The path from cause to effect*. Princeton University Press.

Bhattacharya, J. and Vogt, W. B. (2007). Do instrumental variables belong in propensity scores? Technical report, National Bureau of Economic Research.

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

Cinelli, C. and Hazlett, C. (2020). Making sense of sensitivity: extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1):39–67.

Ding, P. and Miratrix, L. W. (2015). To adjust or not to adjust? sensitivity analysis of m-bias and butterfly-bias. *Journal of Causal Inference*, 3(1):41–57.

Hahn, J. (2004). Functional restriction and efficiency in causal inference. *Review of Economics and Statistics*, 86(1):73–76.

Henckel, L., Perković, E., and Maathuis, M. H. (2019). Graphical criteria for efficient total effect estimation via adjustment in causal linear models. *arXiv preprint arXiv:1907.02435*.

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Middleton, J. A., Scott, M. A., Diakow, R., and Hill, J. L. (2016). Bias amplification and bias unmasking. *Political Analysis*, 24(3):307–323.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4):669–688.

Pearl, J. (2009a). *Causality*. Cambridge University Press.

Pearl, J. (2009b). Letter to the editor: Remarks on the method of propensity score. *Statistics in Medicine*, 28:1420–1423. URL: `https://ucla.in/2NbS14j`.

Pearl, J. (2009c). Myth, confusion, and science in causal analysis. *UCLA Cognitive Systems Laboratory*, Technical Report (R-348). URL: `https://ucla.in/2EihVyD`.

Pearl, J. (2010). On a class of bias-amplifying variables that endanger effect estimates. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pages 417–424. URL: `https://ucla.in/2N8mBMg`.

Pearl, J. (2011). Invited commentary: understanding bias amplification. *American journal of epidemiology*, 174(11):1223–1227. URL: `https://ucla.in/2PORDX2`.

Pearl, J. (2013). Linear models: A useful "microscope" for causal analysis. *Journal of Causal Inference*, 1(1):155–170. URL: `https://ucla.in/2LcpmHz`.

Pearl, J. (2015). Comment on ding and miratrix:"to adjust or not to adjust?". *Journal of Causal Inference*, 3(1):59–60. URL: `https://ucla.in/2PgOWNd`.

Rosenbaum, P. R. (2002). *Observational studies*. Springer.

Rotnitzky, A. and Smucler, E. (2019). Efficient adjustment sets for population average treatment effect estimation in non-parametric causal graphical models. *arXiv preprint arXiv:1912.00306*.

Rubin, D. B. (2009). Should observational studies be designed to allow lack of balance in covariate distributions across treatment groups? *Statistics in Medicine*, 28(9):1420–1423.

Shpitser, I., VanderWeele, T., and Robins, J. M. (2012). On the validity of covariate adjustment for estimating causal effects. *arXiv preprint arXiv:1203.3515*.

Shrier, I. (2009). Propensity scores. *Statistics in Medicine*, 28(8):1317–1318.

Sjölander, A. (2009). Propensity scores and m-structures. *Statistics in medicine*, 28(9):1416–1420.

Steiner, P. M. and Kim, Y. (2016). The mechanics of omitted variable bias: Bias amplification and cancellation of offsetting biases. *Journal of causal inference*, 4(2).

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

Witte, J., Henckel, L., Maathuis, M. H., and Didelez, V. (2020). On efficient adjustment in causal graphs. *arXiv preprint arXiv:2002.06825*.

Wooldridge, J. (2009). Should instrumental variables be used as matching variables. Technical report, Citeseer.