Causal progress with imperfect placebo treatments and outcomes [DRAFT]

Adam Rohde*, Chad Hazlett†

March 2023

Abstract

In the quest to make defensible causal claims from observational data, it is sometimes possible to leverage information from "placebo treatments" and "placebo outcomes" (aka "negative outcome controls") unaffected by the treatment. Traditional approaches employing such information focus largely on point identification and assume (i) "perfect placebos" (placebo treatments have precisely zero effect on the outcome; the real treatment has precisely zero effect on a placebo outcome); and (ii) that the placebo treatment/outcome suffers the same amount of confounding as does the real treatment- outcome relationship on some scale ("equi-confounding"). We take a different approach, showing how the analysis of "omitted variable bias" in regression provides a flexible and powerful way to leverage information from placebo treatments and outcomes while relaxing these assumptions. This provides an easy-to-use partial identification framework for demonstrating the implications of any claim a user chooses to support regarding the relative strengths of confounding suffered by a placebo treatment/outcome compared to the true treatment-outcome relationship. We further permit "imperfect placebos", i.e. placebo treatments with up to some postulated non-zero effect on the outcome, or placebo outcomes experiencing effects of treatment up to some postulated strength. These tools can be easily applied with any choice of placebo treatment or outcome, observed in the same dataset or in a separate one as the main treatment-outcome relationship. Conventional difference-in-difference approaches, in both the repeated cross-section and panel settings, are a special case of this setting in which a strict equi-confounding assumption (regarding the pre-treatment outcome's relationship to the treatment group indicator) is claimed. We demonstrate several relaxations of this that are natural under our framework.

1 Introduction

Unobserved confounding plagues many attempts to draw causal inferences from observational data. Moreover, chance imbalances between the treatment and control groups in randomized experiments, sample selection, measurement error can further threaten the validity of causal inferences for both observational and experimental studies. While standard practice is to check balance on observed covariates and to statistically adjust for these variables in estimating causal effects, these attempts are often unconvincing in their ability to overcome unobserved confounding and other unmeasured threats to validity. When we suspect unobserved confounding exists, we often look for additional sources of information that can aid us in identifying causal effects of interest. This may take the form of an instrumental variable, some sort of discontinuity in the assignment of which units get treated, or an observed variable that is similar to the treatment or outcome. This last approach is to leverage information from "placebo outcomes" or "placebo treatments" (also known as "negative outcome controls" and "negative exposure controls," respectively) to aid in the quest to make defensible causal claims, despite the presence of unobserved confounding (or other issues). Existing approaches using such information typically focus on simply detecting bias or on point identification (i.e., bias correction), assuming (i) that these are "perfect placebos" (treatment has precisely zero effect on a placebo outcome, or placebo treatments have precisely zero effects on the outcome of interest) and (ii) that the placebo outcome/treatment suffers the same amount of confounding as does the real treatment- outcome relationship on some scale ("equi-confounding").

In this paper, we will take a partial identification approach that allows us to relax both of these types of assumptions. We suggest practitioners start from only those assumptions that are credibly defensible based on scientific knowledge. In practice, assumptions of exact equi-confounding or of perfect placebos may not be credible. We should not, as a result, shy away from attempting to make causal inferences; instead, we should try to understand how different the levels of confounding are and how imperfect are the placebos. We can use subject matter knowledge and previous studies to bound these parameters in order to partially identify the causal effect of interest. We can iteratively make stronger and stronger assumptions to investigate how the effect of interest changes when these assumptions are made. But it is important to start from the most credible setting, since this setting provides us with what we can conclude with confidence. Subsequent refinements may also

*Department of Statistics, UCLA. adamrohde@ucla.edu

†Associate Professor, Departments of Statistics & Political Science, UCLA. chazlett@ucla.edu

be interesting, but these should only follow the most credible findings. The approach we are describing is the inverse of how typical applied causal studies proceed. Generally, researchers begin with hard to justify claims like exact equi-confounding, present their results under such strong assumptions as their main results, and only then investigate how violations of these results may change their effect estimates in a sensitivity analysis.[1] We feel that this gets the order, and certainly the emphasis, wrong. An effect estimate based on a indefensible assumptions is meaningless. However, a range of effect estimates based on defensible subject matter expertise focuses on what credibly know. We can subsequently explore what less defensible assumptions would entail, but such results should be appropriately located within the range of plausible estimates. Moreover, an approach that emphasizes credible *ranges* of estimates helps to immunize inquiry from a bias towards "positive" results. Perhaps the plausible range includes zero or perhaps not. Presenting a range however helps lessen pressure for there to be a binary reading of what may be nuanced findings. Fortunately, most existing frameworks built for sensitivity analysis can also be used for partial identification. We develop a novel partial identification framework for for leveraging information about placebo treatments and outcomes.

Simple examples of *perfect* placebo treatments and outcomes are shown in Figure 1(a,b). In this figure N is a placebo outcome and P is a placebo treatment; however, Figure 1(b) is essentially a mirror image of Figure 1(a) with N relabelled as P. What we consider a placebo treatment versus what what we consider a placebo outcome will often have this "Necker cube" quality. A Necker cube is a simple optical illusion in which the same two-dimensional drawing of a cube can be perceived with the cube oriented in two distinct ways. (Necker, 1832) Depending on the substantive context, the same causal graph might arise in two settings, one in which the placebo is viewed as a placebo treatment and one in which the placebo is viewed as a placebo outcome. We explore this more in Section 3 and discuss how the perspective taken by researchers should influence the methods they use. Figure 1(c,d) contains simple examples of *imperfect* placebos. These "imperfections" arise from adding the edge $D \to N$ to Figure 1(a) to get Figure 1(c) and from adding the edge $P \to Y$ to Figure 1(b) to get Figure 1(d).

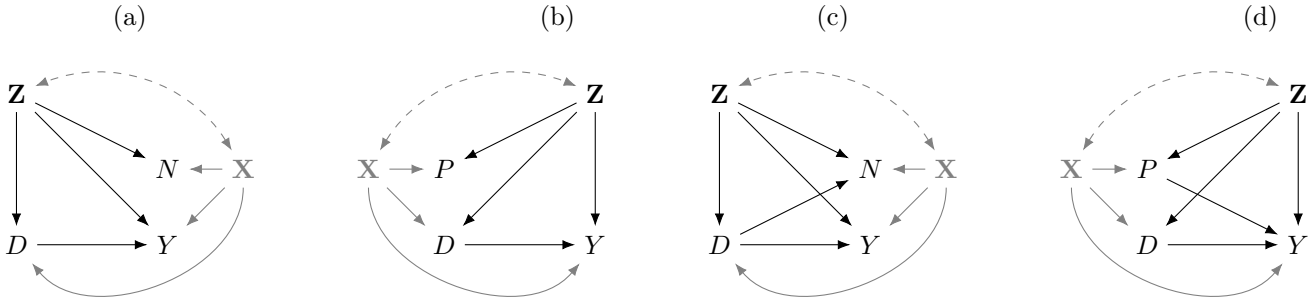
For a recent example of a placebo approach, see Ye et al. (2022). Lipsitch et al. (2010) provide an early discussion of the potential of using placebo treatments and outcomes. Arnold et al. (2016); Arnold and Ercumen (2016) discuss the potential to use placebos to address sample selection and measurement bias as well as bias in randomized experiments. Shi et al. (2020) review many established procedures and recent developments in this growing literature. An important recent area of research has been dubbed "proximal" causal inference. Tchetgen et al. (2020) provides an introduction. This line of work leverages "double-negative control designs" (Shi et al., 2020) in which both a placebo outcome and a placebo treatment are present. Under certain assumptions and conditions, papers like Miao et al. (2018, 2020); Mastouri et al. (2021) show how causal effects are non-parametrically identifiable using both of these placebos. While this work is very promising, it requires two placebos, certain null effect assumptions (i.e., "perfect" placebos), and regularity conditions.[2] Tchetgen et al. (2023a) provides a related approach for the case when only a single placebo is available. In this approach, the placebo variable must be associated with the control potential outcome and must be associated with the treatment only through the control potential outcome; there cannot be any causal relationship between the treatment and the placebo other than those that run through the control potential outcome for the primary outcome. This approach involves untestable assumptions. It also requires that "for any variation in the [primary outcome], there is corresponding variation in the [placebo]." This may preclude us from leveraging placebos in settings where they could be informative. For example, the assumptions are violated if the placebo has a common cause with the treatment that is not also a cause of the primary outcome or if the placebo plays a direct role in the treatment assignment mechanism. Additionally, if we have a placebo treatment, this approach may not apply. Our proposed approach, however, is applicable in these settings.

Additionally, Sofer et al. (2016) point out that placebo outcome approaches can be seen as generalizations of the familiar difference in differences (DID) framework. DID uses a pre-treatment verison of the outcome as a placebo outcome and relies heavily on a "parallel trends" assumption that the baseline trend in the outcome is the same over time for the group that gets treated and the group that does not. (Angrist and Pischke, 2008) Parallel trends is, in fact, an assumption of no unobserved time-varying confounding; meaning that DID approaches have not circumvented the problem of unobserved confounding. Sofer et al. (2016) also show how the parallel trends assumption is equivalent to an equi-confounding assumption. Recently, there have been several attempts to relax the parallel trends assumption within the context of DID or to improve on the current practice of testing for parallel trends in the pre-treatment period. For example, see Manski and Pepper (2018); Bilinski and Hatfield (2018); Freyaldenhoven et al. (2019); Keele et al. (2019); Ryan et al. (2019); Ye et al. (2020); Gibson and Zimmerman (2021); Rambachan and Roth (2021). While these are all useful contributions, they may not be as easy to use as applied researchers might hope and make assumptions that researchers might be interested in relaxing. Gibson and Zimmerman (2021) points out that violations of parallel trends can be considered using an application of the omitted variable bias framework from Cinelli and Hazlett (2020) in a first difference regression. However, this approach is limited to the DID setting in which the pre-treatment outcome is directly comparable to the post-treatment outcome. This approach (and others like Rambachan and Roth (2021)) also do not allow for a relaxation of the assumption that the treatment has no direct causal relationship

[1] Many applied studies forgo both sensitivity analysis and partial identification all together.

[2] In Section 2.3, we show a similar point identification result for regression when both a "perfect" placebo treatment and a "perfect" placebo outcome are available. We also relax the "double placebo" setting to allow for imperfect placebos.

Figure 1: Simple DAGs with a placebo outcome, N, and with a placebo treatment, P. D is treatment, Y is outcome, $\mathbf{Z}$ contains unobserved confounders, and $\mathbf{X}$ contains observed covariates.



with the pre-treatment outcome. Additionally, this approach requires that researchers observe the pre-treatment outcome and the post-treatment outcome for the same set of units; that is, it requires panel data. Tchetgen et al. (2023b) discusses an alternative to the parallel trends assumption which is a particular form of equi-confounding. Finally, the DID approach in general requires data on a pre-treatment version of the outcome, which may not always be available, making the placebo paradigm an attractive generalization.

Our contributions to this literature are simple. Relative to the proximal causal inference approaches, we relax the null effect assumptions that require "perfect" placebos, while also focusing on the case where a single placebo is available. We also discuss the case where two placebos are available. Relative to the literature on relaxing parallel trends in DID, we provide a framework that is applicable when the placebo is not a pre-treatment version of the outcome and relax the assumption that the treatment has no direct relationship with the pre-treatment version of the outcome. Our approach has a very simple and natural parameterization of unequal confouning (i.e., unparallel trends). We achieve all of this by partially identifying regression coefficients, but explore generalizations to partially linear and non-parametric settings in Appendix B. We note that regression estimates can be interpreted as best linear approximations to causal effects when all confounding is accounted for. (Berk et al., 2021) We show how the analysis of "omitted variable bias" in regression provides a simple yet flexible and powerful way to leverage information from placebo treatments and outcomes while relaxing assumptions of equi-confounding and perfect placebos. We provide a partial identification framework in which users can see the implications of any claim they choose to make about the relative strengths of confounding suffered by a placebo treatment/outcome compared to the true treatment-outcome relationship, while allowing for "imperfect placebos" (e.g. placebo treatments with some non-zero effect on the outcome, placebo outcomes somewhat affected by the treatment). These tools can be used with any choice of placebo treatment or outcome, observed in the same dataset or in a separate one as the main treatment-outcome relationship. In fact, our results are simple consequences of ordinary least squares and hold for any choice of placebo. Though researchers will want to chose placebos for which the relative level of confounding is easy to consider. We also discuss how to use our framework under various assumptions about the casual graph. Additionally, we frame the conventional difference-in-difference approach as a special case of our framework, with a strict equi-confounding (which is equivalent to parallel trends) assumption regarding the pre-treatment outcome's relationship to the treatment group indicator, to which our approach allows several useful relaxations. Finally, we demonstrate our approach with two applications that use the forthcoming `placeboLM` `R` package developed by the authors to facilitate the use of the proposed methods.

2 Leveraging placebos

Our proposed framework for using information from placebo treatments and outcomes relies the omitted variable bias analysis for linear regression. However, we also provide relaxations to this parametric approach in Appendix B, in which we illustrate similar methods adapted to partially linear and non-parametric estimation of treatment effects. In either case, we show how the omitted variable bias paradigm can be augmented to incorporate information from placebo treatments and outcomes to make causal progress with unequal levels of confounding for a placebo treatment/outcome compared to the true treatment-outcome relationship and with imperfect placebos. We emphasize this approach as a partial identification strategy, but it can equivalently be seen as providing sensitivity analysis for more traditional placebo approaches. We start by discussing placebo outcomes, then discuss placebo treatments, and finally discuss "double placebos."[3]

[3]As we mentioned, many placebos could be phrased as either a placebo outcome or a placebo treatment depending on the context. One framing may be more natural, depending on the placebo. For example, a pre-treatment version of the actual outcome will typically be best considered as a

2.1 Placebo outcomes

We start our discussion of placebo outcomes by leveraging the traditional omitted variable bias framework. (Angrist and Pischke, 2008; Cinelli and Hazlett, 2020; Hansen, 2022) Suppose that we want to estimate the "long" regression of Y on $D, \mathbf{X}$, and $\mathbf{Z}$ in Equation 1. In particular, we are interested in the coefficient $\beta_{Y\sim D|Z,X}$ as a measure of the causal effect of D on Y or as an approximation to the causal effect of D on Y. $\mathbf{X}$ and $\mathbf{Z}$ are each vectors of covariates. We can consider $Z_{(\text{Y.DX})} = \mathbf{Z}\beta_{Y\sim Z|D,X}$ to be the linear combination of the columns of $\mathbf{Z}$ that "de-confounds" the $D \to Y$ relationship in the same way that including $\mathbf{Z}$ in Equation 1 does.[4] That is $\beta_{Y\sim D|Z,X} = \beta_{Y\sim D|Z_{(\text{Y.DX})},X}$ from Equation 1 and 2.[5]

$$Y = \beta_{Y\sim D|Z,X} D + \mathbf{X}\beta_{Y\sim X|D,Z} + \mathbf{Z}\beta_{Y\sim Z|D,X} + \epsilon_l \tag{1}$$

$$Y = \beta_{Y\sim D|Z_{(\text{Y.DX})},X} D + \mathbf{X}\beta_{Y\sim X|D,Z_{(\text{Y.DX})}} + \beta_{Y\sim Z_{(\text{Y.DX})}|D,X} Z_{(\text{Y.DX})} + \epsilon_l \tag{2}$$

In practice, we typically do not observe all the covariates that would allow us to estimate or approximate the causal effect of D on Y by estimating this long regression. Therefore, let us assume that $\mathbf{Z}$ is unobserved. We will refer to these as "unobserved confounders." Since $\mathbf{Z}$ are unobserved and we must estimate the "short" regression of Y on D and $\mathbf{X}$ in Equation 3, rather than the desired long regression. $\mathbf{Z}$ is "omitted" from the regression, though we wish we could have included it.

$$Y = \beta_{Y\sim D|X} D + \mathbf{X}\beta_{Y\sim X|D} + \epsilon_s \tag{3}$$

Traditional omitted variable bias analysis tells us that $\beta_{Y\sim D|X} - \beta_{Y\sim D|Z_{(\text{Y.DX})},X}$ can be characterised as the product of two regression coefficients, as in Equation 4. $\beta_{Y\sim Z_{(\text{Y.DX})}|D,X}$ captures the relationship between Y and $Z_{(\text{Y.DX})}$ after linearly partialling out D and X. $\beta_{Z_{(\text{Y.DX})}\sim D|X}$ captures the relationship between D and $Z_{(\text{Y.DX})}$ after linearly partialling out X.

$$\text{bias}_{(\text{YD.X})} \triangleq \beta_{Y\sim D|X} - \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = \beta_{Y\sim Z_{(\text{Y.DX})}|D,X}\beta_{Z_{(\text{Y.DX})}\sim D|X} \tag{4}$$

These components alone are enough to build a powerful partial identification framework, as discussed in Cinelli and Hazlett (2020). But can more be done when there also exists an observed placebo outcome, N? If we think that the relationship between D and N suffers from a similar level of unobserved confounding as does the relationship between D and Y, then how might we leverage N to understand $\text{bias}_{(\text{YD.X})}$ and, hence, $\beta_{Y\sim D|Z,X} = \beta_{Y\sim D|Z_{(\text{Y.DX})},X}$? Consider a long and short regression with N as the outcome and with D as the treatment. The long regression of N on $D, \mathbf{X}$, and $\mathbf{Z}$ can be found in Equation 5. We can again consider $Z_{(\text{N.DX})} = \mathbf{Z}\beta_{N\sim Z|D,X}$ to be the linear combination of the columns of $\mathbf{Z}$ that "de-confounds" the $D \to N$ relationship in the same way that including $\mathbf{Z}$ in Equation 5 does. Note that we can consider $\mathbf{Z}$ to be a rich enough set of variables to de-confound both the Y, D and the N, D relationships. Thus, $\beta_{N\sim D|Z,X} = \beta_{N\sim D|Z_{(\text{N.DX})},X}$ from Equation 5 and 6. It is important to note that $Z_{(\text{N.DX})}$ will not equal $Z_{(\text{Y.DX})}$ when $\beta_{N\sim Z|D,X} \neq \beta_{Y\sim Z|D,X}$. The short regression of N on just D and $\mathbf{X}$ can be found in Equation 7. As with Y, we are only able to consider estimating the short regression for N, as $\mathbf{Z}$ is unobserved. Here too $\beta_{N\sim D|X} - \beta_{N\sim D|Z_{(\text{N.DX})},X}$ can be characterised as in Equation 8.

$$N = \beta_{N\sim D|Z,X} D + \mathbf{X}\beta_{N\sim X|D,Z} + \mathbf{Z}\beta_{N\sim Z|D,X} + \xi_l \tag{5}$$

$$N = \beta_{N\sim D|Z_{(\text{N.DX})},X} D + \mathbf{X}\beta_{Y\sim X|D,Z_{(\text{N.DX})}} + \beta_{N\sim Z_{(\text{N.DX})}|D,X} Z_{(\text{N.DX})} + \xi_l \tag{6}$$

$$N = \beta_{N\sim D|X} D + \mathbf{X}\beta_{N\sim X|D} + \xi_s \tag{7}$$

$$\text{bias}_{(\text{ND.X})} \triangleq \beta_{N\sim D|X} - \beta_{N\sim D|Z_{(\text{N.DX})},X} = \beta_{N\sim Z_{(\text{N.DX})}|D,X}\beta_{Z_{(\text{N.DX})}\sim D|X} \tag{8}$$

From Equations 4 and 8, it is easy to see how to leverage information about the placebo outcome, N. There is some $m \in \mathbb{R}$ such that $\text{bias}_{(\text{YD.X})} = m \times \text{bias}_{(\text{ND.X})}$. Thus, we can write Equations 9 and 10; in these expressions, we substitute $\beta_{N\sim D|Z,X}$ for $\beta_{N\sim D|Z_{(\text{N.DX})},X}$ and $\beta_{Y\sim D|Z,X}$ for $\beta_{Y\sim D|Z_{(\text{Y.DX})},X}$, since they are equal. This can be thought of as using the placebo outcome to re-express the confounding path $D \leftarrow Z \to Y$ in Figure 1(a).

$$\text{bias}_{(\text{YD.X})} = \beta_{Y\sim D|X} - \beta_{Y\sim D|Z,X} = m \times \left(\beta_{N\sim D|X} - \beta_{N\sim D|Z,X}\right) \tag{9}$$

$$\iff \beta_{Y\sim D|Z,X} = \beta_{Y\sim D|X} - m \times \left(\beta_{N\sim D|X} - \beta_{N\sim D|Z,X}\right) \tag{10}$$

placebo outcome, since it can easily be compared to the actual outcome.

[4] Note that in this section and others, we will refer to including $\mathbf{Z}$ in the long regression as "de-confounding". Readers should recognize that, in reality, we are only partialling out the linear relationships with $\mathbf{Z}$. This is, therefore, a linear approximation to full non-parametric elimination of confounding from $\mathbf{Z}$.

[5] Indeed $R^2_{Y\sim Z_{(\text{Y.DX})}|D,X}$ will also equal $R^2_{Y\sim Z|D,X}$. But $R^2_{D\sim Z_{(\text{Y.DX})}|X} \leq R^2_{D\sim Z|X}$, since $Z_{(\text{Y.DX})}$ is chosen to maximize the R^2 between $\mathbf{Z}$ and Y not $\mathbf{Z}$ and D. (Cinelli and Hazlett, 2020)

Equation 9 is an expression for $\text{bias}_{(\text{YD.X})}$ in terms of $m \triangleq \frac{\beta_{Y\sim Z_{(\text{Y.DX})}|D,X}\beta_{Z_{(\text{Y.DX})}\sim D|X}}{\beta_{N\sim Z_{(\text{N.DX})}|D,X}\beta_{Z_{(\text{N.DX})}\sim D|X}} = \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{ND.X})}}$, $\beta_{N\sim D|Z_{(\text{N.DX})},X}$, and $\beta_{N\sim D|X}$. We have a similar expression for $\beta_{Y\sim D|Z_{(\text{Y.DX})},X}$ in Equation 10, our target parameter, that also includes $\beta_{Y\sim D|X}$. We can estimate $\beta_{N\sim D|X}$ and $\beta_{Y\sim D|X}$ from the data.[6] The remaining terms in these expressions can be treated as parameters that can be specified by investigators. Thus, we can use Equation 10 for partial identification of $\beta_{Y\sim D|Z,X}$ by establishing plausible ranges of values for m and $\beta_{N\sim D|Z,X}$. m captures the level of relative confounding. This can be thought of as comparing the strength of the $D \leftarrow Z \rightarrow Y$ relationship to the $D \leftarrow Z \rightarrow N$ relationship in Figure 1(a). $\beta_{N\sim D|Z,X} = \beta_{N\sim D|Z_{(\text{N.DX})},X}$ captures any direct causal relationship between D and N, like the $D \rightarrow N$ path in Figure 1(c). Therefore, this $\beta_{N\sim D|Z,X}$ parameter allows for "imperfect" placebo outcomes that are directly affected by the treatment. Since both N and D are observed variables, their relationship conditional on Z, X should be able to be specified as a regression coefficient despite scale considerations. In many cases differing scales for N and Y might lead to difficulty in interpreting or reasoning about the values for m. Can we do better than this?

Re-expressing m Let us consider a reparameterization of m in terms of scale-independent partial-correlation parameters, rather than regression coefficients. This will allow us to circumvent the scale issues we just mentioned. To do this, we use the reparameterizations of omitted variable bias similar to that discussed in Cinelli and Hazlett (2020) shown in Equation 11.

$$\begin{aligned}\text{bias}_{(\text{YD.X})} &= \beta_{Y\sim D|X} - \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}(\text{SD}(Y^{\perp D,X})/\text{SD}(D^{\perp X})) \\ \text{bias}_{(\text{ND.X})} &= \beta_{N\sim D|X} - \beta_{N\sim D|Z_{(\text{N.DX})},X} = R_{N\sim Z_{(\text{N.DX})}|D,X} f_{D\sim Z_{(\text{N.DX})}|X}(\text{SD}(N^{\perp D,X})/\text{SD}(D^{\perp X}))\end{aligned} \tag{11}$$

$$m = \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{ND.X})}} = \frac{R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}}{R_{N\sim Z_{(\text{N.DX})}|D,X} f_{D\sim Z_{(\text{N.DX})}|X}} \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(N^{\perp D,X})} = k \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(N^{\perp D,X})} \tag{12}$$

$$\begin{aligned} k \triangleq \frac{R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}}{R_{N\sim Z_{(\text{N.DX})}|D,X} f_{D\sim Z_{(\text{N.DX})}|X}} &= \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{ND.X})}} \times \frac{1}{\text{SF}} \\ &= \frac{\text{bias}_{(\text{YD.X})} \times (1/\text{SF})}{\text{bias}_{(\text{ND.X})}} = \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{ND.X})} \times \text{SF}} \text{ where } \text{SF} = \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(N^{\perp D,X})} \end{aligned} \tag{13}$$

Equation 12 provides an expression for m in terms of k[7] and SF. SF ("scale factor") captures the scale issues and deals only with observables. The numerator and denominator are moments of the residuals from the two short regressions. k is a scaled version of the ratio of $\text{bias}_{(\text{YD.X})}$ over $\text{bias}_{(\text{ND.X})}$. This is a scale independent measure of the level of confounding of the Y, D relationship relative to the level of confounding of the N, D relationship. Alternatively put, it is the level of unobserved confounding in the Y, D relationship as a percentage of the level of unobserved confounding in the N, D relationship, after resclaing to account for possible differences in the scale of Y and N. If $\text{SD}(N^{\perp D,X}) = \text{SD}(Y^{\perp D,X})$, then k is simply the ratio of biases. Quantities like $R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}$ and $R_{N\sim Z_{(\text{N.DX})}|D,X} f_{D\sim Z_{(\text{N.DX})}|X}$ are called "bias factors" in Cinelli and Hazlett (2020). $f_{D\sim Z|X}$ is Cohen's f which equals $\frac{R}{\sqrt{1-R^2}}$. We can also re-write the bias-factors as $R_{Y\sim Z|D,X} f_{D\sim Z|X} = f_{Y\sim D|X} - f_{Y\sim D|Z,X}\sqrt{\frac{1-R^2_{Y\sim Z|D,X}}{1-R^2_{D\sim Z|X}}}$. See Appendix C for the derivation. Equations 9 and 10 can then be rewritten as Equation 14 and 15.

$$\text{bias}_{(\text{YD.X})} = k \times \left(\beta_{N\sim D|X} - \beta_{N\sim D|Z,X}\right) \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(N^{\perp D,X})} \tag{14}$$

$$\beta_{Y\sim D|Z,X} = \beta_{Y\sim D|X} - k \times \left(\beta_{N\sim D|X} - \beta_{N\sim D|Z,X}\right) \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(N^{\perp D,X})} \tag{15}$$

[6] Note that we do not need the placebo outcome to be observed in the same dataset or for the same units as the primary outcome to estimate $\beta_{N\sim D|X}$ and $\beta_{Y\sim D|X}$.

[7] If partial R^2 are more intuitive, users can substitute in $k = \text{sign}(k) \times \sqrt{k^2}$ and consider $\text{sign}(k) = \text{sign}\left(\frac{R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}}{R_{N\sim Z_{(\text{N.DX})}|D,X} f_{D\sim Z_{(\text{N.DX})}|X}}\right)$ and $k^2 = \frac{R^2_{Y\sim Z_{(\text{Y.DX})}|D,X} f^2_{D\sim Z_{(\text{Y.DX})}|X}}{R^2_{N\sim Z_{(\text{N.DX})}|D,X} f^2_{D\sim Z_{(\text{N.DX})}|X}}$. The denominator of k cannot equal zero. This will be the case for any useful placebo outcome. Otherwise, the placebo outcome will not have a confounded relationship with the treatment and not have any useful information that we can leverage to understand the confounding of the relationship between the treatment and the actual outcome.

Estimator and partial identification Equation 15 provides us with a moment estimator for $\beta_{Y\sim D|Z,X}$ in Equation 16.[8][9]

$$\hat{\beta}_{Y\sim D|Z,X} = \hat{\beta}_{Y\sim D|X} - k \times \left(\hat{\beta}_{N\sim D|X} - \beta_{N\sim D|Z,X}\right) \times \frac{\|Y^{\perp D,X}\|_2}{\|N^{\perp D,X}\|_2} \tag{16}$$

Given k and $\beta_{N\sim D|Z,X}$, the estimator $\hat{\beta}_{Y\sim D|Z,X}$ from Equation 16 is consistent for $\beta_{Y\sim D|Z,X}$ and asymptotically normal. We will discuss estimating standard errors in Section 2.4. Using Equation 16 in a partial identification framework, we can reason about plausible ranges of values for $\beta_{N\sim D|Z,X}$ and k to arrive at a range for plausible estimates of $\beta_{Y\sim D|Z,X}$. In this way, we can partially identify $\beta_{Y\sim D|Z,X}$. In the special cases when there is only a single omitted variable, $Z_{(\text{N.DX})} = Z_{(\text{Y.DX})}$, or $f_{D\sim Z_{(\text{N.DX})}|X} = f_{D\sim Z_{(\text{Y.DX})}|X}$, then k can also be written as $k = \frac{R_{Y\sim Z|D,X}}{R_{N\sim Z|D,X}}$. This can be thought of as comparing the strength of the $Z \to Y$ "arm" of the confounding path $D \leftarrow Z \to Y$ to the relative strength of the $Z \to N$ "arm" of the confounding path $D \leftarrow Z \to N$ in Figure 1(a).

2.2 Placebo treatments

It is possible to use similar machinery to leverage imperfect placebo treatments rather than imperfect placebo outcomes in a partial identification framework for effects of actual treatments on outcomes. Suppose that we want to run the long regression for Y on D (Equation 1) but, again, $\mathbf{Z}$ is an unobserved vector of variables that we would have liked to include in the regression but cannot. Suppose also that we have a placebo treatment, P, for which we believe the confounding of P and Y from $\mathbf{Z}$ is similar to the confounding of D and Y from $\mathbf{Z}$. We might consider running two separate regressions as we did for placebo outcomes. However, that approach is not easily adapted to imperfect placebo treatments (i.e., those for which there is a direct relationship between P and Y).[10] Instead, suppose that P is not a descendant of D, D is not a descendant of P, P is a neutral control for the effect of D on Y, and P is a neutral control for D on Y, conditional on Z and X.[11] Further, suppose that we want to run the "long" regression of Y on $D, P, \mathbf{X}$, and $\mathbf{Z}$ in Equation 17. In particular, we are interested in the coefficient $\beta_{Y\sim D|P,Z,X}$ as a measure of the causal effect of D on Y or as an approximation to the causal effect of D on Y. However, since $\mathbf{Z}$ is unobserved, we must run the "short" regression of Y on D, P and $\mathbf{X}$ in Equation 18, rather than the desired long regression.

$$Y = \beta_{Y\sim D|P,Z,X} D + \beta_{Y\sim P|D,Z,X} P + \mathbf{X}\beta_{Y\sim X|D,P,Z} + \mathbf{Z}\beta_{Y\sim Z|D,P,X} + \epsilon_l \tag{17}$$

$$Y = \beta_{Y\sim D|P,X} D + \beta_{Y\sim P|D,X} P + \mathbf{X}\beta_{Y\sim X|D,P} + \epsilon_s \tag{18}$$

Following a similar omitted variables based approach as we took for placebo outcomes (see Equation 19), we can arrive at similar partial identification expressions for the placebo treatment case.

$$\begin{aligned}
&\text{bias}_{(\text{YD.PX})} = \beta_{Y\sim D|P,X} - \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X} (\text{SD}(Y^{\perp D,P,X})/\text{SD}(D^{\perp P,X})) \\
&\text{bias}_{(\text{YP.DX})} = \beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X} = R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{P\sim Z_{(\text{Y.DPX})}|D,X} (\text{SD}(Y^{\perp D,P,X})/\text{SD}(P^{\perp D,X})) \\
&\text{bias}_{(\text{YD.PX})} = m \times \text{bias}_{(\text{YP.DX})} \implies \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim D|P,X} - m \times (\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X}) \\
&m = \frac{\text{bias}_{(\text{YD.PX})}}{\text{bias}_{(\text{YP.DX})}} = \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{P\sim Z_{(\text{Y.DPX})}|D,X}} \times \frac{\text{SD}(P^{\perp D,X})}{\text{SD}(D^{\perp P,X})} = k \times \frac{\text{SD}(P^{\perp D,X})}{\text{SD}(D^{\perp P,X})}
\end{aligned} \tag{19}$$

$$\text{bias}_{(\text{YD.PX})} \stackrel{\Delta}{=} \beta_{Y\sim D|P,X} - \beta_{Y\sim D|P,Z,X} = k \times \left(\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z,X}\right) \times \frac{\text{SD}(P^{\perp D,X})}{\text{SD}(D^{\perp P,X})} \tag{20}$$

[8] We follow a similar presentation of our estimators and approaches to inference as in Zhang and Ding (2022).

[9] It is also possible to show that $\frac{\widehat{\text{SD}}(Y^{\perp D,X})}{\widehat{\text{SD}}(N^{\perp D,X})} = \frac{\left[\frac{\widehat{\text{SD}}(Y^{\perp D,X})}{\widehat{\text{SD}}(D^{\perp X})}\right]}{\left[\frac{\widehat{\text{SD}}(N^{\perp D,X})}{\widehat{\text{SD}}(D^{\perp X})}\right]} = \frac{\text{se}(\hat{\beta}_{Y\sim D|X})}{\text{se}(\hat{\beta}_{N\sim D|X})} \times \sqrt{\frac{\text{df}_Y}{\text{df}_N}}$, which is an expression containing only commonly reported regression summary statistics from the two short regressions that are estimable from the observed data. $\text{se}(\hat{\beta}_{Y\sim D|X})$ and $\text{se}(\hat{\beta}_{N\sim D|X})$ are the estimated standard errors from standard regression estimation of the short regressions. df_Y and df_N are the degrees of freedom from the estimated short regressions.

[10] Suppose we have a causal model like Figure 1(d). $\beta_{Y\sim Z|P,X}$ (or $R_{Y\sim Z|P,X}$) will capture the causal relationship $D \to Y$, since the path $Z \to D \to Y$ will remain open conditional on P and X alone. Expressing the bias in the regression coefficient on D from the short regression using a two regression approach like we used for placebo outcomes would, therefore, require us to reason about the causal effect of interest itself, which is not useful.

[11] We require that D and P are neutral controls for each other in that they do not change the regression coefficients for the outcome Y when we include them in the same regression. This will simplify interpretation. If a user is concerned that there could be a path opened or closed by conditioning on D or P, we suggest drawing the DAG to check. See Section 3 for further discussion.

$$\beta_{Y\sim D|P,Z,X} = \beta_{Y\sim D|P,X} - k \times \left(\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z,X}\right) \times \frac{\text{SD}(P^{\perp D,X})}{\text{SD}(D^{\perp P,X})} \tag{21}$$

Equations 20 and 21 contain the results. All components of these expressions except for $k \triangleq \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{P\sim Z_{(\text{Y.DPX})}|D,X}}$ and $\beta_{Y\sim P|D,Z,X}$ can be estimated from the data. We can use this for partial identification by considering plausible ranges of values for k and $\beta_{Y\sim P|D,Z,X}$. $\beta_{Y\sim P|D,Z,X}$ captures any direct causal relationship between P and Y. Since both P and Y are observed variables, their relationship conditional on D, Z, X should be able to be specified as the regression coefficient $\beta_{Y\sim P|D,Z,X}$ despite scale considerations. As with placebo outcomes, k captures the relative levels of confounding. Note that we can compare the level of confounding of the $P \to Y$ relationship from $\mathbf{Z}$ to the level of confounding of the $D \to Y$ relationship from $\mathbf{Z}$ by using the same linear combination, $Z_{(\text{Y.DPX})} = \mathbf{Z}\beta_{Y\sim Z|D,P,X}$. This is because in the placebo treatment approach we run a single regression with Y as the outcome that includes both D and P on the right-hand side. Since we assume that $\mathbf{Z}$ contains a rich enough set of unobserved variables to de-confound both $P \to Y$ and $D \to Y$, $Z_{(\text{Y.DPX})}$ is also sufficient for this. k is the ratio of the level of total confounding of $D \to Y$ from $D \leftarrow Z \to Y$ to the level of total confounding of $P \to Y$ from $P \leftarrow Z \to Y$, after re-scaling one of these biases to account for scale differences between P and D.[12]

Estimator and partial identification Equation 21 provides us with a moment estimator for $\beta_{Y\sim D|P,Z,X}$ in Equation 22.[13]

$$\hat{\beta}_{Y\sim D|P,Z,X} = \hat{\beta}_{Y\sim D|P,X} - k \times \left(\hat{\beta}_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z,X}\right) \times \frac{\|P^{\perp D,X}\|_2}{\|D^{\perp P,X}\|_2} \tag{22}$$

Given k and $\beta_{Y\sim P|D,Z,X}$, the estimator $\hat{\beta}_{Y\sim D|P,Z,X}$ from Equation 22 is consistent for $\beta_{Y\sim D|P,Z,X}$ and asymptotically normal. We discuss computing standard errors in Section 2.4. Using Equation 22 in a partial identification framework, we can reason about plausible ranges of values for $\beta_{Y\sim P|D,Z,X}$ and k to arrive at a range of plausible estimates of $\beta_{Y\sim D|P,Z,X}$. In the special case when there is only a single omitted variable, then k can also be written as $k = \frac{f_{D\sim Z|P,X}}{f_{P\sim Z|D,X}}$. This can be thought of as comparing the strength of the $Z \to D$ "arm" of the confounding path $D \leftarrow Z \to Y$ to the relative strength of the $Z \to P$ "arm" of the confounding path $P \leftarrow Z \to Y$ in Figure 1(b).

2.3 Double placebos

Finally, we consider the setting in which we have two placebos. Suppose we are interested in $\beta_{Y\sim D|Z,X}$ from the "long" regression in Equation 1. However, Z is unobserved and so we must only consider estimating "short" regression in Equation 3. Further, suppose we are in a setting like Figure 2 - that is, we observe both a placebo outcome (N) and a placebo treatment (P). In such a "double placebo" setting, we can also develop a partial identification framework. Consider the following two sets of short and long regressions. Again, $\mathbf{Z}$ is an unobserved confounder. So we can only consider estimating the short regressions (i.e., Equations 23b and 23d). $Z_{(\text{Y.DPX})} = \mathbf{Z}\beta_{Y\sim Z|D,P,X}$ is the linear combination that de-confounds the Y, D and the Y, P relationship. $Z_{(\text{N.DPX})} = \mathbf{Z}\beta_{N\sim Z|D,P,X}$ is the linear combination that de-confounds the N, D and the N, P relationship. We choose $\mathbf{Z}$ to be a rich enough set of variables to allow for this interpretation.

$$Y = \beta_{Y\sim D|P,Z,X} D + \beta_{Y\sim P|D,Z,X} P + \mathbf{X}\beta_{Y\sim X|D,P,Z} + \mathbf{Z}\beta_{Y\sim Z|D,P,X} + \epsilon_{y,l} \tag{23a}$$
$$Y = \beta_{Y\sim D|P,X} D + \beta_{Y\sim P|D,X} P + \mathbf{X}\beta_{Y\sim X|D,P} + \epsilon_{y,s} \tag{23b}$$
$$N = \beta_{N\sim D|P,Z,X} D + \beta_{N\sim P|D,Z,X} P + \mathbf{X}\beta_{N\sim X|D,P,Z} + \mathbf{Z}\beta_{N\sim Z|D,P,X} + \epsilon_{n,l} \tag{23c}$$
$$N = \beta_{N\sim D|P,X} D + \beta_{N\sim P|D,X} P + \mathbf{X}\beta_{N\sim X|D,P} + \epsilon_{n,s} \tag{23d}$$

In Appendix A, we detail a similar omitted variables bias based approach for this setting. The approach results in the expression for $\beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim D|P,Z,X}$ in Equation 24.

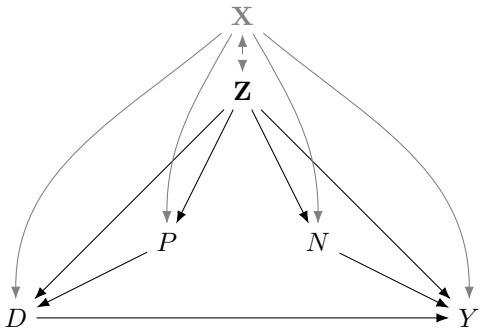
$$\beta_{Y\sim D|P,Z,X} = \beta_{Y\sim D|P,X} - k_{\left(\frac{\text{YD}}{\text{YP}}\right)} \times k_{\left(\frac{\text{NP}}{\text{ND}}\right)} \times \frac{(\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z,X})(\beta_{N\sim D|P,X} - \beta_{N\sim D|P,Z,X})}{(\beta_{N\sim P|D,X} - \beta_{N\sim P|D,Z,X})} \tag{24}$$

[12] Note that we do not cancel terms in $k = \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{P\sim Z_{(\text{Y.DPX})}|D,X}}$ since this would force us to reason about how $f_{D\sim Z_{(\text{Y.DPX})}|P,X}$ differs from $f_{D\sim\mathbf{Z}|P,X}$ and how $f_{P\sim Z_{(\text{Y.DPX})}|D,X}$ differs from $f_{P\sim\mathbf{Z}|D,X}$. It is more intuitive to reason about relative levels of bias, which does not require us to directly consider those differences.

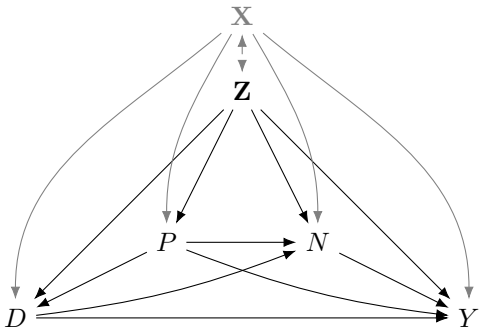
[13] It is also possible to show that $\frac{\widehat{\text{SD}}(P^{\perp D,X})}{\widehat{\text{SD}}(D^{\perp P,X})} = \frac{\widehat{\text{SD}}(P^{\perp D,X})}{\widehat{\text{SD}}(D^{\perp P,X})} \times \frac{\widehat{\text{SD}}(Y^{\perp D,P,X})}{\widehat{\text{SD}}(Y^{\perp D,P,X})} = \frac{\text{se}(\hat{\beta}_{Y\sim D|P,X})}{\text{se}(\hat{\beta}_{Y\sim P|D,X})} \times \sqrt{\frac{\text{df}_D}{\text{df}_P}}$, which is an expression containing only commonly reported regression summary statistics from the short regression that is estimable from the observed data. $\text{se}(\hat{\beta}_{Y\sim D|P,X})$ and $\text{se}(\hat{\beta}_{Y\sim P|D,X})$ are the estimated standard errors from standard regression estimation of the short regression. df_Y and df_N are the degrees of freedom from the estimated short regression, these will be equal in this case.

Figure 2: A single causal graph with both a placebo outcome, N, and with a placebo treatment, P. D is treatment, Y is outcome, and $\mathbf{Z}$ contains unobserved confounders. $\mathbf{X}$ contains observed covariates.

(a) Perfect Double Placebo



(a) Imperfect Double Placebo



Equation 24 provides us with a moment estimator for $\beta_{Y\sim D|P,Z,X}$ in Equation 25.

$$\hat{\beta}_{Y\sim D|P,Z,X} = \hat{\beta}_{Y\sim D|P,X} - k_{\left(\frac{\text{YD}}{\text{YP}}\right)} \times k_{\left(\frac{\text{NP}}{\text{ND}}\right)} \times \frac{(\hat{\beta}_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z,X})(\hat{\beta}_{N\sim D|P,X} - \beta_{N\sim D|P,Z,X})}{(\hat{\beta}_{N\sim P|D,X} - \beta_{N\sim P|D,Z,X})} \tag{25}$$

Given $k_{\left(\frac{\text{YD}}{\text{YP}}\right)}, k_{\left(\frac{\text{NP}}{\text{ND}}\right)}, \beta_{Y\sim P|D,Z,X}, \beta_{N\sim D|P,Z,X}$, and $\beta_{N\sim P|D,Z,X}$, the estimator $\hat{\beta}_{Y\sim D|P,Z,X}$ from Equation 25 is consistent for $\beta_{Y\sim D|P,Z,X}$ and asymptotically normal. We discuss computing standard errors in Section 2.4. Using Equation 25 in a partial identification framework, we can reason about plausible ranges of values for $k_{\left(\frac{\text{YD}}{\text{YP}}\right)}, k_{\left(\frac{\text{NP}}{\text{ND}}\right)}, \beta_{Y\sim P|D,Z,X}, \beta_{N\sim D|P,Z,X}$, and $\beta_{N\sim P|D,Z,X}$ to arrive at a range of plausible estimates of $\beta_{Y\sim D|P,Z,X}$.

In Figure 2(b), the partial identification parameters for double placebos capture paths as follows: $\beta_{Y\sim P|D,Z,X} = \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X}$ captures $P \to Y$ and $P \to N \to Y$ (the effect of P on Y conditional on D); $\beta_{N\sim D|P,Z,X} = \beta_{N\sim D|P,Z_{(\text{N.DPX})},X}$ captures $D \to N$; $\beta_{N\sim P|D,Z,X} = \beta_{N\sim P|D,Z_{(\text{N.DPX})},X}$ captures $P \to N$. We can disentangle the $P \to Y$ and $P \to N \to Y$ components of $\beta_{Y\sim P|D,Z,X}$ for Figure 2(b) by using the following omitted variable expression: $\beta_{Y\sim P|D,Z,X} = \beta_{Y\sim P|N,D,Z,X} + \beta_{Y\sim N|P,D,Z,X}\beta_{N\sim P|D,Z,X}$. $\beta_{Y\sim P|N,D,Z,X}$ captures $P \to Y$; $\beta_{Y\sim N|P,D,Z,X}$ captures $N \to Y$; $\beta_{N\sim P|D,Z,X}$ captures $P \to N$ and already a partial identification parameter. $k_{\left(\frac{\text{YD}}{\text{YP}}\right)} \times k_{\left(\frac{\text{NP}}{\text{ND}}\right)} = \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{Z_{(\text{Y.DPX})}\sim D|P,X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{Z_{(\text{Y.DPX})}\sim P|D,X}} \times \frac{R_{N\sim Z_{(\text{N.DPX})}|D,P,X} f_{Z_{(\text{N.DPX})}\sim P|D,X}}{R_{N\sim Z_{(\text{N.DPX})}|D,P,X} f_{Z_{(\text{N.DPX})}\sim D|P,X}}$. $k_{\left(\frac{\text{YD}}{\text{YP}}\right)}$ captures the scaled ratio of the level of confounding of the Y, D relationship (conditional on P) to the level of confounding of the Y, P relationship (conditional on D), after re-scaling one of these biases, or the ratio of bias factors. $k_{\left(\frac{\text{NP}}{\text{ND}}\right)}$ captures the scaled ratio of the level of confounding of the N, P relationship (conditional on D) to the level of confounding of the N, D relationship (conditional on P), after re-scaling one of these biases, or the ratio of bias factors.

This type of expression can be used in a wide variety of settings. Though the interpretation of the partial identification parameters and of $\beta_{Y\sim D|P,Z,X}$ will depend on the underlying causal model being considered. See Section 3 for more details. With two imperfect placebos, users have a choice between the "single placebo" approaches and the "double placebo" approach. If we are in a setting like in Figure 2(a), where "perfect" placebos are available, then $\beta_{Y\sim P|D,Z,X}, \beta_{N\sim D|P,Z,X}$, and $\beta_{N\sim P|D,Z,X}$ will all equal zero. If it is reasonable to assume that either we have a single unobserved variable, $Z_{(\text{Y.DPX})} = Z_{(\text{N.DPX})}$, or $R_{Z_{(\text{Y.DPX})}\sim D|P,X} = R_{Z_{(\text{N.DPX})}\sim D|P,X}$ and $R_{Z_{(\text{Y.DPX})}\sim P|D,X} = R_{Z_{(\text{N.DPX})}\sim P|D,X}$, then $k_{\left(\frac{\text{YD}}{\text{YP}}\right)} \times k_{\left(\frac{\text{NP}}{\text{ND}}\right)} = 1$.[14] If we are in both of these settings at the same time, then we can point identify $\beta_{Y\sim D|Z,X}$ using observed data. This is similar to the proximal causal inference point identification results (Miao et al., 2020). See Liu et al. (2022) for an application of a double placebo approach.

2.4 Standard Errors

For fixed values for the partial identification parameters, the moment estimators in Equations 16, 22, and 25 are consistent for the target regression coefficient on the treatment from the "long" regression and are asymptotically normal. We can use the non-parametric bootstrap to estimate standard errors. See Zhang and Ding (2022); Freidling and Zhao (2023) for similar partial identification approaches that rely on the non-parametric bootstrap for standard errors.

[14] If these relationships approximately hold, we could also consider a narrow range of values for $k_{\left(\frac{\text{YD}}{\text{YP}}\right)} \times k_{\left(\frac{\text{NP}}{\text{ND}}\right)}$ close to 1.

3 A taxonomy of causal graphs

How do we know if we have a useful placebo? How do we know if it is better to treat a placebo as a placebo outcome or a placebo treatment? Are there settings in which the above approaches should not be used? Are there settings in which the proposed methods can be used for problems other than unobserved confounding? We will address these questions by developing a taxonomy of causal graphs when a single placebo is available and when two are available. For introductions to graphical causal models see Pearl (2009), Richardson and Robins (2013), Matthay and Glymour (2020), and Cinelli et al. (2022). We re-emphasize the Necker cube quality of placebos. Two causal graphs, each with a placebo, that have the same structure may invite different interpretations for whether the placebo is a placebo treatment or outcome. These choices are context specific. Sometimes the placebo may feel more like an outcome, while in other contexts the placebo feels more like a treatment.

3.1 Single Placebos

We start by listing all possible causal graphs when a single placebo is observed in Figure 3. We assume that all relevant unobserved common causes are bundled into the vector of unobserved variables $\mathbf{Z}$. Observed covariates $\mathbf{X}$ are omitted from the graphs for clarity. We refer to the placebo as P. Figure 3(a) contains a template from which the rest of the graphs in Figure 3 are derived. The two dashed edges can point in either direction or not exist.[15] Different approaches should be used when dealing with each causal structure.[16]

Placebos Figure 3(b) is the simplest setting in which P has no direct relationship with either the treatment, D, or the outcome, Y. That is, P is a "perfect" placebo. There are two options for how we might proceed in this case. We can look at P as a placebo outcome and follow the approach laid out in Section 2.1, or we can look at P as a placebo treatment and follow the approach laid out in Section 2.2. In either case, since P is a perfect placebo, we will only have one partial identification parameter that captures the relative levels of confounding.

Figure 3(c) has the addition of $D \to P$. Since the placebo is post-treatment in this setting, it is likely that we will want to treat the placebo as an imperfect placebo outcome. Therefore, we can follow the approach laid out in Section 2.1 [17] Figure 3(d) has $P \to Y$. Here there are also two options. The first would be to treat P as an imperfect placebo treatment and follow the approach laid out in Section 2.2. The second would be to view P as a placebo outcome, as in the case that P is a pre-treatment version of the outcome. In this setting, we can employ the approach laid out in Section 2.1 or Equation 26. The key in using Equation 26 is the inclusion of P in the regression of Y on D.[18]

$$
\begin{aligned}
&\text{bias}_{(\text{YD.PX})} = \beta_{Y\sim D|P,X} - \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}(\text{SD}(Y^{\perp D,P,X})/\text{SD}(D^{\perp P,X}))\\
&\text{bias}_{(\text{PD.X})} = \beta_{P\sim D|X} - \beta_{P\sim D|Z_{(\text{P.DX})},X} = R_{P\sim Z_{(\text{P.DX})}|D,X} f_{D\sim Z_{(\text{P.DX})}|X}(\text{SD}(P^{\perp D,X})/\text{SD}(D^{\perp X}))\\
&\text{bias}_{(\text{YD.PX})} = m \times \text{bias}_{(\text{PD.X})} \implies \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim D|P,X} - m \times (\beta_{P\sim D|X} - \beta_{P\sim D|Z_{(\text{P.DX})},X})\\
&m = \frac{\text{bias}_{(\text{YD.PX})}}{\text{bias}_{(\text{PD.X})}} = \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}}{R_{P\sim Z_{(\text{P.DX})}|D,X} f_{D\sim Z_{(\text{P.DX})}|X}} \times \frac{\text{SD}(Y^{\perp D,P,X})}{\text{SD}(D^{\perp P,X})} \times \frac{\text{SD}(D^{\perp X})}{\text{SD}(P^{\perp D,X})} = k \times \frac{\text{SD}(Y^{\perp D,P,X})}{\text{SD}(D^{\perp P,X})} \times \frac{\text{SD}(D^{\perp X})}{\text{SD}(P^{\perp D,X})}\\
&\therefore \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim D|P,X} - k \times (\beta_{P\sim D|X} - \beta_{P\sim D|Z_{(\text{P.DX})},X}) \times \frac{\text{SD}(Y^{\perp D,P,X})}{\text{SD}(D^{\perp P,X})} \times \frac{\text{SD}(D^{\perp X})}{\text{SD}(P^{\perp D,X})}
\end{aligned}
\tag{26}
$$

In Equation 26, $k \triangleq \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}}{R_{P\sim Z_{(\text{P.DX})}|D,X} f_{D\sim Z_{(\text{P.DX})}|X}}$ is the ratio of the level of confounding D and Y, conditional on P, to the level of confounding D and P, after re-scaling one of these biases to account for scale differences. Since we included P in the regression of Y on D, we see that the confounding of D and Y does not include the $Z \to P \to Y$ path, possibly making comparison with the confounding of D and P easier. $\beta_{P\sim D|Z_{(\text{P.DX})},X} = \beta_{P\sim D|Z,X}$ measures the causal effect of D on P, which based on Figure 3(d) should be zero.

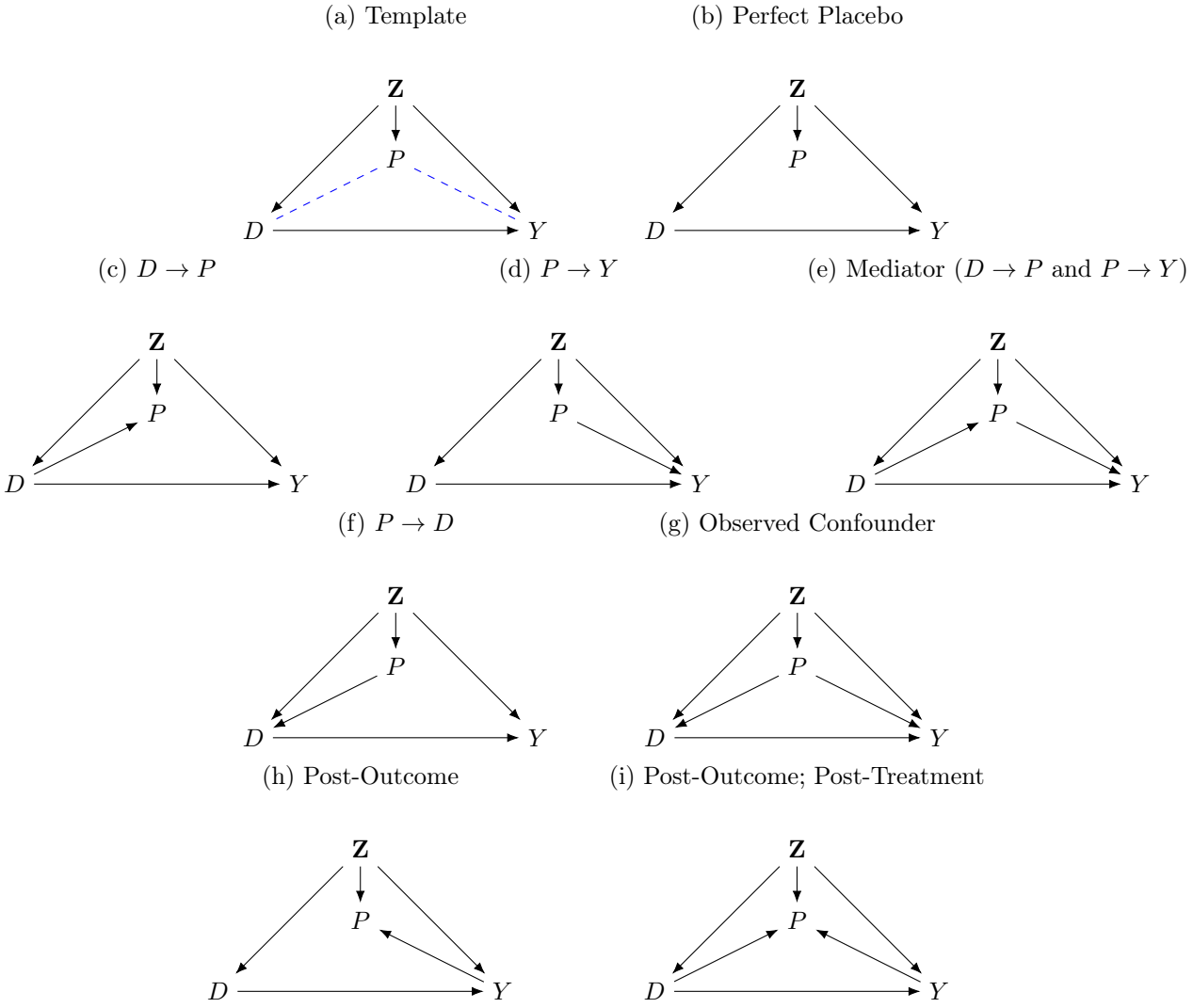
[15] There are $3^2 - 1 = 8$ possible graphs, excluding those with cycles.

[16] The approaches presented in Section 2.1 and 2.2 hold for any causal graph. However, the interpretation of the partial identification parameters depend on the causal graph. We, therefore, provide guidance on which approaches we suggest for each causal graph. Assuming a causal graph (or set of causal graphs) that are plausible in any application is a pre-requisite to making any credible causal claims. This section includes some additional approaches not yet discussed. Details of these can be found in the Appendix.

[17] In Figure 3(c), the approach from Section 2.2 will not work well. The regression $Y \sim D + P$ will run into the problem that $\beta_{Y\sim D|P} = \beta_{Y\sim D|P,Z} + \beta_{Y\sim Z|D,P}\beta_{Z\sim D|P}$ and $\beta_{Z\sim D|P}$ captures both $Z \to D$ and $Z \to P \leftarrow D$. This means that a parameter that we will need to reason about will capture the effect of conditioning on a collider.

[18] Note that Chabé-Ferret (2017), Ham and Miratrix (2022), and others warn that conditioning on pre-treatment outcomes in a difference-in-differences type approach can create amplify bias under certain circumstances. However, since we take a partial identification approach, this should not present any issues for using Equation 26.

Figure 3: Single placebo DAGs.



Mediators Figure 3(e) has $D \to P$ and $P \to Y$, making P a mediator between D and Y. In this case there is more than one option. First, Zhang and Ding (2022) provide guidance on how to take an omitted variables bias approach to partial identification of the direct effect and indirect effect. The direct effect approach deals with the fact that conditioning on P opens the collider path $D \to P \leftarrow Z$ and involves specifying partial identification parameters for each of the $Z \to D$, $Z \to P$, and $Z \to Y$ relationships. Suppose we want to consider the total effect. We could use the approach outlined in Cinelli and Hazlett (2020), but this approach for the total effect applied to a setting with a mediator requires that we reason about $R^2_{Y \sim Z|D}$ which captures the $Z \to Y$ and $Z \to P \to Y$ relationships. A wrinkle is that $P \to Y$ is a component of the total effect we want to identify. An alternative would be to follow the approach laid out in Section 2.1, where the mediator is treated as a placebo outcome. Here there is again the wrinkle that we will be required to reason about $D \to P$ (above we refer to the placebo outcomes as N), which is a component of the total effect we want to identify. We could also follow the approach in Equation 27. This does something similar but where we treat the mediator as a placebo treatment. The wrinkle remains that we need to reason about $P \to Y$, a component of the total effect we want to identify.

$$
\begin{aligned}
&\text{bias}_{(\text{YD.X})} = \beta_{Y\sim D|X} - \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}(\text{SD}(Y^{\perp D,X})/\text{SD}(D^{\perp X}))\\
&\text{bias}_{(\text{YP.DX})} = \beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X} = R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{P\sim Z_{(\text{Y.DPX})}|D,X}(\text{SD}(Y^{\perp D,P,X})/\text{SD}(P^{\perp D,X}))\\
&\text{bias}_{(\text{YD.X})} = m \times \text{bias}_{(\text{YP.DX})} \implies \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = \beta_{Y\sim D|X} - m \times (\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X})\\
&m = \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{YP.DX})}} = \frac{R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{P\sim Z_{(\text{Y.DPX})}|D,X}} \frac{\text{SD}(P^{\perp D,X})}{\text{SD}(D^{\perp X})} \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(Y^{\perp D,P,X})} = k \times \frac{\text{SD}(P^{\perp D,X})}{\text{SD}(D^{\perp X})} \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(Y^{\perp D,P,X})}\\
&\therefore \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = \beta_{Y\sim D|X} - k \times (\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X}) \times \frac{\text{SD}(P^{\perp D,X})}{\text{SD}(D^{\perp X})} \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(Y^{\perp D,P,X})}
\end{aligned}
\tag{27}
$$

In Equation 27, $k \triangleq \frac{R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{P\sim Z_{(\text{Y.DPX})}|D,X}}$ is the ratio of the level of confounding D and Y to the level of confounding P and Y, conditional on D, after re-scaling one of these biases to account for scale differences. $\beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim P|D,Z,X}$ measures the causal effect of P on Y.

Observed Confounders Figure 3(f) has $P \to D$. This is similar to an observed confounder. Figure 3(g) has $P \to D$ and $P \to Y$, making P an observed confounder. In both of these cases, there are two options. First, we could follow the approach laid out in Cinelli and Hazlett (2020), where P is an observed covariate. Second, we could follow the approach in Equation 28. The key here is to include P in the regression of Y on D and to also consider the regression of D on P.

$$
\begin{aligned}
&\text{bias}_{(\text{YD.PX})} = \beta_{Y\sim D|P,X} - \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}(\text{SD}(Y^{\perp D,P,X})/\text{SD}(D^{\perp P,X}))\\
&\text{bias}_{(\text{DP.X})} = \beta_{D\sim P|X} - \beta_{D\sim P|Z_{(\text{D.PX})},X} = R_{D\sim Z_{(\text{D.PX})}|P,X} f_{P\sim Z_{(\text{D.PX})}|X}(\text{SD}(D^{\perp P,X})/\text{SD}(P^{\perp X}))\\
&\text{bias}_{(\text{YD.PX})} = m \times \text{bias}_{(\text{DP.X})} \implies \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim D|P,X} - m \times (\beta_{D\sim P|X} - \beta_{D\sim P|Z_{(\text{D.PX})},X})\\
&m = \frac{\text{bias}_{(\text{YD.PX})}}{\text{bias}_{(\text{DP.X})}} = \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}}{R_{D\sim Z_{(\text{D.PX})}|P,X} f_{P\sim Z_{(\text{D.PX})}|X}} \times \frac{\text{SD}(Y^{\perp D,P,X})}{\text{SD}(D^{\perp P,X})} \times \frac{\text{SD}(P^{\perp X})}{\text{SD}(D^{\perp P,X})} = k \times \frac{\text{SD}(Y^{\perp D,P,X})}{\text{SD}(D^{\perp P,X})} \times \frac{\text{SD}(P^{\perp X})}{\text{SD}(D^{\perp P,X})}\\
&\therefore \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim D|P,X} - k \times (\beta_{D\sim P|X} - \beta_{D\sim P|Z_{(\text{D.PX})},X}) \times \frac{\text{SD}(Y^{\perp D,P,X})}{\text{SD}(D^{\perp P,X})} \times \frac{\text{SD}(P^{\perp X})}{\text{SD}(D^{\perp P,X})}
\end{aligned}
\tag{28}
$$

In Equation 28, $k \triangleq \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{D\sim Z_{(\text{Y.DPX})}|P,X}}{R_{D\sim Z_{(\text{D.PX})}|P,X} f_{P\sim Z_{(\text{D.PX})}|X}}$ is the ratio of the level of confounding D and Y, conditional on P, to the level of confounding D and P, after re-scaling one of these biases to account for scale differences. Since we included P in the regression of Y on D, we see that the confounding of D and Y does not include the paths that run through P, making comparison with the confounding of D and P easier. $\beta_{D\sim P|Z_{(\text{D.PX})},X} = \beta_{D\sim P|Z,X}$ measures the causal effect of P on D.

Post-outcome Figure 3(h) has $Y \to P$. Figure 3(i) has $Y \to P$ and $D \to P$. In both cases, P is a post-outcome variable. In both of these cases, there are two options. First, we could follow the approach laid out in Cinelli and Hazlett (2020), where we ignore P. Second, we could follow the approach in Equation 29. The key is to consider the regression of P on Y and D.

$$
\begin{aligned}
&\text{bias}_{(\text{YD.X})} = \beta_{Y\sim D|X} - \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}(\text{SD}(Y^{\perp D,X})/\text{SD}(D^{\perp X}))\\
&\text{bias}_{(\text{PY.DX})} = \beta_{P\sim Y|D,X} - \beta_{P\sim Y|D,Z_{(\text{P.YDX})},X} = R_{P\sim Z_{(\text{P.YDX})}|Y,D,X} f_{Y\sim Z_{(\text{P.YDX})}|D,X}(\text{SD}(P^{\perp Y,D,X})/\text{SD}(Y^{\perp D,X}))\\
&\text{bias}_{(\text{YD.X})} = m \times \text{bias}_{(\text{PY.DX})} \implies \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = \beta_{Y\sim D|X} - m \times (\beta_{P\sim Y|D,X} - \beta_{P\sim Y|D,Z_{(\text{P.YDX})},X})\\
&m = \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{PY.DX})}} = \frac{R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}}{R_{P\sim Z_{(\text{P.YDX})}|Y,D,X} f_{Y\sim Z_{(\text{P.YDX})}|D,X}} \times \frac{\text{SD}(Y^{\perp D,P,X})}{\text{SD}(D^{\perp P,X})} \times \frac{\text{SD}(P^{\perp X})}{\text{SD}(D^{\perp P,X})} = k \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(D^{\perp X})} \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(P^{\perp Y,D,X})}\\
&\therefore \beta_{Y\sim D|Z_{(\text{Y.DX})},X} = \beta_{Y\sim D|X} - k \times (\beta_{P\sim Y|D,X} - \beta_{P\sim Y|D,Z_{(\text{P.YDX})},X}) \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(D^{\perp X})} \times \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(P^{\perp Y,D,X})}
\end{aligned}
\tag{29}
$$

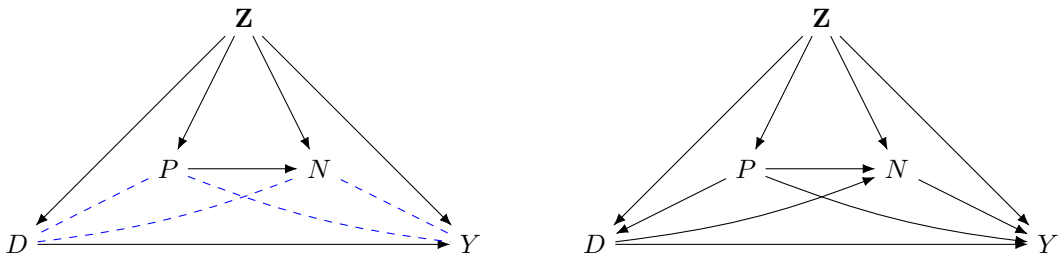
In Equation 29, $k \triangleq \frac{R_{Y\sim Z_{(\text{Y.DX})}|D,X} f_{D\sim Z_{(\text{Y.DX})}|X}}{R_{P\sim Z_{(\text{P.YDX})}|Y,D,X} f_{Y\sim Z_{(\text{P.YDX})}|D,X}}$ is the ratio of the level of confounding of D and Y to the level of confounding of Y and P, after re-scaling one of these biases to account for scale differences. $\beta_{P\sim Y|D,Z_{(\text{P.YDX})},X} = \beta_{P\sim Y|D,Z,X}$ measures the causal effect of Y on P.

3.2 Double Placebos

In the double placebo case, there are many more possible graphs. Again, we assume that all relevant unobserved common causes are bundled into the vector of unobserved variables $\mathbf{Z}$. Observed covariates $\mathbf{X}$ are omitted from the graphs for clarity. All the dashed edges in Figure 4(a) could have either direction or not exist[19] (excluding combinations of edge directions that lead to cycles). However, we limit our analysis to the specific case shown in Figure 4(b) and detailed in Section 2.3. This may cover many practical settings and researchers interested in alternative graphs can reason through the implications of their favored graph and its implications for interpreting the partial identification parameters based on Section 2.3. There may be other partial identification approaches available in the double placebo case but we omit an exploration of these for brevity.

Figure 4: Double placebo causal graph.

(a) Template (b) Observed confounder and mediator

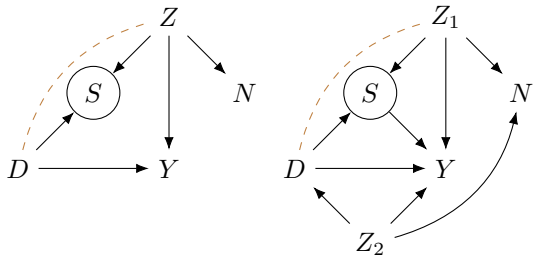


3.3 Sample selection and other applications

Arnold et al. (2016) suggests that placebos can be used to address sample selection and measurement bias in addition to unobserved confounding. This might look like Figure 5. Take Figure 5(a) as an example. Here selection of units into the study sample is indicated by a special node S. Since we are limited to the study sample, we are implicitly conditioning on S, which opens a non-causal path from D to Y. However, if sample selection also opens a similar non-causal path from D to N, then we may be able to use information about the placebo outcome N to inform how sample selection biases the estimate of the effect of D on Y. See Rohde and Hazlett (20XX) for a discussion of causal graphs and sample selection.

Figure 5: Example causal graphs with a placebo outcome, N. D is treatment; Y is outcome; Z contains unobserved confounders; S represents sample selection. Additional observed covariates are omitted for simplicity.

(a) (b)



[19]There are $3^4 = 81$ possible graphs, assuming that users know which of the two placebos is causally prior to the other; but the graphs with cycles should be excluded. Making the conservative assumption that the dashed edges in Figure 4(a) do exist, then there are $2^4 = 16$ possible fully connected graphs; again, the graphs with cycles should be excluded.

4 Difference in differences as special case

In this section, we explore how the approaches put forth in this paper correspond to various relaxations to the standard assumptions of difference in differences identification strategies. Let's start by reviewing the typical difference in differences (DID) setup and identification strategy. Suppose we have panel data for two periods for a cohort of units. We observe the outcome of interest in period 1, which we will label N, as well as the outcome of interest in period 2, which we will call Y. In period 1, no units are treated but in period 2 some units are treated. We indicate the group of units that is treated with $G = 1$. We indicate actual treatment with $D = 1$. We therefore have that in period 1 $D = 0$ for all units and that in period 2 $G = D$. For now, we assume that we do not need to adjust for any covariates. The standard DID identification strategy identifies the ATT for Y (ATT_Y) in the following way. Note that Y_d and N_d are the potential outcomes when D is set, perhaps counterfactually, to $D = d$ with $d \in \{0, 1\}$. We start by revisiting a standard decomposition of the simple difference in means (DIM) between the treated and control groups for Y.

$$\begin{aligned}
\text{DIM}_Y &= \mathbb{E}[Y|G=1] - \mathbb{E}[Y|G=0] \\
&= \mathbb{E}[Y_1|G=1] - \mathbb{E}[Y_0|G=0] \text{ by consistency and since } D = G \text{ in period 2} \\
&= \underbrace{\mathbb{E}[Y_1|G=1] - \mathbb{E}[Y_0|G=1]}_{\text{ATT}_Y} + \underbrace{\mathbb{E}[Y_0|G=1] - \mathbb{E}[Y_0|G=0]}_{\text{Bias}_Y} \\
&= \text{ATT}_Y + \text{Bias}_Y
\end{aligned} \tag{30}$$

We make a "parallel trends" assumption that the trend in the mean control potential outcomes for the treated group is the same as the trend for the control group (i.e., $\mathbb{E}[Y_0|G=1] - \mathbb{E}[N_0|G=1] = \mathbb{E}[Y_0|G=0] - \mathbb{E}[N_0|G=0]$). The parallel trends assumption can be re-ordered into an "equi-confounding" assumption that the mean baseline difference between the treated and control groups in period 2 is the same as the mean baseline difference between the treated and control groups in period 1. The equivalence of parallel trends and equi-confounding is discussed in Sofer et al. (2016).

$$\begin{aligned}
\text{Parallel Trends} &\triangleq \underbrace{\mathbb{E}[Y_0|G=1] - \mathbb{E}[N_0|G=1]}_{\text{Trend}_{G=1}} = \underbrace{\mathbb{E}[Y_0|G=0] - \mathbb{E}[N_0|G=0]}_{\text{Trend}_{G=0}} \\
&\iff \underbrace{\mathbb{E}[Y_0|G=1] - \mathbb{E}[Y_0|G=0]}_{\text{Bias}_Y} = \underbrace{\mathbb{E}[N_0|G=1] - \mathbb{E}[N_0|G=0]}_{\text{Bias}_N} \triangleq \text{Equi-Confounding}
\end{aligned} \tag{31}$$

Now, we recall our constraint that N is a pre-treatment version of the outcome. This means that the treatment cannot have an effect on the pre-treatment outcome (and that no units are treated in period 1). So $N_1 = N_0 = N$ for both the treated and control groups. So we see that $\text{DIM}_N = \text{ATT}_N + \text{Bias}_N = \text{Bias}_N$.

$$\begin{aligned}
\text{DIM}_N &= \mathbb{E}[N|G=1] - \mathbb{E}[N|G=0] \\
&= \mathbb{E}[N_1|G=1] - \mathbb{E}[N_0|G=0] \\
&= \mathbb{E}[N_1|G=1] - \mathbb{E}[N_0|G=1] + \mathbb{E}[N_0|G=1] - \mathbb{E}[N_0|G=0] \\
&= \underbrace{\mathbb{E}[N_1 - N_0|G=1]}_{\text{ATT}_N} + \underbrace{\mathbb{E}[N_0|G=1] - \mathbb{E}[N_0|G=0]}_{\text{Bias}_N} \\
&= \mathbb{E}[N - N|G=1] + \text{Bias}_N = \text{Bias}_N
\end{aligned} \tag{32}$$

Therefore, we have that $\text{Bias}_Y = \text{Bias}_N = \text{DIM}_N$ and that $\text{DIM}_Y = \text{ATT}_Y + \text{Bias}_Y$. Plugging into the later, we have $\text{DIM}_Y = \text{ATT}_Y + \text{DIM}_N$, which can be rearranged to yield the standard difference in differences identification result, $\text{ATT}_Y = \text{DIM}_Y - \text{DIM}_N$, which makes clear where the DID name comes from.

Parallel Trends and Equi-Confounding While it is common to make parallel trends assumptions, equi-confounding is mathematically equivalent and is not any less intuitive to consider. Equi-confounding assumes that the difference in mean control potential outcomes for the treated and control groups in period 2 is the same as the difference for the treated and control groups in period 1. This means that, on average, the way that the treatment and control groups differ for the outcome in period 1 can be used as a stand in for the way that the treatment and control groups differ for the outcome in period 2. Researchers should be able to consider this and possible violations to it just as easily as the parallel trends assumptions and its violations. In what follows, we will stick to equi-confounding type assumptions, rather than parallel trends assumptions since they have a more direct connection to the bias that we are trying to overcome, namely Bias_Y. But every equi-confounding type assumption we might consider has analogous parallel trends assumption, as we will show.

Relaxing Assumptions It is easy to see that we can relax the assumption that there is no effect of D on N by allowing ATT_N to be non-zero. Further, we can relax the assumption that Bias_Y exactly equals Bias_N by recognizing that, for some $m \in \mathbb{R}$, $\text{Bias}_Y = m \times \text{Bias}_N$. Together, these relaxations give us the expression for ATT_Y in Equation 33.

$$\begin{aligned} \text{ATT}_Y &= \text{DIM}_Y - \text{Bias}_Y \\ &= \text{DIM}_Y - m \times \text{Bias}_N \\ &= \text{DIM}_Y - m \times (\text{DIM}_N - \text{ATT}_N) \end{aligned} \tag{33}$$

That is, we can express ATT_Y as a difference between DIM_Y and an "adjusted" version of DIM_N, where the adjustment is parameterized by $m = \frac{\text{Bias}_Y}{\text{Bias}_N}$ and ATT_N. The statement $\text{Bias}_Y = m \times \text{Bias}_N$ is similar to the equi-confounding assumption from the standard DID approach. And the expression $\text{ATT}_Y = \text{DIM}_Y - m \times (\text{DIM}_N - \text{ATT}_N)$ is similar to a difference in differences type identification result. So we are not too far from familiar results, but this approach would apply to more general placebo outcomes and settings outside the standard DID setting. Equation 33 could be used for partial identification in cases with binary treatments and a placebo outcome. Researchers would estimate DIM_Y and DIM_N and then reason about plausible values for m and ATT_N to partially identify ATT_Y.[20] But how does this all relate to the regression-based results we discuss in Section 2?

Connecting to Regression Since G is binary, a regression of Y or N on G will give the difference in means DIM_Y or DIM_N, respectively. Additionally, we can use traditional omitted variable bias analysis to show that,

$$\begin{aligned} \text{DIM}_Y &= \beta_{Y\sim G} = \beta_{Y\sim G|Z_{(\text{Y.G})}} + \beta_{Y\sim Z_{(\text{Y.G})}|G}\beta_{Z_{(\text{Y.G})}\sim G} \\ \text{DIM}_N &= \beta_{N\sim G} = \beta_{N\sim G|Z_{(\text{N.G})}} + \beta_{N\sim Z_{(\text{N.G})}|G}\beta_{Z_{(\text{N.G})}\sim G} \end{aligned} \tag{34}$$

where $Z_{(\text{Y.D})}$ and $Z_{(\text{N.D})}$ are linear combinations of a vector $\mathbf{Z}$ of all the time-varying and non-time-varying confounders of the G, Y and G, N relationships, like in previous sections. $\beta_{Y\sim G}$ is the regression coefficient on G from the regression $Y = \beta_{Y\sim G}G + \epsilon_1$. $\beta_{Y\sim G|Z} = \beta_{Y\sim G|Z_{(\text{Y.G})}}$ is the regression coefficient on G from the regression $Y = \beta_{Y\sim G|Z}G + \mathbf{Z}\beta_{Y\sim Z|G} + \epsilon_2$. All the other regression coefficients are defined similarly. We can also see that Equation 35 holds.

$$\text{ATT}_Y + \text{Bias}_Y = \text{DIM}_Y = \beta_{Y\sim G} = \beta_{Y\sim G|Z} + \beta_{Y\sim Z_{(\text{Y.G})}|G}\beta_{Z_{(\text{Y.G})}\sim G} \tag{35}$$

Under the standard DID assumptions, we could write $\text{ATT}_Y = \beta_{Y\sim G} - \beta_{N\sim G}$, but we want to relax these assumptions. Under an assumption of effect homogeneity, or using OLS regression as an approximation, we have that $\text{ATT}_Y = \beta_{Y\sim G|Z}$, which means that $\text{Bias}_Y = \beta_{Y\sim Z_{(\text{Y.G})}|G}\beta_{Z_{(\text{Y.G})}\sim G}$. We can also see that, $\text{DIM}_N = \beta_{N\sim G}$ and so $\text{DIM}_N - \beta_{N\sim G|Z} = \beta_{N\sim G} - \beta_{N\sim G|Z} = \beta_{N\sim Z_{(\text{N.D})}|G}\beta_{Z_{(\text{N.D})}\sim G}$. Again, there is some $m \in \mathbb{R}$ such that $\beta_{Y\sim G} - \beta_{Y\sim G|Z} = m \times (\beta_{N\sim G} - \beta_{N\sim G|Z})$. Plugging this into the expression for ATT_Y, we get Equation 36.

$$\text{ATT}_Y = \beta_{Y\sim G|Z} = \beta_{Y\sim G} - \beta_{Y\sim Z|G}\beta_{Z\sim G} = \beta_{Y\sim G} - m \times (\beta_{N\sim G} - \beta_{N\sim G|Z}) \tag{36}$$

$m = \frac{\beta_{Y\sim Z_{(\text{Y.D})}|G}\beta_{Z_{(\text{Y.D})}\sim G}}{\beta_{N\sim Z_{(\text{N.D})}|G}\beta_{Z_{(\text{N.D})}\sim G}}$ is the ratio of biases for the regression coefficients. Under the assumption of effect homogeneity for the effect of D on N or using OLS as an approximation, we see that $\beta_{N\sim G|Z} = \text{ATT}_N$, $m = \frac{\text{Bias}_Y}{\text{Bias}_N}$, and $\beta_{N\sim G} = \text{DIM}_N$. Thus, Equations 33 and 36 are equivalent. Note that we don't need to make the effect homogeneity for the effect of D on N for Equation 36 to be useful, however. We can simply parameterize the expression for ATT_Y with m and $\beta_{N\sim G|Z}$. Additionally, if we do not want to make such an assumption for the effect of D on Y, we can make $\beta_{Y\sim G|Z}$ our inferential target and still make use of Equation 36. This framing connects us back to the framework laid out in Section 2.

We can also define k as in Section 2. If the scale of N and Y differ, it will be easier to reason about the value of k than to reason about the value of m. So we now can see how we can relax both the assumption that N is a pre-treatment outcome with no treatment effect from G and the assumption that we have perfect equi-confounding. We also have seen how the latter

[20] We suspect there are more settings in which multiple applications of the decomposition DIM = ATT + Bias could yield useful expressions for causal effects with binary treatments, similar to how Section 2 involves multiple applications of the omitted variables bias formula. For example, another possibility is for a binary placebo treatment P with actual binary treatment D and outcome Y. We may have that

$$\underbrace{\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0]}_{\text{DIM}_{Y\sim D}} = \underbrace{\mathbb{E}[Y_{D=1}|D=1] - \mathbb{E}[Y_{D=0}|D=1]}_{\text{ATT}_{Y\sim D}} + \underbrace{\mathbb{E}[Y_{D=0}|D=1] - \mathbb{E}[Y_{D=0}|D=0]}_{\text{Bias}_{Y\sim D}}$$

$$\underbrace{\mathbb{E}[Y|P=1] - \mathbb{E}[Y|P=0]}_{\text{DIM}_{Y\sim P}} = \underbrace{\mathbb{E}[Y_{P=1}|P=1] - \mathbb{E}[Y_{P=0}|P=1]}_{\text{ATT}_{Y\sim P}} + \underbrace{\mathbb{E}[Y_{P=0}|P=1] - \mathbb{E}[Y_{P=0}|P=0]}_{\text{Bias}_{Y\sim P}}$$

There is some m such that $\text{Bias}_{Y\sim D} = m \times \text{Bias}_{Y\sim P}$. So we can write $\text{ATT}_{Y\sim D} = \text{DIM}_{Y\sim D} - m \times (\text{DIM}_{Y\sim P} - \text{ATT}_{Y\sim P})$. We could then partially identify $\text{ATT}_{Y\sim D}$ by estimating $\text{DIM}_{Y\sim D}$ and $\text{DIM}_{Y\sim P}$ and reasoning about plausible values for $m \triangleq \frac{\text{Bias}_{Y\sim D}}{\text{Bias}_{Y\sim P}}$ and $\text{ATT}_{Y\sim P}$.

relaxation could be parameterized in terms of $\beta_{N\sim G|Z}$ and m or k. This all can be done while maintaining an identification strategy with roots that connect easily to the standard difference in differences assumptions and identification strategy.

Moreover, we can see standard difference in differences as the special case where we assume that $\beta_{Y\sim G|Z} = 0$ and $m = 1$ or equivalently $k = \frac{\text{SD}(N^{\perp G})}{\text{SD}(Y^{\perp G})}$. This value of k arising could be the result of simply a fortunate balancing coincidence for scale differences and levels of confounding. But it is more likely to be interpreted as the scale of N and Y being the same (i.e., $\text{SD}(N^{\perp G}) = \text{SD}(Y^{\perp G})$) in addition to equi-confounding. In this way, the assumption that $k = 1$ can be thought of as less demanding than the assumption that $m = 1$. So we could consider a "new" form of difference in difference where we assume $k = 1$ but do not assume that $m = 1$. However, the above framework need not be limited to any specific choice for k. Rather, we can consider a range of plausible values for k, perhaps including the value that would bring the true effect to zero under the assumption that $\beta_{N\sim G|Z} = 0$, $k = \frac{\text{DIM}_Y}{\text{DIM}_N} \times \frac{\text{SD}(N^{\perp G})}{\text{SD}(Y^{\perp G})}$, as well as $k = 1$ and $k = \frac{\text{SD}(N^{\perp G})}{\text{SD}(Y^{\perp G})}$.

Transforming unequal confounding assumptions into unequal trend assumptions Finally, we show that any assumption for m corresponds to an assumption on the trend in control potential outcomes. We will show that any assumption for m can be turned into an assumption about how the trends in control potential outcomes compare. An assumption about a difference in confounding would be something like Equation 37. An assumption about a difference in trend would be something like Equation 38.

$$\underbrace{\mathbb{E}[Y_0|G=1] - \mathbb{E}[Y_0|G=0]}_{\text{Bias}_Y} = m \times \underbrace{(\underbrace{\mathbb{E}[N|G=1] - \mathbb{E}[N|G=0]}_{\text{DIM}_N} - \beta_{N\sim G|Z})}_{\text{Bias}_N} \tag{37}$$

$$\iff \mathbb{E}[Y_0|G=1] = \mathbb{E}[Y_0|G=0] + m \times (\mathbb{E}[N|G=1] - \mathbb{E}[N|G=0] - \beta_{N\sim G|Z})$$

$$\underbrace{\mathbb{E}[Y_0|G=1] - \mathbb{E}[N|G=1]}_{\text{Trend}_{G=1}} = w \times (\underbrace{\mathbb{E}[Y_0|G=0] - \mathbb{E}[N|G=0]}_{\text{Trend}_{G=0}}) \tag{38}$$

$$\iff \mathbb{E}[Y_0|G=1] = \mathbb{E}[N|G=1] + w \times (\mathbb{E}[Y_0|G=0] - \mathbb{E}[N|G=0])$$

For this exposition, we assume $\beta_{N\sim G|Z} = 0$ for simplicity. Both Equations 37 and 38 provide an expression for $\mathbb{E}[Y_0|G=1]$, which, in the DID setting, is unobserved while all the other components of the expressions are estimable from the data. We can set these two expressions for $\mathbb{E}[Y_0|G=1]$ equal to each other and then solve for w to get the trend assumption that corresponds to our assumption on m.

$$\mathbb{E}[Y_0|G=0] + m \times (\mathbb{E}[N|G=1] - \mathbb{E}[N|G=0]) = \mathbb{E}[N|G=1] + w \times (\mathbb{E}[Y_0|G=0] - \mathbb{E}[N|G=0]) \tag{39}$$

$$\begin{aligned} \iff w &= \frac{\mathbb{E}[Y_0|G=0] + m \times (\mathbb{E}[N|G=1] - \mathbb{E}[N|G=0]) - \mathbb{E}[N|G=1]}{\mathbb{E}[Y_0|G=0] - \mathbb{E}[N|G=0]} \\ \iff m &= \frac{\mathbb{E}[N|G=1] + w \times (\mathbb{E}[Y_0|G=0] - \mathbb{E}[N|G=0]) - \mathbb{E}[Y_0|G=0]}{\mathbb{E}[N|G=1] - \mathbb{E}[N|G=0]} \end{aligned} \tag{40}$$

We could similarly write m in terms of w, if we preferred to make an assumption on w, the difference in the trend, and wanted to know what this implies about the unequal confounding. Additionally, we know that $m = k \times \frac{\text{SD}(Y^{\perp G})}{\text{SD}(N^{\perp G})}$ so that we can make an assumption on k and see what this implies about w, the trend, or make a trend assumption on w and see what this implies about k.

5 Application: National Supported Work Demonstration

Let us consider a substantive example. The National Supported Work Demonstration (NSW) was a job training program aimed at helping disadvantaged worker build basic skills in the 1970s. The program randomly assigned participants to training positions. These consisted of a group that received the benefits of the program (the treatment group) and a group that did not receive any benefits (the control group). (LaLonde, 1986) The data also include a non-experimental dataset in which individuals not from the NSW program but from the Panel Study of Income Dynamics (PSID) are added as additional control units. The data that we work with is "a subset that consists of males with three (2 pre-training and 1 post-training) years of earnings data (the outcome of interest). This subset has been considered in (Dehejia and Wahba 1999; Smith and Todd 2005; Firpo 2007)." (Callaway, 2019) The data can be found in the "`qte`" `R` package by running the command `data(lalonde,package = "qte")`. We use the non-experimental dataset (called `lalonde.psid`) to illustrate the methods

discussed in this paper and then compare these results with the estimates from the experimental data (called `lalonde.exp`).[21] The difference in means estimate of the treatment effect from the experimental data is 1,795.55 higher 1978 earnings in dollars for the treated group relative to the control group. The estimate is 1,693.12 when we control for observed covariates.

We use 1974 earnings (`re74`) as a placebo outcome for 1978 earnings (`re78`). The treatment (`treat`) is experiencing the job training program. We also use the set of covariates (X) in the data including age, education (years and degree or not), marital status, and demographic indicators. We recognize that there are unobserved covariates (Z) that we wish we could condition on but cannot. Ideally, we would like to estimate the regression `lm(re78 ~ treat + X + Z)`. Due to the unobserved confounders, we will not be able to point identify the coefficient on `treat` from this "long" regression. Instead, we partially identify $\beta_{Y\sim D|X,Z}$ with the approach proposed by the present paper and leverage information on 1974 earnings as a placebo outcome. We note that it is likely that 1974 earnings influence 1978 earnings. However, Chabé-Ferret (2017); Ham and Miratrix (2022) warn against conditioning on pre-treatment outcomes in a difference-in-differences type approach. We follow that guidance here for simplicity. Thus we use the approach from Section 2.1. Since the placebo outcome is pre-treatment earnings, it is reasonable to assume that the treatment has no causal effect on the the placebo outcome. Therefore, we primarily consider that $\beta_{P\sim D|Z,X} = 0$. We carryout the partial identification analysis using the "`placeboLM`" `R` package the authors built for this purpose. `placeboLM` is currently under development; a software paper will accompany its release.

```
# install.packages("devtools")
devtools::install_github("Adam-Rohde/placeboLM")
library(placeboLM)
```

This package can be used in a manner reminiscent of the `lm()` function in `R`. Arguments to the `placeboLM` function include the dataset, the outcome variable, the treatment variable, the placebo variable (and whether we want to think of it as a placebo outcome or treatment), the direction or presence of the edge from the treatment (D) to the placebo (P), the direction or presence of the edge from the placebo (P) to the outcome (Y), the observed covariates, and finally the ranges of the partial identification parameters that we want to consider. These ranges can be set to be the range that is plausible or something more conservative than that. The first partial identification parameter is k, which captures the ratio of the level of confounding of D and Y, conditional on P, to the level of confounding of D and P, after accounting for scale differences. We can think of this as the D, Y confounding as a percentage of the D, P confounding, where scale differences between P and Y have been accounted for. The second partial identification parameter is $\beta_{P\sim D|Z,X}$, which captures the direct causal relationship between D and P and we already noted should be zero. This parameter is referred to as `coef_P_D_given_XZ` in the software. For this example, we set the partial identification parameter ranges to be conservative and assume that the placebo (1974 earnings) directly causes the outcome (1978 earnings).

```
plm = placeboLM(
  data = "lalonde.psid",
  outcome = "re78",
  treatment = "treat",
  placebo_outcome = "re74",
  placebo_treatment = "",
  DP = "",
  PY = "",
  observed_covariates = c("age", "education", "black", "hispanic", "married", "nodegree"),
  partialIDparam_minmax = list(k = c(-2,2), coef_P_D_given_XZ = c(-15000,15000)) )
```

Running the `placeboLM` function provides us with a quick summary of the setting that we have chosen and the regressions that we will estimate as part of the partial identification framework. We can then use the three main analysis functions of the `placeboLM` package to do partial identification for $\beta_{Y\sim D|X,Z}$, which is the regression coefficient that we would get if we could include the unobserved confounders in our regression. The first is `placeboLM_table()`. We can specify which percentiles of the range of plausible values for the partial identification parameters we would like to report estimates for. The rows with these estimates are labelled "`Grid`". We get bootstrapped standard errors and confidence intervals in addition to the estimates. Finally, we also get the estimates under common point identification assumptions: selection on observables (SOO), standard difference in differences (DID), and DID assuming $k = 1$.

```
set.seed(0)
placeboLM_table(plm, n_boot = 1000, ptiles = c(0.25,0.5,0.75), alpha = 0.05)
```

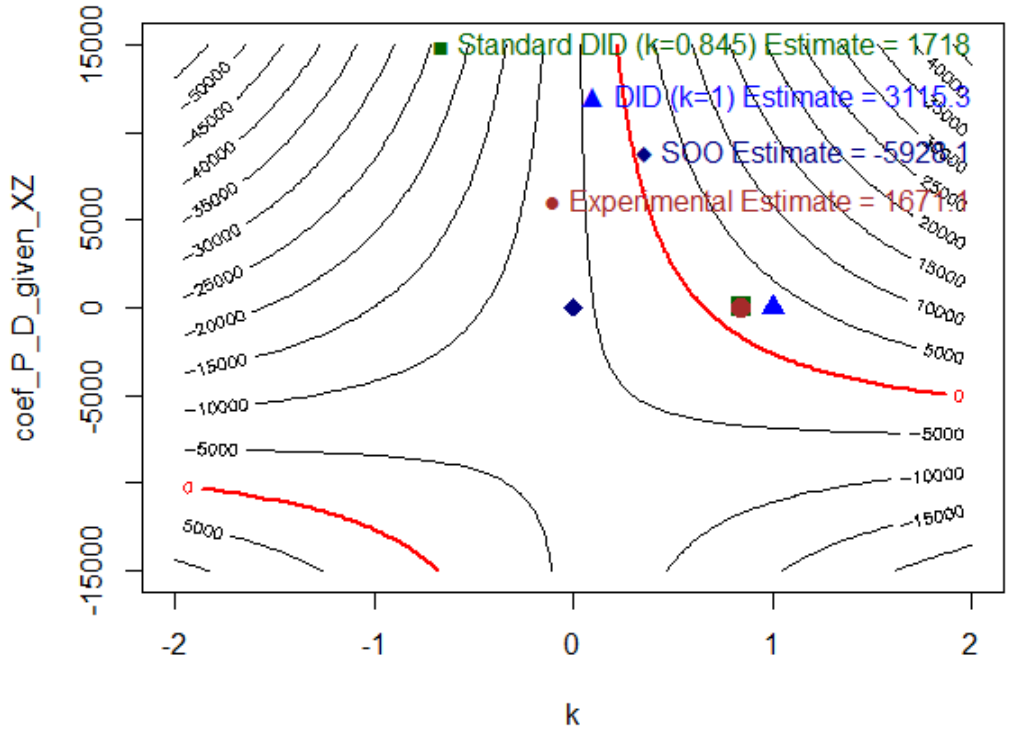
[21] Since this data includes a randomized experiment, we can also estimate the effect using the experimental data to get an unbiased estimate of the true treatment effect. This experimental estimate can be used to back out the true value for k, assuming that $\beta_{N\sim D|Z,X} = 0$.

	k	coef_P_D_given_XZ	Estimate	Std. Error	CI Low	CI High
SOO	0	0	-5928.11	822.57	-7563.28	-4337.09
Standard DID	0.84549	0	1718.01	852.98	27.35	3393.96
k=1 DID	1	0	3115.31	882.42	1337.4	4783.3
Grid	-1	-7500	-6100.92	1375.32	-8843.01	-3539.27
Grid	-1	0	-14971.53	1379.99	-17615.87	-12160.46
Grid	-1	7500	-23842.13	1484.07	-26566.69	-20663
Grid	0	-7500	-5928.11	809.4	-7473.88	-4304.65
Grid	0	0	-5928.11	842.29	-7549.07	-4215.56
Grid	0	7500	-5928.11	832.39	-7523.88	-4262.12
Grid	1	-7500	-5755.3	871.06	-7441.71	-4051.86
Grid	1	0	3115.31	926.64	1389.42	4907.9
Grid	1	7500	11985.91	984.87	10176.39	13963.14

The second key partial identification function is `placeboLM_contour_plot()` which provides a contour plot of the estimates as we vary the two partial identification parameters within the ranges we originally provided. The plot also includes the SOO and DID estimates for reference. We highlight the contour on which estimates are zero. See Figure 8.

```
placeboLM_contour_plot(plm, gran = 100)
```

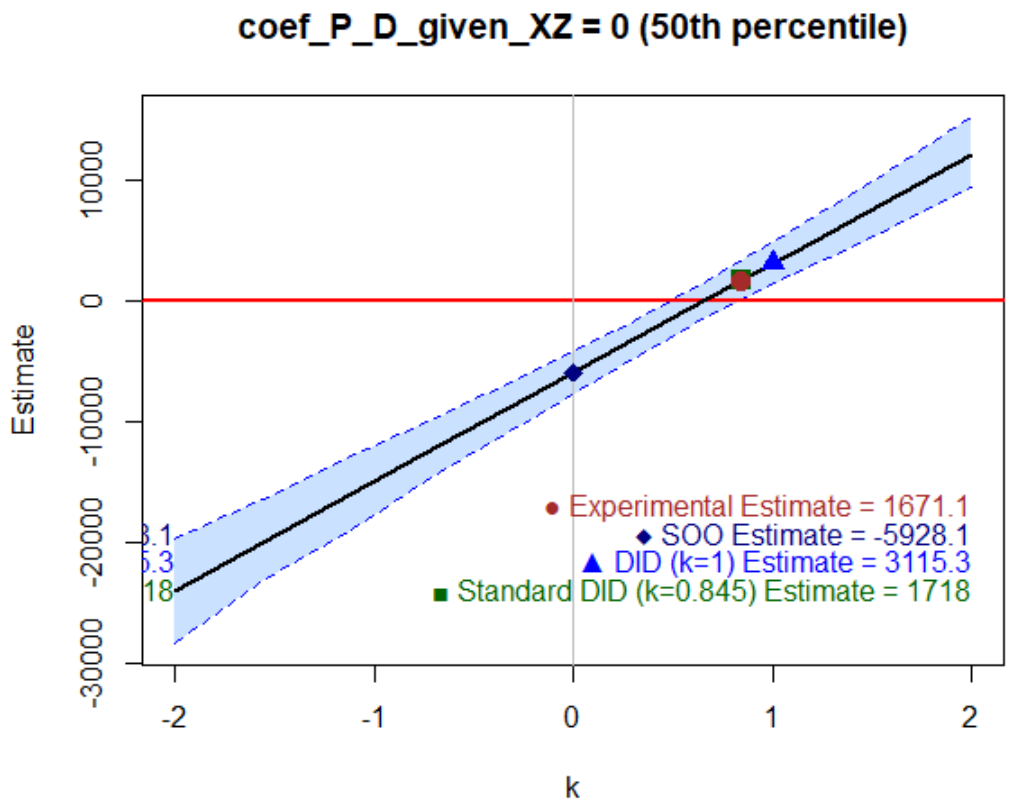
Figure 6: Contour plot for National Supported Work example with 1974 earnings as placebo outcome.



The final key partial identification function is `placeboLM_contour_plot()` which provides a line plot of the estimates as we vary one of the two partial identification parameters. We can specify percentiles of the other partial identification parameter and have a line plot generated for each. Since we believe `coef_P_D_given_XZ` is zero in this example, we focus only on this case. The line plot also includes 95% confidence intervals. See Figure 7.

```
placeboLM_line_plot(plm, bootstrap=TRUE, n_boot=1000, ptiles = c(0.5),
                    focus_param = "k", ptile_param = "coef_P_D_given_XZ",
                    gran= 10, alpha = 0.05)
```

Figure 7: Line plot for National Supported Work example with 1974 earnings as placebo outcome.



The contour plot shows how the effect estimate for $\beta_{Y\sim D|X,Z}$ changes as k and $\beta_{P\sim D|Z,X}$ change. We've noted that it is likely that $\beta_{P\sim D|Z,X} = 0$. The line plot shows us how the effect estimate for $\beta_{Y\sim D|X,Z}$ changes as k changes, under the assumption that $\beta_{P\sim D|Z,X} = 0$. The SOO estimate is negative. This is not the sign we would intuitively expect, since training should probably boost earnings. The standard DID estimate, however, is positive and implicitly makes an assumption that the bias for the treatment and 1978 earnings is 85% of the bias for the treatment and 1974 earnings, after accounting for scale differences. The estimate under this assumption is an increase of 1,718.01 dollars as a result of experiencing the job training. The $k = 1$ DID estimate is also positive and provides an estimate of an increase of 3,115.31 dollars as a result of experiencing the job training. $k = 1$ corresponds to an assumption that the bias for the treatment and 1978 earnings is 100% of the bias for the treatment and 1974 earnings. Under the assumption that $\beta_{P\sim D|Z,X} = 0$, the experimental estimate (1,693.12) implies that the "true" value for k is about 0.84 or that the bias for the treatment and 1978 earnings is 84% of the bias for the treatment and 1974 earnings, after accounting for scale differences.

Without knowledge of the experimental estimate, however, we might consider the line plot and make an argument that the bias for the treatment and 1978 earnings and the bias for the treatment and 1974 earnings have the same sign. This would put us on the right half of the line plot. We might further guess that the levels of confounding or bias are somewhat similar though perhaps not identical, since the two earnings measures are separated in time by a few years. So we might consider values for k that are relatively close to 1. This analysis would lead us to consider effect estimates similar to the two DID estimates, which it turns out would get us close to the experimental estimate. Most importantly, we would likely hone in on a range of plausible effect estimates in which most or all effect estimates are positive. Moreover, we are not tied to a strict parallel trends or equi-confounding assumption and can consider a range of plausible levels of relative confounding.

6 Application: Zika Virus

In our second application, we investigate the effect Zika virus had on birth rates in Brazil. Zika virus can cause birth complications. Brazil was heavily impacted by the outbreak of Zika virus in 2015. We use data that has also been analyzed by Taddeo et al. (2022); Tchetgen et al. (2023a). The virus had a dramatic impact on the state of Pernambuco but zero cases were reported in the state of Rio Grande do Sul. As in previous works, we treat the former state as the treated group and the later as the control group. We use municipality level overall birth rates per 1000 people in 2016 as our primary outcome of interest. We use municipality level overall birth rates per 1000 people in 2014 as our placebo outcome. We refer curious readers to the previously cited paper for additional discussion of the virus and its spread in Brazil. We make the reasonable assumption that the treatment has no direct causal relationship with the pre-treatment outcome. We use the data from Amorim (2022) and run the following application of `placeboLM`. We can get a full picture of partial identification just using `placeboLM_line_plot()`.

```
plm = placeboLM(
  data = "data_reshape",
  placebo_data = NULL,
  outcome = "Rate_2016",
  treatment = "trt",
  placebo_outcome = "Rate_2014",
  placebo_treatment = "",
  DP = "",
  PY = "",
  observed_covariates = c("1"),
  partialIDparam_minmax = list(k = c(0,2), coef_P_D_given_XZ = c(-10,10))
)
```

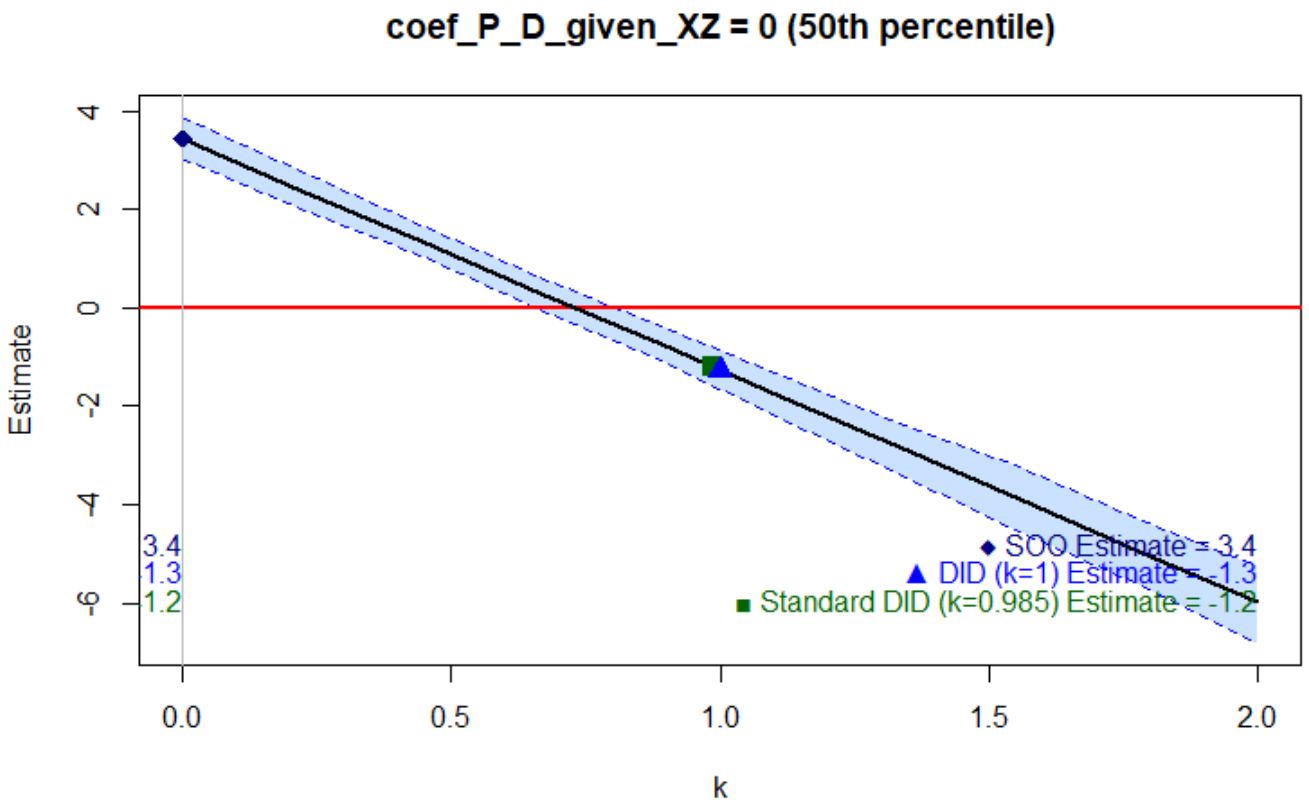
```
set.seed(0)
placeboLM_line_plot(plm,
                    bootstrap=TRUE,
                    n_boot=1000,
                    ptiles = c(0.5),
                    focus_param = "k",
                    ptile_param = "coef_P_D_given_XZ",
                    gran= 10,
                    alpha = 0.05)
```

We find that the naive difference in means estimate is that Zika is associated with an rise in the birth rate by 3.4 births per 1000 people. However, the two difference in difference estimates are negative and indicate that there were about 1.2 to 1.3 fewer births per 1000 people in Pernambuco as a result of the Zika virus. We can also see that if the level of confounding of 2016 birth rate and treatment is above approximately 70% of the level of confounding of 2014 birth rate and treatment, then the effect estimates are always negative. If the relative level of confounding is something like 120%, the effect estimate may be around 2. It is reasonable to believe that the relative level of confounding is not too far from 100%. So we might conclude that the Zika virus reduces birth rate or at least does not increase it. These findings are consistent with the results in Taddeo et al. (2022); Tchetgen et al. (2023a); but using `placeboLM` and the partial identification framework outlined in the present paper builds in a skepticism about whether untestable assumptions like those made in Tchetgen et al. (2023a) actually hold. Rather than hold to a single assumption, we can explore the range of plausible possibilities. Additionally, we could conceive of the state indicator, which is used as a proxy for whether or not the population was exposed to Zika virus, as a direct cause of 2014 birth rates, in that the two states likely differ in important ways that lead to divergent birth rates in 2014. Such a view could easily be accommodated in our framework by considering values for `coef_P_D_given_XZ` other than zero. This would, on the other hand, be a violation of key assumptions in Tchetgen et al. (2023a).

7 Discussion

We have explored how the analysis of "omitted variable bias" in regression provides a simple yet powerful way to leverage information from placebo treatments and outcomes for partial identification of causal effects or approximations of them. We have seen that the framework we developed allows for violations of assumptions made in traditional difference in differences,

Figure 8: Line plot for Zika Virus example with 2014 birth rate as placebo outcome.



as well as in standard negative outcome (and treatment) control settings. These can be viewed as special cases of our approach. Additionally, our framework requires only a single placebo to be observed and does not require assumptions of null effects, certain conditional independencies, or unique solutions to integral equations as in Tchetgen et al. (2020, 2023a). Moreover, our approach is applicable to various causal graphs. We obtain these gains by relying on linear regression machinery but do not require any distributional assumptions. Relaxations of this are explored in Appendix B. We also saw that the forthcoming `placeboLM` `R` package facilitates easy application of the proposed framework. Finally, we emphasized the partial identification approach as a way to emphasize credible findings over those that rely on restrictive assumptions. Researchers should start from only those assumptions that are credibly defensible based on scientific knowledge. Iterative strengthening of causal assumptions can be explored but should not be emphasized over results that researchers feel are most credible. This credible-first approach is the inverse of how typical applied causal studies proceed, but gives us the best chance of advancing science and steering away from biases toward "positive" findings by emphasizing the ranges of effect estimates that are well supported by our scientific knowledge.

References

Amorim, L. (2022). Replication Data for: Zika Epidemic and Birth Rates in Brazil.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arnold, B. F. and Ercumen, A. (2016). Negative Control Outcomes: A Tool to Detect Bias in Randomized Trials. *JAMA*, 316(24):2597–2598.

Arnold, B. F., Ercumen, A., Benjamin-Chung, J., and John M. Colford, J. (2016). Negative Controls to Detect Selection Bias and Measurement Bias in Epidemiologic Studies. *Epidemiology*, 27(5):637–641.

Berk, R., Buja, A., Brown, L., George, E., Kuchibhotla, A. K., Su, W., and Zhao, L. (2021). Assumption lean regression. *The American Statistician*, 75(1):76–84.

Bilinski, A. and Hatfield, L. A. (2018). Nothing to see here? non-inferiority approaches to parallel trends and other model assumptions.

Callaway, B. (2019). *qte: Quantile Treatment Effects*. R package version 1.3.0.

Chabé-Ferret, S. (2017). Should We Combine Difference In Differences with Conditioning on Pre-Treatment Outcomes? TSE Working Papers 17-824, Toulouse School of Economics (TSE).

Chernozhukov, V., Cinelli, C., Newey, W., Sharma, A., and Syrgkanis, V. (2022). Long story short: Omitted variable bias in causal machine learning. Working Paper 30302, National Bureau of Economic Research.

Cinelli, C., Forney, A., and Pearl, J. (2022). A crash course in good and bad controls. *Sociological Methods & Research*, 0(0):00491241221099552.

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.

Freidling, T. and Zhao, Q. (2023). Sensitivity analysis with the r^2-calculus.

Freyaldenhoven, S., Hansen, C., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–38.

Gibson, L. and Zimmerman, F. (2021). Measuring the sensitivity of difference-in-difference estimates to the parallel trends assumption. *Research Methods in Medicine & Health Sciences*, 2(4):148–156.

Ham, D. W. and Miratrix, L. (2022). Benefits and costs of matching prior to a difference in differences analysis when parallel trends does not hold.

Hansen, B. (2022). *Econometrics*. Princeton University Press.

Keele, L. J., Small, D. S., Hsu, J. Y., and Fogarty, C. B. (2019). Patterns of effects and sensitivity analysis for differences-in-differences.

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review*, 76(4):604–620.

Lipsitch, M., Tchetgen, E. T., and Cohen, T. (2010). Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies. *Epidemiology*, 21(3):383–388.

Liu, R. A., Wei, Y., Qiu, X., Kosheleva, A., and Schwartz, J. D. (2022). Short term exposure to air pollution and mortality in the us: a double negative control analysis. *Environmental Health*, 21.

Manski, C. F. and Pepper, J. V. (2018). How Do Right-to-Carry Laws Affect Crime Rates? Coping with Ambiguity Using Bounded-Variation Assumptions. *The Review of Economics and Statistics*, 100(2):232–244.

Mastouri, A., Zhu, Y., Gultchin, L., Korba, A., Silva, R., Kusner, M., Gretton, A., and Muandet, K. (2021). Proximal causal learning with kernels: Two-stage estimation and moment restriction. In Meila, M. and Zhang, T., editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7512–7523. PMLR.

Matthay, E. C. and Glymour, M. M. (2020). A graphical catalog of threats to validity. *Epidemiology*, 31(3):376–384.

Miao, W., Geng, Z., and Tchetgen Tchetgen, E. J. (2018). Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*, 105(4):987–993.

Miao, W., Shi, X., and Tchetgen, E. T. (2020). A confounding bridge approach for double negative control inference on causal effects.

Necker, L. A. (1832). LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid.

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

Rambachan, A. and Roth, J. (2021). An honest approach to parallel trends.

Richardson, T. and Robins, J. (2013). Single world intervention graphs: a primer. *Working Paper, University of Washington, Seattle*.

Rohde, A. and Hazlett, C. (20XX). Revisiting sample selection as a threat to internal validity: New lessons, examples, and tools. *XXXX*, XX(XX):XXX–XXX.

Ryan, A. M., Kontopantelis, E., Linden, A., and James F Burgess, J. (2019). Now trending: Coping with non-parallel trends in difference-in-differences analysis. *Statistical Methods in Medical Research*, 28(12):3697–3711. PMID: 30474484.

Shi, X., Miao, W., and Tchetgen, E. (2020). A selective review of negative control methods in epidemiology. *Current Epidemiology Reports*, 7:190–202.

Sofer, T., Richardson, D. B., Colicino, E., Schwartz, J., and Tchetgen, E. J. T. (2016). On Negative Outcome Control of Unobserved Confounding as a Generalization of Difference-in-Differences. *Statistical Science*, 31(3):348 – 361.

Taddeo, M. M., Amorim, L. D., and Aquino, R. (2022). Causal measures using generalized difference-in-difference approach with nonlinear models. *Statistics and Its Interface*, 15(4):399 – 413.

Tchetgen, E. J. T., Ying, A., Cui, Y., Shi, X., and Miao, W. (2020). An introduction to proximal causal learning.

Tchetgen, E. T., Park, C., and Richardson, D. (2023a). Single proxy control.

Tchetgen, E. T., Park, C., and Richardson, D. (2023b). Universal difference-in-differences for causal inference in epidemiology.

Ye, T., Chen, S., and Zhang, B. (2022). The role of placebo samples in observational studies.

Ye, T., Keele, L., Hasegawa, R., and Small, D. S. (2020). A negative correlation strategy for bracketing in difference-in-differences.

Zhang, M. and Ding, P. (2022). Interpretable sensitivity analysis for the baron-kenny approach to mediation with unmeasured confounding.

A Double placebo with multiple unobserved confounders

Let us assume we have the causal graph in Figure 4(b). We now want to assume that there are multiple unobserved confounders captured by the vector of variables $\mathbf{Z}$. We do not assume that the same linear combination of the components of $\mathbf{Z}$ will de-confound all the relationships we care about. $Z_{(\text{Y.DPX})} = \mathbf{Z}\beta_{Y\sim Z|D,P,X}$ is the linear combination that de-confounds the Y, D and the Y, P relationship. $Z_{(\text{N.DPX})} = \mathbf{Z}\beta_{N\sim Z|D,P,X}$ is the linear combination that de-confounds the N, D and the N, P relationship. Consider the following OVB expressions, where we use the relevant linear combination for each relationship.

$$\beta_{Y\sim D|P,X} - \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim Z_{(\text{Y.DPX})}|D,P,X}\beta_{Z_{(\text{Y.DPX})}\sim D|P,X} \tag{41a}$$
$$\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim Z_{(\text{Y.DPX})}|D,P,X}\beta_{Z_{(\text{Y.DPX})}\sim P|D,X} \tag{41b}$$
$$\beta_{N\sim D|P,X} - \beta_{N\sim D|P,Z_{(\text{N.DPX})},X} = \beta_{N\sim Z_{(\text{N.DPX})}|D,P,X}\beta_{Z_{(\text{N.DPX})}\sim D|P,X} \tag{41c}$$
$$\beta_{N\sim P|D,X} - \beta_{N\sim P|D,Z_{(\text{N.DPX})},X} = \beta_{N\sim Z_{(\text{N.DPX})}|D,P,X}\beta_{Z_{(\text{N.DPX})}\sim P|D,X} \tag{41d}$$

We can re-write the middle two of these as follows.

$$\beta_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} = \frac{\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X}}{\beta_{Z_{(\text{Y.DPX})}\sim P|D,X}} \tag{42a}$$
$$\beta_{Z_{(\text{N.DPX})}\sim D|P,X} = \frac{\beta_{N\sim D|P,X} - \beta_{N\sim D|P,Z_{(\text{N.DPX})},X}}{\beta_{N\sim Z_{(\text{N.DPX})}|D,P,X}} \tag{42b}$$

Now we note that there are m_1, m_2, m_3, and m_4 that satisfy the following.

$$\beta_{Z_{(\text{Y.DPX})}\sim D|P,X} = m_{\text{A}} \times \beta_{Z_{(\text{N.DPX})}\sim D|P,X} \tag{43a}$$
$$\beta_{Z_{(\text{Y.DPX})}\sim P|D,X} = m_{\text{B}} \times \beta_{Z_{(\text{N.DPX})}\sim P|D,X} \tag{43b}$$

Then we can get the following expression for $\beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X}$.

$$\begin{aligned}
\beta_{Y\sim D|P,X} - \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} &= \beta_{Y\sim Z_{(\text{Y.DPX})}|D,P,X}\beta_{Z_{(\text{Y.DPX})}\sim D|P,X} \\
&= m_{\text{A}} \times \beta_{Y\sim Z_{(\text{Y.DPX})}|D,P,X}\beta_{Z_{(\text{N.DPX})}\sim D|P,X} \\
&= m_{\text{A}} \times \left[\frac{\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X}}{\beta_{Z_{(\text{Y.DPX})}\sim P|D,X}}\right]\left[\frac{\beta_{N\sim D|P,X} - \beta_{N\sim D|P,Z_{(\text{N.DPX})},X}}{\beta_{N\sim Z_{(\text{N.DPX})}|D,P,X}}\right] \\
&= \frac{m_{\text{A}}}{m_{\text{B}}} \times \frac{(\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X})(\beta_{N\sim D|P,X} - \beta_{N\sim D|P,Z_{(\text{N.DPX})},X})}{[\beta_{N\sim P|D,X} - \beta_{N\sim P|D,Z_{(\text{N.DPX})},X}]} \\
\therefore \beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} &= \beta_{Y\sim D|P,X} - \frac{m_{\text{A}}}{m_{\text{B}}} \times \frac{(\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X})(\beta_{N\sim D|P,X} - \beta_{N\sim D|P,Z_{(\text{N.DPX})},X})}{(\beta_{N\sim P|D,X} - \beta_{N\sim P|D,Z_{(\text{N.DPX})},X})}
\end{aligned} \tag{44}$$

$$\begin{aligned}
\text{Now, } \frac{m_{\text{A}}}{m_{\text{B}}} &= \frac{\beta_{Z_{(\text{Y.DPX})}\sim D|P,X}}{\beta_{Z_{(\text{N.DPX})}\sim D|P,X}}\frac{\beta_{Z_{(\text{Y.DPX})}\sim P|D,X}}{\beta_{Z_{(\text{N.DPX})}\sim P|D,X}} \\
&= \frac{R_{Z_{(\text{Y.DPX})}\sim D|P,X}R_{Z_{(\text{N.DPX})}\sim P|D,X}}{R_{Z_{(\text{N.DPX})}\sim D|P,X}R_{Z_{(\text{Y.DPX})}\sim P|D,X}} \times \frac{\text{SD}(Z^{\perp P,X}_{(\text{Y.DPX})})}{\text{SD}(Z^{\perp P,X}_{(\text{N.DPX})})}\frac{\text{SD}(Z^{\perp D,X}_{(\text{N.DPX})})}{\text{SD}(Z^{\perp D,X}_{(\text{Y.DPX})})} \\
&= \frac{R_{Z_{(\text{Y.DPX})}\sim D|P,X}R_{Z_{(\text{N.DPX})}\sim P|D,X}}{R_{Z_{(\text{N.DPX})}\sim D|P,X}R_{Z_{(\text{Y.DPX})}\sim P|D,X}} \times \frac{\text{SD}(Z^{\perp P,X}_{(\text{Y.DPX})})}{\text{SD}(Z^{\perp D,P,X}_{(\text{Y.DPX})})}\frac{\text{SD}(Z^{\perp D,P,X}_{(\text{N.DPX})})}{\text{SD}(Z^{\perp P,X}_{(\text{N.DPX})})}\frac{\text{SD}(Z^{\perp D,X}_{(\text{N.DPX})})}{\text{SD}(Z^{\perp D,P,X}_{(\text{N.DPX})})}\frac{\text{SD}(Z^{\perp D,P,X}_{(\text{Y.DPX})})}{\text{SD}(Z^{\perp D,X}_{(\text{Y.DPX})})} \\
&= \frac{R_{Z_{(\text{Y.DPX})}\sim D|P,X}R_{Z_{(\text{N.DPX})}\sim P|D,X}}{R_{Z_{(\text{N.DPX})}\sim D|P,X}R_{Z_{(\text{Y.DPX})}\sim P|D,X}} \times \frac{\sqrt{1-R^2_{Z_{(\text{N.DPX})}\sim D|P,X}}\sqrt{1-R^2_{Z_{(\text{Y.DPX})}\sim P|D,X}}}{\sqrt{1-R^2_{Z_{(\text{Y.DPX})}\sim D|P,X}}\sqrt{1-R^2_{Z_{(\text{N.DPX})}\sim P|D,X}}} \\
&= \frac{f_{Z_{(\text{Y.DPX})}\sim D|P,X}f_{Z_{(\text{N.DPX})}\sim P|D,X}}{f_{Z_{(\text{N.DPX})}\sim D|P,X}f_{Z_{(\text{Y.DPX})}\sim P|D,X}} \\
\therefore \frac{m_{\text{A}}}{m_{\text{B}}} &= \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X}f_{Z_{(\text{Y.DPX})}\sim P|D,X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X}f_{Z_{(\text{Y.DPX})}\sim P|D,X}} \times \frac{R_{N\sim Z_{(\text{N.DPX})}|D,P,X}f_{Z_{(\text{N.DPX})}\sim P|D,X}}{R_{N\sim Z_{(\text{N.DPX})}|D,P,X}f_{Z_{(\text{N.DPX})}\sim D|P,X}} \triangleq k_{(\text{YD / YP})} \times k_{(\text{NP / ND})}
\end{aligned} \tag{45}$$

Thus, finally, we arrive at the expression for $\beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X}$ below.

$$\beta_{Y\sim D|P,Z_{(\text{Y.DPX})},X} = \beta_{Y\sim D|P,X} - k_{(\text{YD / YP})} \times k_{(\text{NP / ND})} \times \frac{(\beta_{Y\sim P|D,X} - \beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X})(\beta_{N\sim D|P,X} - \beta_{N\sim D|P,Z_{(\text{N.DPX})},X})}{(\beta_{N\sim P|D,X} - \beta_{N\sim P|D,Z_{(\text{N.DPX})},X})} \tag{46}$$

Some observations:

- When $Z_{(\text{Y.DPX})} = Z_{(\text{N.DPX})}$ or both $R_{Z_{(\text{Y.DPX})}\sim D|P,X} = R_{Z_{(\text{N.DPX})}\sim D|P,X}$ and $R_{Z_{(\text{Y.DPX})}\sim P|D,X} = R_{Z_{(\text{N.DPX})}\sim P|D,X}$, then $k_{(\text{YD / YP})} \times k_{(\text{NP / ND})} = 1$.
- $k_{(\text{YD / YP})}$ captures the scaled ratio of the level of confounding of the Y, D relationship (conditional on P) to the level of confounding of the Y, P relationship (conditional on D), after re-scaling one of these biases, or the ratio of bias factors.
- $k_{(\text{NP / ND})}$ captures the scaled ratio of the level of confounding of the N, P relationship (conditional on D) to the level of confounding of the N, D relationship (conditional on P), after re-scaling one of these biases, or the ratio of bias factors.
- $\beta_{Y\sim P|D,Z_{(\text{Y.DPX})},X}$ measures the causal of P on Y (conditional on D).
- $\beta_{N\sim D|P,Z_{(\text{N.DPX})},X}$ measures the causal of D on N (conditional on P).
- $\beta_{N\sim P|D,Z_{(\text{N.DPX})},X}$ measures the causal of P on N (conditional on D).
- We could also use $k_{(\text{YD / ND})} \times k_{(\text{NP / YP})} = \frac{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{Z_{(\text{Y.DPX})}\sim P|D,X}}{R_{N\sim Z_{(\text{N.DPX})}|D,P,X} f_{Z_{(\text{N.DPX})}\sim D|P,X}} \times \frac{R_{N\sim Z_{(\text{N.DPX})}|D,P,X} f_{Z_{(\text{N.DPX})}\sim P|D,X}}{R_{Y\sim Z_{(\text{Y.DPX})}|D,P,X} f_{Z_{(\text{Y.DPX})}\sim P|D,X}}$ and have similar interpretation.

B Partially linear and non-parametric framework

If investigators are interested in partially linear of fully non-parametric approaches to estimating treatment effects, we can use the omitted variables frameworks from Chernozhukov et al. (2022) to arrive at partial identification frameworks similar to those presented in the main text for settings in which placebo outcomes are available. We do not demonstrate the partially linear or non-parametric case for placebo treatments, but in principle this is also possible.

B.1 Partially linear framework

An investigator may be interested in estimating a partially linear model. Y is the outcome; N is a placebo outcome; D is the treatment; $\mathbf{X}$ is a vector of observed covariates; $\mathbf{Z}$ is a vector of unobserved covariates. We consider the following short and long partially linear models.

$$Y = \theta_{l,Y} D + f_l(\mathbf{X}, \mathbf{Z}) + \epsilon_l \tag{47a}$$
$$Y = \theta_{s,Y} D + f_s(\mathbf{X}) + \epsilon_s \tag{47b}$$

$$N = \theta_{l,N} D + g_l(\mathbf{X}, \mathbf{Z}) + \xi_l \tag{48a}$$
$$N = \theta_{s,N} D + g_s(\mathbf{X}) + \xi_s \tag{48b}$$

There is some m such that $\text{bias}_{(\text{YD.X})} = m \times \text{bias}_{(\text{ND.X})}$, where $\text{bias}_{(\text{YD.X})} \triangleq \theta_{s,Y} - \theta_{l,Y}$ and $\text{bias}_{(\text{ND.X})} \triangleq \theta_{s,N} - \theta_{l,N}$. This implies that Equation 49 holds. Using Equation 49 for partial identification will run into the issue that $m \triangleq \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{ND.X})}}$ may be difficult to interpret when Y and N have differing scales. Can we do better?

$$\theta_{l,Y} = \theta_{s,Y} - m \times (\theta_{s,N} - \theta_{l,N}) \tag{49}$$

Chernozhukov et al. (2022) show that omitted variable bias in partially linear setting can be bounded in the following way.

$$|\theta_{s,Y} - \theta_{l,Y}|^2 = |\text{bias}_{(\text{YD.X})}|^2 \leq S_Y^2 \times \eta_{Y\sim Z|D,X}^2 \times \frac{\eta_{D\sim Z|X}^2}{1 - \eta_{D\sim Z|X}^2} = S_Y^2 \times \text{BF}_{(\text{YD.X})} \tag{50a}$$

$$|\theta_{s,N} - \theta_{l,N}|^2 = |\text{bias}_{(\text{ND.X})}|^2 \leq S_N^2 \times \eta_{N\sim Z|D,X}^2 \times \frac{\eta_{D\sim Z|X}^2}{1 - \eta_{D\sim Z|X}^2} = S_N^2 \times \text{BF}_{(\text{ND.X})} \tag{50b}$$

$\text{BF}_{(\text{YD.X})} \triangleq \eta_{Y\sim Z|D,X}^2 \frac{\eta_{D\sim Z|X}^2}{1-\eta_{D\sim Z|X}^2}$ and $\text{BF}_{(\text{ND.X})} \triangleq \eta_{N\sim Z|D,X}^2 \frac{\eta_{D\sim Z|X}^2}{1-\eta_{D\sim Z|X}^2}$ are "bias factors" that are scaled versions of the maximum level of unobserved confounding from $\mathbf{Z}$. S_Y^2 and S_N^2 are identified and therefore estimable from the observed data.

$\eta^2_{Y\sim Z|D,X}$ and $\eta^2_{D\sim Z|X}$ are Pearson's correlation ratios (or the non-parametric R^2s).[22] $\eta^2_{Y\sim Z|D,X}$ is the proportion of residual variation in Y explained by Z. $\eta^2_{D\sim Z|X}$ is the proportion of residual variation in D explained by Z. See Chernozhukov et al. (2022) for further discussion of how to interpret partial η^2s and for more on S^2_Y and S^2_N.

There is some k such that $\text{BF}_{(\text{YD.X})} = k \times |\theta_{s,N} - \theta_{l,N}|^2 \times \frac{1}{S^2_N}$, where the right hand side contains a scaled version of the actual level of confounding from $\mathbf{Z}$ for N and D. The left hand side contains the scaled version of the maximum level of confounding from $\mathbf{Z}$ for Y and D. We can plug this into the expression for $|\text{bias}_{(\text{YD.X})}|^2$, to get the following.

$$\begin{aligned} |\text{bias}_{(\text{YD.X})}|^2 &\leq \left[k \times |\theta_{s,N} - \theta_{l,N}|^2 \times \frac{1}{S^2_N}\right] \times S^2_Y \\ \iff |\text{bias}_{(\text{YD.X})}| &\leq \sqrt{k} \times |\theta_{s,N} - \theta_{l,N}| \times \frac{S_Y}{S_N} \end{aligned} \tag{51}$$

This is similar to our bias expression for the fully linear version. We have a bound on bias in terms of quantities that can be estimated from the data, k, and $\theta_{l,N}$. We know that $k = \frac{\text{BF}_{(\text{YD.X})}}{|\theta_{s,N}-\theta_{l,N}|^2 \times \frac{1}{S^2_N}}$ is the ratio of the maximum level of confounding from $\mathbf{Z}$ for Y and D to the actual level of confounding from $\mathbf{Z}$ for N and D. We also have that $\text{BF}_{(\text{ND.X})}$ is a bound on $|\theta_{s,N} - \theta_{l,N}|^2 \times \frac{1}{S^2_N}$. Therefore, we see that $k = \frac{\text{BF}_{(\text{YD.X})}}{|\theta_{s,N}-\theta_{l,N}|^2 \times \frac{1}{S^2_N}} \geq \frac{\text{BF}_{(\text{YD.X})}}{\text{BF}_{(\text{ND.X})}} \triangleq k^*$. If we assume what Chernozhukov et al. (2022) call "adversarial confounding," i.e. that $|\text{bias}_{(\text{ND.X})}|^2 = S^2_N \times \text{BF}_{(\text{ND.X})}$ for N and D, then we can reason in terms of k^*, rather than k. See Chernozhukov et al. (2022) for a discussion of what such an assumption entails. If we additionally assume adversarial confounding for Y and D, then we can write the following bias expression, where k^* represents the relative level of confounding where differences in scale have been accounted for.

$$|\text{bias}_{(\text{YD.X})}| = \sqrt{k^*} \times |\theta_{s,N} - \theta_{l,N}| \times \frac{S_Y}{S_N} \tag{52}$$

Finally, note that we could write $k^* = \frac{\text{BF}_{(\text{YD.X})}}{\text{BF}_{(\text{ND.X})}} = \frac{\eta^2_{Y\sim Z|D,X} \frac{\eta^2_{D\sim Z|X}}{1-\eta^2_{D\sim Z|X}}}{\eta^2_{N\sim Z|D,X} \frac{\eta^2_{D\sim Z|X}}{1-\eta^2_{D\sim Z|X}}} = \frac{\eta^2_{Y\sim Z|D,X}}{\eta^2_{N\sim Z|D,X}}$, which allow for a second interpretation for k^* as the ratio of the strengths of relationship between the outcomes and the unobserved $\mathbf{Z}$.

B.2 Non-parametric framework

An investigator may also be interested in estimating a linear functional of the conditional expectation function of the outcome in a fully non-parametric setting, like $\theta_{l,Y} = \mathbb{E}[Y_1 - Y_0] = \mathbb{E}[f_Y(1, \mathbf{X}, \mathbf{Z}) - f_Y(0, \mathbf{X}, \mathbf{Z})]$ for a binary treatment D, where $Y_d = f_Y(d, \mathbf{X}, \mathbf{Z}, U_Y)$ is the equation for Y in the structural causal model under intervention to set $D = d$. Again, the investigator is only able to estimate $\theta_{s,Y} = \mathbb{E}[f^*_Y(1, \mathbf{X}) - f^*_Y(0, \mathbf{X})]$, where $f^*_Y(D, X) \triangleq \mathbb{E}[Y|D, \mathbf{X}] = \mathbb{E}[f_Y(D, X, \mathbf{Z})|D, \mathbf{X}]$. We define $\theta_{l,N}$ and $\theta_{s,N}$ similarly for a placebo outcome. We could again consider an approach like Equation 53. But again $m \triangleq \frac{\text{bias}_{(\text{YD.X})}}{\text{bias}_{(\text{ND.X})}}$ may be difficult to interpret when Y and N have differing scales.

$$\theta_{l,Y} = \theta_{s,Y} - m \times (\theta_{s,N} - \theta_{l,N}) \tag{53}$$

Chernozhukov et al. (2022) also show that omitted variable bias in fully non-parametric setting can be bounded by an expression in terms of $\eta^2_{Y\sim Z|D,X}$ and a second term that, in the case of targeting $\theta_{l,Y} = \mathbb{E}[Y_1 - Y_0]$ with a binary treatment D, is the "average gain in the conditional precision with which we predict D by using Z in addition to X," which is somewhat similar to $\eta^2_{D\sim Z|X}$. Specifically, we have the following.

$$|\theta_{s,Y} - \theta_{l,Y}|^2 = |\text{bias}_{(\text{YD.X})}|^2 \leq S^2_Y \times \eta^2_{Y\sim Z|D,X} \times \frac{1 - R^2_{\alpha\sim\alpha_s}}{R^2_{\alpha\sim\alpha_s}} = S^2_Y \times \text{BF}_{(\text{YD.X})} \tag{54a}$$

$$|\theta_{s,N} - \theta_{l,N}|^2 = |\text{bias}_{(\text{YD.X})}|^2 \leq S^2_N \times \eta^2_{N\sim Z|D,X} \times \frac{1 - R^2_{\alpha\sim\alpha_s}}{R^2_{\alpha\sim\alpha_s}} = S^2_N \times \text{BF}_{(\text{ND.X})} \tag{54b}$$

$\text{BF}_{(\text{YD.X})} \triangleq \eta^2_{Y\sim Z|D,X} \frac{1-R^2_{\alpha\sim\alpha_s}}{R^2_{\alpha\sim\alpha_s}}$ and $\text{BF}_{(\text{ND.X})} \triangleq \eta^2_{N\sim Z|D,X} \frac{1-R^2_{\alpha\sim\alpha_s}}{R^2_{\alpha\sim\alpha_s}}$ are "bias factors" that are scaled versions of the maximum level of unobserved confounding from $\mathbf{Z}$. S^2_Y and S^2_N are identified and therefore estimable from the data. See

[22] $\eta^2_{D\sim Z|X} = \frac{\text{Var}(\mathbb{E}[D|Z,X]) - \text{Var}(\mathbb{E}[D|X])}{\text{Var}(D) - \text{Var}(\mathbb{E}[D|X])} = \frac{\eta^2_{D\sim Z,X} - \eta^2_{D\sim X}}{1-\eta^2_{D\sim X}}$. $\eta^2_{Y\sim Z|D,X}$ can be similarly interpreted.

Chernozhukov et al. (2022) for discussion of S_Y^2, S_N^2, and $R^2_{\alpha\sim\alpha_s}$. Proceeding as in the last section, we can arrive at Equation 55. Again, $k \triangleq \frac{\text{BF}_{(\text{YD.X})}}{|\theta_{s,N}-\theta_{l,N}|^2\times\frac{1}{S_N^2}}$. If we assume adversarial confounding for both Y and D and N and D, then we can write Equation 56, where $k^* \triangleq \frac{\text{BF}_{(\text{YD.X})}}{\text{BF}_{(\text{ND.X})}}$ represents the relative level of confounding where differences in scale have been accounted for.

$$|\text{bias}_{(\text{YD.X})}| \leq \sqrt{k} \times |\theta_{s,N} - \theta_{l,N}| \times \frac{S_Y}{S_N} \tag{55}$$

$$|\text{bias}_{(\text{YD.X})}| = \sqrt{k^*} \times |\theta_{s,N} - \theta_{l,N}| \times \frac{S_Y}{S_N} \tag{56}$$

Again, note that we could write $k^* = \frac{\text{BF}_{(\text{YD.X})}}{\text{BF}_{(\text{ND.X})}} = \frac{\eta^2_{Y\sim Z|D,X}\frac{1-R^2_{\alpha\sim\alpha_s}}{R^2_{\alpha\sim\alpha_s}}}{\eta^2_{N\sim Z|D,X}\frac{1-R^2_{\alpha\sim\alpha_s}}{R^2_{\alpha\sim\alpha_s}}} = \frac{\eta^2_{Y\sim Z|D,X}}{\eta^2_{N\sim Z|D,X}}$, which allow for a second interpretation for k^* as the ratio of the strengths of relationship between the outcomes and the unobserved $\mathbf{Z}$.

C Re-expression of bias factors

$$\begin{aligned}
\beta_{Y\sim D|X} - \beta_{Y\sim D|Z,X} &= R_{Y\sim Z|D,X} f_{D\sim Z|X} \frac{\text{SD}(Y^{\perp D,X})}{\text{SD}(D^{\perp X})} \\
\implies R_{Y\sim Z|D,X} f_{D\sim Z|X} &= (\beta_{Y\sim D|X} - \beta_{Y\sim D|Z,X}) \frac{\text{SD}(D^{\perp X})}{\text{SD}(Y^{\perp D,X})} \\
&= R_{Y\sim D|X} \frac{\text{SD}(Y^{\perp X})}{\text{SD}(D^{\perp X})} \frac{\text{SD}(D^{\perp X})}{\text{SD}(Y^{\perp D,X})} - R_{Y\sim D|Z,X} \frac{\text{SD}(Y^{\perp X,Z})}{\text{SD}(D^{\perp X,Z})} \frac{\text{SD}(D^{\perp X})}{\text{SD}(Y^{\perp D,X})} \frac{\text{SD}(Y^{\perp D,X,Z})}{\text{SD}(Y^{\perp D,X,Z})} \\
&= f_{Y\sim D|X} - f_{Y\sim D|Z,X} \frac{\text{SD}(Y^{\perp D,X,Z})}{\text{SD}(D^{\perp X,Z})} \frac{\text{SD}(D^{\perp X})}{\text{SD}(Y^{\perp D,X})} = f_{Y\sim D|X} - f_{Y\sim D|Z,X} \sqrt{\frac{1-R^2_{Y\sim Z|D,X}}{1-R^2_{D\sim Z|X}}} \\
\implies 1 &= \frac{f_{Y\sim D|X}}{R_{Y\sim Z|D,X} f_{D\sim Z|X}} - \frac{f_{Y\sim D|Z,X}}{f_{Y\sim Z|D,X} R_{D\sim Z|X}}
\end{aligned}$$