Title:

How Conditioning on Pretests of the Outcome Removes or even Increases Bias in Effect Estimates from Observational Data

Authors and Affiliations:

Yongnam Kim, University of Wisconsin-Madison
Peter M. Steiner, University of Wisconsin-Madison
Background

One of the most popular methods to reduce confounding bias from observational studies is to condition on pretest measures of the outcome via regression (ANCOVA), matching or propensity score adjustments. Many studies strongly support the special role of pretests in bias-reduction (e.g., Cook & Steiner, 2010; Hallberg, Cook, Steiner, & Clark, 2016; Wong, Valentine, & Miller-Bains, 2017). They frequently argue that pretests often succeed in removing a major part of confounding bias because of their high correlation with the posttest (outcome). However, Hallberg et al. (2016) also found in their empirical analyses that a high correlation does not always ensure a significant bias-reduction. Morgan and Winship (2015) even argued that conditioning on pretests can induce a new bias. Despite some empirical evidence that pretests often work well for reducing bias, it is still not clear how and under which conditions pretests reduce (or increase) the overall bias in effect estimates.

Focus of Study

In this paper, we use graphical models to investigate and demonstrate the conditions under which conditioning on a pretest decreases or increases bias. Since measurement errors in pre- and posttests are often correlated (because they are often measured with the same instrument in a similar setting), we particularly focus on their error structure, and show in a linear regression framework how the errors’ (in)dependence affects bias-reduction.

Findings

Suppose we are interested in the causal effect of attending a math camp on math achievement. Figure 1 shows the causal graph for a pretest-posttest design where the unmeasured math ability, $A$, confounds the causal relation between attending the camp, $Z$, and the math posttest score, $Y$. The math pretest score, $P$, is also affected by the ability, $A$, but is unaffected by camp attendance, $Z$, because $P$ is measured before the camp starts. The pretest and posttest are contaminated with measurement errors, $e_p$ and $e_y$, which may be correlated with each other, indicated by $\rho$.

From the graph in Figure 1, we see that conditioning on $P$ leaves two non-causal paths open:

(i) $Z \leftarrow A \rightarrow Y$

(ii) $Z \leftarrow A \rightarrow P \leftarrow e_p \leftarrow \cdots e_y \rightarrow Y$.

The first path is the original confounding path due to $A$ which cannot be fully blocked by the pretest $P$, because $P$ is not on this path. The second path is a new open path because $P$ is a collider on the path (i.e., $A \rightarrow P \leftarrow e_p$) and conditioning on a collider opens the path (Elwert & Winship, 2014). Because of these two open non-causal paths, the causal effect of $Z$ on $Y$ is generally not identified and a corresponding effect estimate would remain biased (backdoor criterion; Pearl, 2009). While the bias due to the first open path is referred to as confounding bias, the bias induced by the second path is referred to as collider bias. Morgan and Winship (2015) warned us about this collider bias when conditioning on a pretest.

Although conditioning on a pretest cannot eliminate the entire bias, it may nonetheless decrease the overall bias. In linear systems, we can show that the overall bias is the sum of the confounding bias (via the first path) and collider bias (via the second path). In many practical settings it is plausible that the impact of the ability on the pretest and posttest has the same sign, $\text{sgn}(\beta_i) = \text{sgn}(\beta_z)$, and the correlation between the error terms is positive, $\text{sgn}(\rho) \geq 0$. Under
these assumptions, one can then show that the signs of the confounding and collider biases are different,

\[ \text{sgn}(Z \leftarrow A \rightarrow Y) \neq \text{sgn}(Z \leftarrow A \rightarrow P \leftarrow e_P \leftarrow\cdots e_y \rightarrow Y) \]

such that conditioning on \( P \) results in a (partial) offsetting of the confounding and collider biases. This means that even if the pretest is not able to fully remove the confounding bias, the collider bias via the correlated error terms might further reduce the bias. The finally remaining bias depends on the extent of confounding bias removed by \( P \) and of collider bias induced by conditioning on \( P \). However, whether correlated error terms between the pre- and posttest are beneficial for removing bias depends on the specific data-generating scenario.

Figure 2 presents simulation results for three different data-generating scenarios. The plots show the remaining bias after conditioning on the pretest in a regression analysis (y-axis) as a function of the error terms’ correlation, \( \rho \) (x-axis). The three plots differ with respect to the ability’s effect on the pretest and posttest (i.e., \( \beta_1 \) and \( \beta_2 \)). For \( \beta_1 = \beta_2 = 1.0 \) (left panel), a stronger correlation between the error terms implies less remaining bias. This is because the negative collider bias (since all parameters are positive) further offsets the positive confounding bias left due to unreliable pretest measure. However, stronger error correlations do not always imply that more bias is removed. For instance, if the ability strongly affects the pretest than the posttest, \( \beta_1 = 1.0 \) and \( \beta_2 = .1 \) (middle panel), then the absolute remaining bias resulting from highly correlated errors can be larger than the initial bias (solid line). That is, the negative collider bias strongly dominates the positive confounding bias and result in negative bias. Such a situation is impossible if the ability has a stronger effect on the posttest than the pretest, \( \beta_1 = .1 \) and \( \beta_2 = 1.0 \) (right panel). In our settings, a higher correlation between the error terms (\( \rho \)) directly implies a higher correlation between the pretest and posttest. But, as the middle panel demonstrates, a higher pretest-posttest correlation (mainly due to the correlated errors) does not imply more bias-reduction as several articles suggested.

Conclusion

This paper is a first step in theoretically explaining the special role of the pretest in removing bias from observational data. First, we show that Morgan and Winship’s (2015) claim that conditioning on a pretest can induce bias is compatible with empirical findings. We show that such collider bias can occur because of correlated measurement errors between the pretest and posttest. In many pretest-posttest designs, the collider bias will likely partially offset confounding bias, therefore, contribute to reducing overall bias. Second, we also show that if the impact of the unmeasured confounder on the posttest is smaller than on the pretest, the collider bias may dominate the confounding bias and result in more absolute bias even though the pretest and posttest are highly correlated.
References
Figure 1. Causal graph of a pretest-posttest design with correlated measurement errors. Math ability $A$ and two measurement errors, $e_P$ and $e_Y$, are represented by vacant nodes because they are unmeasured.

Figure 2. Bias of ANCOVA estimators (i.e., regressing $Y$ on $Z$ and $P$), varying the correlation of the error terms (\(\rho\)) and the relative magnitudes of the impacts of $A$ on $P$ and $Y$ (\(\beta_1\) and \(\beta_2\)). Across the simulations, we use the following parameters: $\alpha = 1$, $\tau = 0$, $\text{Var}(A) = 1$, $\text{Var}(e_P) = 1$, $\text{Var}(e_Y) = 1$ and $\text{Var}(e_Z) = 1$, where $e_Z$ denotes the error term of the treatment $Z$. The horizontal solid lines represent the initial biases when $P$ is not conditioned on (i.e., regressing $Y$ on $Z$).