

Demystifying the Relationship Between Fixed/Random Effects and Unmeasured Confounding in Panel Data Analysis

Yicheng Shen, Yanjiao Yang, Huiyin Lin

23 April, 2023

1 Introduction

Panel data analysis is a widely used statistical tool in econometrics to study the relationships between variables over time. However, this type of analysis is susceptible to confounding variables, which can bias the estimated coefficients and distort the interpretation of the results. Fixed and random effects models frequently emerge in panel data analysis to account for confounding variables that are time-invariant (Gunasekara et al. 2014) or time-varying (Li, Chen, and Gao 2011; Ahn, Lee, and Schmidt 2013). A common belief among econometricians is that fixed and random effects models can absorb unmeasured confounding variables (Angrist and Pischke 2009), but the mechanism behind this claim is mysterious and not well-understood. In this research paper, we aim to explore the relationship between fixed/random effects and time-invariant unmeasured confounding in panel data analysis and provide insights into whether and in what sense these models can address this issue.

Our simulation suggests that fixed/random effects can remedy the issue of unmeasured confounding to a certain extent, but there are still systematic bias resulting from the relationship between unmeasured confounding and the treatment assignment. **[more stuff here]**

2 Background

In typical observational studies, failing to capture significant unmeasured confounding gives rise to biased estimates of treatment effects, which compels practitioners to develop methods of assessing and handling uncontrolled confounding (VanderWeele and Arah 2011). The question of whether fixed/random effects models can account for unmeasured confounding in panel data analysis has been the subject of much debate in the econometrics literature. A number of studies have explored this issue from different angles and with varying degrees of empirical evidence.

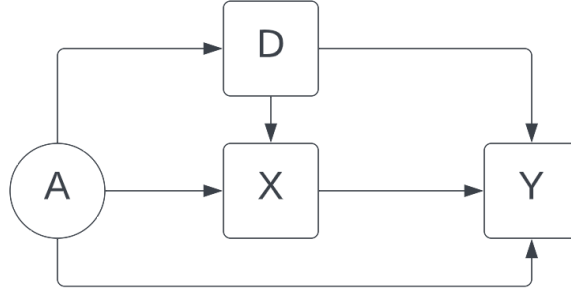
One line of research has focused on theoretical arguments for why fixed/random effects models might be effective at absorbing unmeasured confounding. For example, Angrist and Pischke (2009) discuss the strategies that use data with a time or cohort dimension to control for unobserved-but-fixed omitted variables. Hausman and Taylor (1981) argue that fixed effects models can control for the time-invariant confounders by essentially differencing them out, while random effects models can account for time-varying confounders that are uncorrelated with the fixed effects. More recently, Wooldridge (2010) has suggested that fixed effects models can be viewed as a form of quasi-experimental design that mimics a randomized controlled trial, and thus can address the unobserved component to the extent that such designs do.

However, empirical evidence for the effectiveness of fixed/random effects models at absorbing unmeasured confounding has been mixed. Some studies have found that these models can indeed account for unmeasured confounding to some extent. For example, Bell and Jones (2015) used a simulation study to show that fixed effects models can provide unbiased estimates of treatment effects even in the presence of unobserved heterogeneity. Similarly, Chandrasekhar and Zhan (2014) found that random effects models can help to address omitted variable bias in studies of firm-level productivity.

Other studies, however, have challenged the notion that fixed/random effects models can fully account for unmeasured confounding. For example, Mundlak (1978) argued that random effects models are biased when unobserved heterogeneity is correlated with observed variables, and that fixed effects models are limited by the fact that they cannot estimate time-invariant covariates. Similarly, Bell and Jones (2015) noted that while fixed effects models can provide reasonable estimates of treatment effects, they may still suffer from omitted variable bias if the unobserved confounding variable is correlated with the treatment.

3 Method

We presume that specifying a random effect (intercept) for every individual in the study is an appropriate method to address unmeasured confounding, A .



Our data generation process (DGP) uses the following model specifications:

$$E(Y_{it} | A_i, X_{it}, t, D_{it}) = \alpha + \lambda_t + \rho D_{it} + \gamma A_i + \delta X_{it}$$

$$Y_{it} = \alpha + \lambda_t + \rho D_{it} + \gamma A_i + \delta X_{it} + \epsilon_{it}$$

Difference-in-difference (DID) estimator

$$\tau^{DID} = (\bar{Y}_{1,t+1} - \bar{Y}_{1,t}) - (\bar{Y}_{0,t+1} - \bar{Y}_{0,t})$$

Regression model

$$Y_{it} = \alpha + \lambda_t + \rho D_{it} + \delta X_{it} + \epsilon_{it}$$

Fixed effects model

$$\alpha_i \equiv \alpha + \gamma A_i$$

$$Y_{it} = \alpha_i + \lambda_t + \rho D_{it} + \delta X_{it} + \epsilon_{it}$$

Random effects model

$$\alpha_i \equiv \alpha + \gamma A_i$$

$$Y_{it} = \alpha_i + \lambda_t + \rho D_{it} + \delta X_{it} + \epsilon_{it}$$

$$\text{where } \alpha_i \sim N(\mu_i, \sigma^2)$$

4 Result

5 Discussion

Limitation and further direction:

Time-variant confounding?

More time points?

Coefficients of latent variable?

Overall, the literature suggests that while fixed/random effects models may be useful in controlling for unmeasured confounding in panel data analysis, they are not a panacea. Other methods, such as instrumental variables or regression discontinuity designs, may be more necessary in certain cases to fully address this issue.

6 Reference

- Ahn, Seung C, Young H Lee, and Peter Schmidt. 2013. “Panel Data Models with Multiple Time-Varying Individual Effects.” *Journal of Econometrics* 174 (1): 1–14.
- Angrist, Joshua D, and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist’s Companion, Chapter 5*. Princeton university press.
- Gunasekara, Fiona Imlach, Ken Richardson, Kristie Carter, and Tony Blakely. 2014. “Fixed Effects Analysis of Repeated Measures Data.” *International Journal of Epidemiology* 43 (1): 264–69.
- Hausman, Jerry A, and William E Taylor. 1981. “Panel Data and Unobservable Individual Effects.” *Econometrica: Journal of the Econometric Society*, 1377–98.
- Li, Degui, Jia Chen, and Jiti Gao. 2011. “Non-Parametric Time-Varying Coefficient Panel Data Models with Fixed Effects.” *The Econometrics Journal* 14 (3): 387–408.
- VanderWeele, Tyler J, and Onyebuchi A Arah. 2011. “Unmeasured Confounding for General Outcomes, Treatments, and Confounders: Bias Formulas for Sensitivity Analysis.” *Epidemiology (Cambridge, Mass.)* 22 (1): 42.