Common Errors: How to (and Not to) Control for Unobserved Heterogeneity*

Todd A. Gormley† and David A. Matsa‡

June 6, 2012

Abstract

Controlling for unobserved heterogeneity (or “common errors”), such as industry-specific shocks, is a fundamental challenge in empirical research, as failing to do so can introduce omitted variables biases and preclude causal inference. This paper discusses limitations of two approaches commonly used to control for unobserved group-level heterogeneity in finance research—demeaning the dependent variable with respect to the group (e.g., “industry-adjusting”) and adding the group’s mean as a control. We show that these techniques, which are used widely in both asset pricing and corporate finance research, typically provide inconsistent coefficients and can lead researchers to incorrect inferences. In contrast, the fixed effects estimator is consistent and should be used instead. We also explain how to estimate the fixed effects model when traditional methods are computationally infeasible. (JEL G12, G2, G3, C01, C13)

Keywords: unobserved heterogeneity, group fixed effects, industry-adjust, bias

* For helpful comments, we thank Michael Anderson, Joshua Angrist, Christian Hansen, Alexander Ljungqvist, Mitchell Petersen, Nagpurmanand R. Prabhala, Michael Roberts, Nick Souleles, Michael R. Wagner, and Jeffrey Wooldridge, as well as the seminar participants at Wharton and MIT. Matthew Denes, Christine Dobridge, Jingling Guan, and Kanis Saengchote provided helpful research assistance. Gormley thanks the Rodney L. White Center for Financial Research Brandywine Global Investment Management Research Fellowship and the Cynthia and Bennett Golub Endowed Faculty Scholar Award for financial support.
† The Wharton School, University of Pennsylvania, 3620 Locust Walk, Suite 2400, Philadelphia, PA, 19104. Phone: (215) 746-0496. Fax: (215) 898-6200. E-mail: tgormley@wharton.upenn.edu
‡ Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208. Phone: (847) 491-8337. Fax: (847) 491-5719. E-mail: dmatsa@kellogg.northwestern.edu
Controlling for unobserved heterogeneity is a fundamental challenge in empirical finance research because most corporate policies—including financing and investment—depend on factors that are unobservable to the econometrician. If these factors are correlated with the variables of interest, then without proper treatment, omitted variables bias infects the estimated parameters and precludes causal inference. In many settings, important sources of unobserved heterogeneity are common within groups of observations, creating a component of the regression errors that is common across observations. For example, unobserved factors—like investment opportunities—are often common across firms in an industry and affect many corporate decisions. Failing to control for such factors can cause serious identification challenges.\footnote{Potential unobserved factors abound: Unobserved differences in local economic environments, management quality, and the cost of capital, to name a few, can also pose difficult identification problems.}

Although the existing literature uses various estimation strategies to control for unobserved group-level heterogeneity, there is little understanding of how these approaches differ and under which circumstances each provides consistent estimates. Our paper examines this question and shows that some commonly used approaches typically lead to inconsistent estimates and can distort inferences.

We focus on two popular estimation strategies. The first, which we refer to as “adjusted-Y” (AdjY), demeanes the dependent variable with respect to the group before estimating the model with ordinary least squares (OLS). A common example is when researchers “industry-adjust” their dependent variable so as to remove common industry factors in a firm-level analysis. A second approach, which we refer to as “average effects” (AvgE), uses the group’s mean as a control in the OLS specification. A common implementation of AvgE uses observations’ state-year mean to control for time-varying differences in local economic environments.

Both AdjY and AvgE are widely used in empirical finance research. Articles published in top finance journals—including the Journal of Finance, Journal of Financial Economics, and Review of Financial Studies—have used both approaches since at least the late 1980s, and they continue to be used today.\footnote{The exact origin of the two estimators in finance is unclear; we suspect they were inspired by the event studies literature, in which stock returns are regressed on market-average returns. AdjY may have been inspired by analyses of market-adjusted returns, and AvgE by estimations of the market model.} Among articles published in these three journals in 2008–2010, we found over 60 articles that
employed at least one of the two techniques. The techniques are used to study a variety of finance topics, including asset pricing, banking, capital structure, corporate boards, governance, executive compensation, and corporate control. Articles using these estimation methods have also been published in the *American Economic Review*, *Journal of Political Economy*, and *Quarterly Journal of Economics*.

Our paper shows that, despite their popularity, the AdjY and AvgE estimators rarely provide consistent estimates of models with unobserved group-level heterogeneity; both estimators can exhibit severe biases. In the presence of such heterogeneity, the AdjY estimator suffers from an omitted variable bias if there is any within-group correlation across observations—either among or across independent variables in the model. Such correlations are likely in the presence of any unobserved group-level variation. The AvgE estimator suffers from a measurement error bias because—even in the absence of any within-group correlations—the group’s sample mean measures the true unobserved heterogeneity with error. For both estimators, the bias can be large and complicated; trying to predict even the sign of the bias is typically impractical because it depends on numerous correlations.

The shortcomings of the AdjY and AvgE estimators stand in stark contrast to the “fixed effects” (FE) estimator—another approach available to control for unobserved group-level heterogeneity. The FE estimator, which instead adds group indicator variables to the OLS estimation, is consistent in the presence of unobserved group-level heterogeneity. When there is only one source of unobserved group-level heterogeneity, the FE estimator is equivalent to demeaning all of the dependent and independent variables with respect to the group and then estimating using OLS.

The differences between the estimators are important because the AdjY and AvgE estimators can lead researchers to make incorrect inferences. We show that AdjY and AvgE estimates can be more biased than OLS and even yield estimates with the opposite sign of the true coefficient. AdjY and AvgE can also be inconsistent even in circumstances in which the original OLS estimates would be consistent.

---

3 Our analysis focuses on settings where the underlying data structure exhibits unobserved group-level heterogeneity; we do not analyze the performance of AdjY and AvgE estimators in other settings. Note, however, that even when the underlying data structure exactly matches the AdjY or AvgE specifications (implying there are peer effects), both estimators are still biased. See Section 2.4 for more details.

4 Our analysis compares the consistency of estimators as the number of observations, \( N \), goes to infinity. Any references in this paper to “bias” do not refer to the finite-sample properties of the estimator but rather the difference between the probability limit of the estimate and the true parameter as \( N \) goes to infinity.
When estimating a few textbook finance models using each of the different techniques to control for unobserved heterogeneity, we find large differences between the AdjY, AvgE, and FE estimates and confirm that AdjY and AvgE can exhibit larger biases than OLS and can yield coefficients of the opposite sign as FE. These differences confirm the presence of unobserved group-level heterogeneity in these settings and of correlations within these commonly used data structures that cause the AdjY and AvgE estimators to be inconsistent and potentially quite misleading in practice.

Based on these findings, we argue that AdjY and AvgE and related estimators should not be used to control for unobserved group-level heterogeneity. The same is also true of related techniques used in the literature to remove unobserved heterogeneity. Any estimation that transforms the dependent variable but not the independent variables will typically yield inconsistent estimates. For example, subtracting the group median or the mean or median of a comparable set of firms from the dependent variable will yield inconsistent estimates by failing to account for correlations across independent variables within these groups. The method of characteristically adjusting stock returns in asset pricing—which subtracts the return of a benchmark portfolio containing stocks with similar characteristics—before comparing these stock returns across subsamples is another transformation that can be problematic, because it does not control for how the benchmark characteristics vary across these subsamples. Regression analyses of conglomerates’ diversification “discount” are also subject to a similar critique. Even a simple comparison of industry-adjusted means before and after events—as is common in analyses of corporate control transactions, stock issues, and other sets of 0/1 events—does not reveal the true causal effect of the events, because the comparison fails to adjust the implicit independent variable, the event indicator, to control for the share of firms in the affected industry.\(^5\)

FE estimators should be used instead of these other approaches. FE estimators are consistent because they are equivalent to transforming both the dependent and independent variables so as to remove the unobserved heterogeneity. For any AdjY or AvgE estimator, there is a corresponding FE estimator that properly accounts for correlations in both the dependent and independent variables. For example,

\(^5\) Put differently, by demeaning the data using an industry mean that includes both affected and unaffected firms, the simple comparison incorrectly removes part of the events’ effect from the estimate.
rather than industry-adjusting a dependent variable or controlling for the dependent variable mean, researchers should instead estimate a model with industry fixed effects. Likewise, rather than correlating benchmark-portfolio-adjusted stock returns with explanatory variables of interest, a researcher should instead estimate a model with fixed effects for each benchmark portfolio. FE estimation still transforms the stock returns using the average returns for the benchmark portfolios but also accounts for correlations between the benchmark portfolios and the explanatory variables of interest.

The FE estimator, however, also has limitations. Although the FE estimator controls for unobserved factors that vary across groups, it is unable to control for unobserved within-group heterogeneities. FE estimation also cannot identify the effect of independent variables that do not vary within groups and is subject to severe attenuation bias in the presence of measurement error. We discuss these limitations and provide guidance on when FE estimation is appropriate. We also describe how researchers can use instrumental variable (IV) techniques within the FE framework to address concerns about attenuation bias and to identify the effects of independent variables that do not vary within groups.

We also address another limitation of FE that has motivated some researchers to use AdjY or AvgE rather than FE—computational difficulties that arise when trying to estimate FE models with multiple types of unobserved heterogeneity. As the size and detail of datasets has increased, researchers are increasingly interested in controlling for multiple sources of unobserved heterogeneity. When there are multiple sources of unobserved group heterogeneity in an unbalanced panel, demeaning the data multiple times is not equivalent to fixed effects. FE estimation of such models requires a large number of indicator variables, which can pose computational problems. The computer memory required to estimate these models can exceed the resources available to most researchers.

We discuss techniques that provide consistent estimates for models with multiple, high-dimensional group effects, while avoiding the computational constraints of a standard FE estimator. One
approach is to interact all values of the multiple group effects to create a large set of fixed effects in one dimension that can be removed by transforming the data. A second approach, which helps to avoid potential attenuation biases and allows the researcher to estimate a larger set of parameters, is to maintain the multidimensional structure but to make estimation feasible by reducing the amount of information that needs to be stored in memory. This can be accomplished by using the properties of sparse matrices and/or by employing iterative algorithms. We discuss the relative advantages of each approach and how these techniques can be implemented easily in the widely used statistical package Stata.\(^7\)

Overall, our paper provides practical guidance on empirical estimation in the presence of unobserved group-level heterogeneity—a pervasive identification challenge in empirical finance research. A small, but impactful, set of recent articles have addressed other challenges researchers face. For example, Bertrand, Duflo, and Mullainathan (2004) and Petersen (2009) recommend methods to account for correlation across residuals in computing standard errors; Almeida, Campello, and Galvao (2010) and Erickson and Whited (2011) compare methods used to account for measurement error in investment regressions; and Fee, Hadlock, and Pierce (2011) evaluate the use of F-tests on indicator variables in managerial style regressions.

The remainder of this paper is organized as follows. Section 1 describes the underlying identification concern of estimating a model with unobserved group-level heterogeneity and describes why \(\text{Adj}Y\) and \(\text{Avg}E\) provide inconsistent estimates. Section 2 contrasts the \(\text{Adj}Y\) and \(\text{Avg}E\) estimators with the FE estimator and discusses why other related estimation techniques, which are commonly used in the literature, will yield inconsistent estimates. Section 3 shows that differences between the estimation techniques can be important in practice. Section 4 discusses limitations of the FE estimator and provides guidance on when its use is appropriate. Section 5 concludes.

1. Estimation Techniques Used to Control for Unobserved Group-Level Heterogeneity

We consider the case in which an independent variable of interest, \(X\), has a causal effect on a dependent variable, \(y\), which also has unobserved group-level heterogeneity that is correlated with \(X\).

---

\(^7\) To help interested researchers, we have also posted code and additional resources on our website to show how common implementations of \(\text{Adj}Y\) and \(\text{Avg}E\) can be transformed into consistent FE estimators.
Assume the data exhibits the following structure:

\[
y_{i,j} = \alpha + \beta X_{i,j} + f_j + \epsilon_{i,j}
\]

var(\epsilon) = \sigma^2_{\epsilon}, \mu_{\epsilon} = 0

var(X) = \sigma^2_X

var(f) = \sigma^2_f

\text{corr}(f_j, \epsilon_{i,j}) = 0

\text{corr}(X_{i,j}, \epsilon_{i,j}) = \text{corr}(X_{i,j}, \epsilon_{-i,j}) = 0

\text{corr}(X_{i,j}, f_j) = \rho_{Xf} \neq 0,

(1)

where the unit of analysis is \(i\) (e.g., firm), these units are organized into larger groups, \(j\) (e.g., industry), and there is some constant unobserved group effect, \(f_j\), for each group \(j\). There are a total of \(N\) observations for the unit of analysis \(i\) with \(N_j\) observations in each group \(j\).

The model in Equation (1) can be easily augmented to reflect more complicated sources of unobserved heterogeneity without affecting our subsequent analysis. Time-varying omitted factors (such as industry shocks that vary over time) can be captured by adding an additional subscript \(t\) to each variable, including the unobserved heterogeneity \(f\). A model with two types of unobserved heterogeneity, such as firm and year group effects in panel data, can be captured by adding a second type of unobserved heterogeneity to Equation (1).

Because there exists nonzero correlation, \(\rho_{Xf}\), between the unobserved group term and the independent variable of interest, failing to account for the unobserved heterogeneity causes an omitted variable problem. We assume the unobserved group term, \(f\), is the only omitted variable, that is, \(f\) and the independent variable of interest, \(X\), do not covary with the residual, \(\epsilon\). For ease of exposition, we further assume that both the group term and independent variable are mean zero (i.e., \(\mu_X = \mu_f = 0\)).

It is well known that using ordinary least squares (OLS) to estimate \(\beta\) — the causal effect of \(X\) on \(y\) — when the data exhibit the structure specified in Equation (1) will yield an inconsistent estimate. This assumption only simplifies the analysis and has no effect on the estimate of \(\beta\) under the different estimation techniques we analyze in Section 2. Assuming nonzero means affects the estimate of the constant, \(\alpha\).

Throughout the paper, we use the standard large-sample approach to determine an estimate’s consistency by taking the number of observations, \(N\), to infinity, while holding group size, \(N_j\), constant. In our later analysis of average effects, we discuss what happens when the number of observations per group, \(N_j\), also goes to infinity. The adjusted-\(Y\) estimates are not affected by group size.
estimates the following specification:

\[ y_{i,j} = \alpha^{\text{OLS}} + \beta^{\text{OLS}} x_{i,j} + u_{i,j}^{\text{OLS}}. \]  

(2)

And, the OLS estimate is

\[ \hat{\beta}^{\text{OLS}} = \beta + \rho_{xy} \left( \frac{\sigma_y}{\sigma_x} \right). \]  

(3)

The OLS estimate of \( \beta \) is inconsistent when the correlation between the group term, \( f \), and the independent variable of interest, \( X \), is nonzero because of an omitted variable problem. By failing to control for the group term, \( \beta^{\text{OLS}} \) will reflect the causal effects of both \( X \) and \( f \) on the dependent variable \( y \).

Because OLS is inconsistent in the presence of unobserved heterogeneity, researchers must rely on other estimation techniques. Two popular approaches are “adjusted-Y” (\( \text{Adj}Y \)), which demeans the dependent variable with respect to the group before estimating the model with OLS, and “average effects” (\( \text{Avg}E \)), which uses the group’s mean of the dependent variable as a control in an OLS specification. In this section, we describe the \( \text{Adj}Y \) and \( \text{Avg}E \) estimates and discuss why they typically lead to inconsistent estimates of the coefficient of interest, \( \beta \). The source of the bias extends to models with more complicated data structures.

1.1 Adjusted-Y estimation

The \( \text{Adj}Y \) estimator attempts to remove the influence of the group term on the dependent variable of interest by demeaning the dependent variable within each group. Adjusted-Y estimation is applied, for example, at the industry level in firm-level panel datasets by subtracting the industry-mean from the dependent variable of interest. When this adjustment is applied at the industry or industry-year level, researchers typically refer to the dependent variable as being “industry-adjusted.”

Although there are a variety of approaches used in practice, a common implementation of \( \text{Adj}Y \) is to demean the dependent variable using the sample group’s mean, excluding the observation at hand.\(^{10}\)

More specifically, the researcher calculates the group mean, \( \bar{y}_{i,j} \), as

\[ \bar{y}_{i,j} = \frac{\sum_{i,j} y_{i,j}}{N_{i,j}}. \]

\(^{10}\) In some cases, the authors do not exclude the observation at hand, and in other cases, the median is used. Both of these alternative approaches yield similarly inconsistent estimates, and proof of this, in the case when the mean of all observations is used, is provided in the Appendix A6.
\[ \bar{y}_{.-i,j} = \frac{1}{N_j-1_{k \in \text{group}}} \sum_{k \in \text{group}} \left( \alpha + \beta X_{k,j} + f_j + e_{k,j} \right) \]  

(4)

and estimates the following model using OLS:

\[ y_{i,j} - \bar{y}_{.-i,j} = \alpha^{Adj} + \beta^{Adj} X_{i,j} + u^{Adj}_{i,j}. \]  

(5)

The adjusted-Y estimation, however, does not generally provide a consistent estimate of \( \beta \) in the presence of unobserved group-level heterogeneity because the estimation suffers from an omitted variable problem. To see this, it is helpful to re-express the sample group mean as

\[ \bar{y}_{.-i,j} = \alpha + f_j + \bar{r}_{i,j}, \]  

(6)

where

\[ \bar{r}_{i,j} = \frac{1}{N_j-1_{k \in \text{group}}} \sum_{k \in \text{group}} \left( \beta X_{k,j} + e_{k,j} \right). \]  

(7)

In the presence of unobserved group-level heterogeneity, as in Equation (1), the dependent variable in \( AdjY \) estimation can be written as

\[ y_{i,j} - \bar{y}_{.-i,j} = \beta X_{i,j} - \bar{r}_{i,j} + e_{i,j}. \]  

(8)

Comparing Equations (5) and (8), we see that the adjusted-Y estimation fails to control for \( \bar{r}_{i,j} \), leading to a biased estimate for \( \beta \) if \( \bar{r}_{i,j} \) is correlated with the independent variable, \( X \). In other words, the covariance between \( X_{i,j} \) and the estimation error, \( u^{Adj}_{i,j} \), is not zero when the correlation between \( \bar{r}_{i,j} \) and \( X_{i,j} \) is nonzero. But, because \( \bar{r}_{i,j} \) includes the mean of the other \( X \) observations in the group, \( \bar{X}_{.-i,j} \), it will be correlated with \( X_{i,j} \) if the \( X \)'s are correlated within groups. Such a correlation is likely when the group effect, \( f_j \), is correlated with \( X_{i,j} \).

Letting \( \rho_{X_{i,j}X_{.-i,j}} \) represent the within-group correlation between \( X \)'s, we can derive the sign and magnitude of the potential bias in the \( AdjY \) estimate of Equation (1).

**Proposition 1.** In the presence of unobserved group-level heterogeneity, as in Equation (1), the adjusted-Y estimator yields an inconsistent estimate for \( \beta \). Specifically,

\[ \hat{\beta}^{Adj} = \beta \left( 1 - \rho_{X_{i,j}X_{.-i,j}} \right). \]
The proof of Proposition 1 is provided in the Appendix.

As with any omitted variable bias, the direction and magnitude of the bias depend on the correlation of $X$ and the omitted variable, which in this case is the within-group correlation between $X$’s. If the $X$’s are positively correlated within groups, such that $\rho_{x_{i,i}} \in (0,1]$, then the adjusted-$Y$ estimate for $\beta$ exhibits an attenuation bias. And, vice versa, if the $X$’s are negatively correlated within groups, the adjusted-$Y$ estimate of $\beta$ is biased away from zero.\footnote{In practice, the bias on $\beta$ is typically attenuating because it is unusual for the $X$’s to be negatively correlated within groups when there is also a group component, $f$, that has nonzero correlation with $X$ (as in Equation (1)).}

In practice, such within-group correlations are commonplace, leading to inconsistent $AdjY$ estimates. For example, consider a standard firm-level capital structure estimation, where leverage is regressed onto multiple independent variables—such as return on assets, bankruptcy risk, and market-to-book ratio. Because firms in the same industry are subject to common demand and technology shocks, their return on assets, bankruptcy risk, and the other regressors exhibit positive within-industry correlation. Within-industry correlation in leverage—the dependent variable—results as well. By industry-adjusting only leverage, the $AdjY$ approach treats the correlation in the dependent variable but fails to address the correlation in the regressors. By failing to control for the correlations in return on assets, bankruptcy risk, and the other regressors, the $AdjY$ (or “industry-adjusting” approach to removing the underlying industry-level heterogeneity) is inconsistent.

The bias in the adjusted-$Y$ estimation is present even with very large groups and even when standard OLS estimates are consistent. Because the $AdjY$ estimator suffers from an omitted variable bias, increasing group size does not lessen the identification problem—the estimation’s error term will always contain the summation of the other within-group $X$’s, which are correlated with the independent variable $X_{i,i}$ when $\rho_{x_{i,i}} \neq 0$. This is also why $AdjY$ is inconsistent even when the correlation between $X$ and $f$ is zero and the OLS estimate is consistent. In this case, $AdjY$ introduces a new omitted variable problem in its attempt to control for a nonexistent omitted variable problem in the original OLS specification.

The bias of the $AdjY$ estimator becomes considerably more complex when there is more than one
independent variable of interest. Suppose the true model is given by the following:

\[ y_{i,j} = \alpha + \beta X_{i,j} + \gamma Z_{i,j} + f_j + \epsilon_{i,j} \]

\[ \text{var}(\epsilon) = \sigma_\epsilon^2, \mu_\epsilon = 0 \]

\[ \text{var}(X) = \sigma_X^2 \]

\[ \text{var}(f) = \sigma_f^2 \]

\[ \text{var}(Z) = \sigma_Z^2 \]

\[ \text{cov}(f, \epsilon_{i,j}) = 0 \]

\[ \text{cov}(X_{i,j}, \epsilon_{i,j}) = \text{cov}(X_{i,j}, \epsilon_{-i,j}) = 0 \]

\[ \text{cov}(Z_{i,j}, \epsilon_{i,j}) = \text{cov}(Z_{i,j}, \epsilon_{-i,j}) = 0 \]

\[ \text{cov}(X_{i,j}, f_j) = \rho_{Xf} \]

\[ \text{cov}(X_{i,j}, Z_{i,j}) = \rho_{XZ} \]

\[ \text{cov}(Z_{i,j}, f_j) = \rho_{Zf}. \]

There is still just one type of unobserved group-level heterogeneity, \( f_j \), but there are two independent variables of interest, \( X \) and \( Z \). As before, assume that the independent variables, \( X_{i,j} \) and \( Z_{i,j} \), and the unobserved group-level effect, \( f_j \), are uncorrelated with the errors. The two independent variables, however, are correlated with each other and with the unobserved group-level heterogeneity, \( f_j \). For ease of exposition, again assume without loss of generality that both the group term and independent variables are mean zero (i.e., \( \mu_X = \mu_Z = \mu_f = 0 \)).

**Proposition 2.** In the presence of unobserved group-level heterogeneity and two independent variables as in Equation (9), the adjusted-\( Y \) estimator yields inconsistent estimates for both \( \beta \) and \( \gamma \). Specifically,

\[ \hat{\beta}_{\text{adj}} = \beta + \beta \left( \frac{\rho_{XZ} \rho_{X_{i,j}Z_{i,j}} - \rho_{X_{i,j}X_{i,j}}}{1 - \rho_{XZ}^2} \right) + \gamma \left( \frac{\sigma_X}{\sigma_Z} \right) \left( \frac{\rho_{XZ} \rho_{X_{i,j}Z_{i,j}} - \rho_{X_{i,j}Z_{i,j}}}{1 - \rho_{XZ}^2} \right) \]

\[ \hat{\gamma}_{\text{adj}} = \gamma + \gamma \left( \frac{\rho_{XZ} \rho_{X_{i,j}Z_{i,j}} - \rho_{X_{i,j}Z_{i,j}}}{1 - \rho_{XZ}^2} \right) + \beta \left( \frac{\sigma_X}{\sigma_Z} \right) \left( \frac{\rho_{XZ} \rho_{X_{i,j}Z_{i,j}} - \rho_{X_{i,j}Z_{i,j}}}{1 - \rho_{XZ}^2} \right), \]

where \( \rho_{X_{i,j}Z_{i,j}} \) and \( \rho_{X_{i,j}X_{i,j}} \) represent the within-group correlations among the independent variables and \( \rho_{X_{i,j}Z_{i,j}} \) represents the within-group correlations across these variables.

The proof of Proposition 2 is provided in the Appendix.
The direction and magnitude of the bias in the AdjY estimator is ambiguous when there are two or more independent variables. The sign of $\hat{\beta}^{AdjY}$ may not even match the sign of the true $\beta$. The expression for $\hat{\beta}^{AdjY}$ is more complicated in Proposition 2 than in Proposition 1 because the omitted variables includes both $X_{-i,j}$ and $Z_{-i,j}$, which may each be correlated with the independent variables, $X_{i,j}$ and $Z_{i,j}$. Even a within-group correlation across independent variables contributes to the inconsistency of AdjY. Such correlations are again commonplace in practice. For example, in the capital structure regression for industry-adjusted leverage discussed earlier, the independent variables bankruptcy risk and return on assets are likely to be positively correlated within industries because of common demand and technology shocks.

1.2 Average effects estimation

Average effects estimation approaches the problem of unobserved heterogeneity a different way. Instead of adjusting the dependent variable, AvgE uses a proxy—the group’s sample mean, $\bar{y}_{-i,j}$—to control for the unobserved variation, $f_j$. A common implementation of AvgE uses observations’ state-year mean to control for time-varying differences in local economic environments. It is again common to exclude the current observation when calculating the group sample mean for each observation.\textsuperscript{12} When there is one independent variable of interest, the average effects approach estimates the following regression:

$$y_{i,j} = \alpha^{AvgE} + \beta^{AvgE}X_{i,j} + \lambda^{AvgE}\bar{y}_{-i,j} + u_{i,j}^{AvgE}. \quad (10)$$

In the presence of unobserved group heterogeneity, as in Equation (1), average effects estimation is inconsistent. The underlying problem is that AvgE suffers from a measurement error bias. As seen in Equation (6), $f_j = \bar{y}_{-i,j} - \alpha - \bar{f}_{-i,j}$, the groups’ sample means, $\bar{y}_{-i,j}$, measure the unobserved variation, $f_j$, with error $-\alpha - \bar{f}_{-i,j}$. Such measurement error does not only bias the coefficient on the mismeasured variable; it also biases coefficients for other variables in the estimated equation.

\textsuperscript{12} As shown in Appendix A6, not excluding the observation at hand also yields inconsistent estimates.
Proposition 3. In the presence of unobserved group-level heterogeneity, as in Equation (1), the average effects estimator yields an inconsistent estimate for $\beta$. Specifically,

$$
\hat{\beta}^{AvgE} = \beta + \frac{\rho_{Xf} \sigma_{f} \left( \frac{\beta^2 \left( 1 - \rho_{X_i,X_{-ij}} \right) + \sigma_{X}^2}{N_j - 1} \right) + \beta \left( \frac{\sigma_{f}^2}{\sigma_{X}^2} \right) (\rho_{Xf} - \rho_{X_{-ij}})}{\left( 1 - \rho_{X_{ij},X_{-ij}} \right) \left( \frac{\sigma_{f}^2}{\sigma_{X}^2} + \beta \rho_{Xf} \right)^2 - \left( \rho_{Xf}^2 - \rho_{X_{ij},X_{-ij}} \right) \left( \beta^2 \left( 1 - \rho_{X_{ij},X_{-ij}} \right) + \sigma_{f}^2 \right) + \frac{\sigma_{f}^2}{\sigma_{X}^2}} \cdot
$$

The proof of Proposition 3 is provided in the Appendix.

The sign and magnitude of the bias for $\hat{\beta}^{AvgE}$ is considerably more complicated than for $\hat{\beta}^{AdjY}$. The bias is complicated because the measurement error, $-\alpha - \bar{\epsilon}_{ij}$, is correlated with both the mismeasured variable, $\bar{Y}_{-ij}$, and with $X_{ij}$ when the within-group correlation among the $X$’s is nonzero. Put another way, the measurement error can be viewed as an omitted variable in the AvgE specification that is not just correlated with $X_{ij}$ (as is the omitted variable in AdjY) but is also correlated with $\bar{Y}_{-ij}$ — the proxy for the unobserved variation $f$.

The $AvgE$ estimator is likely to be inconsistent in practice. Similar to AdjY, the AvgE estimator will be inconsistent when there is correlation among independent variables within groups. For example, in a firm-level AvgE estimation, where a researcher uses the state-year mean to control for time-varying differences in local economic environments, any correlation among the independent variables of firms within a state, such as return on assets, would cause AvgE to be inconsistent. Such correlations are likely to exist because firms located in the same state are subject to similar local demand shocks and investment opportunities. Moreover, even when $\rho_{Xf} = 0$ (and OLS is consistent) and $\rho_{X_{ij},X_{-ij}} = 0$ (and AdjY is consistent), AvgE will still be inconsistent because of the measurement error. Likewise, large group size, $N_j$, does not eliminate the bias when either $\rho_{Xf} \neq 0$ or $\rho_{X_{ij},X_{-ij}} \neq 0$. Although increasing $N_j$ reduces the noise of the measurement error, it does not disappear completely; as seen in Proposition 3, AvgE is still asymptotically inconsistent even when the number of observations per group, $N_j$, goes to infinity.$^{13}$

---

$^{13}$ The variance of the measurement error, $-\alpha - \bar{\epsilon}_{ij}$, is given by $\beta^2 \sigma_{X_{ij}}^2 + \left( \beta^2 \sigma_{X_{ij}}^2 + \sigma_{f}^2 \right) / (N_j - 1)$ and only goes to zero if $\beta = 0$ and $N_j$ also goes to infinity.
1.3 Relative performance of OLS, AdjY, and AvgE estimators

Type 1 error. The potential for incorrect inferences differs across the three approaches. When $\beta = 0$, both the AdjY and AvgE estimation techniques provide less-biased coefficients than does OLS when there is unobserved group-level heterogeneity.

Proposition 4. Assume $\beta = 0$ in Equation (1). The adjusted-Y estimator for $\beta$ is consistent. The average-effects estimator for $\beta$ is inconsistent when $N_j < \infty$ but is closer than the ordinary least squares estimator to the true $\beta$.

The proof of Proposition 4 is provided in the Appendix.

AdjY entirely avoids Type 1 errors because AdjY’s bias is multiplicative. AvgE, however, can incorrectly reject the null hypothesis because the bias in AvgE is also additive, similar to the bias in OLS. When $\beta = 0$, the bias for AvgE, however, is less than that for OLS, and unlike OLS, the bias approaches zero when the number of observations per group, $N_j$, increases. In fact, the AvgE estimator is asymptotically consistent when both $N$ and $N_j$ go to infinity and $\beta = 0$.

Type 2 error. Although both AdjY and AvgE are less biased than OLS when $\beta = 0$, this is not necessarily the case when $\beta \neq 0$. In some cases, AvgE and AdjY will provide a less-biased estimate of $\beta$ than does OLS, and in other cases, they will be more biased and possibly even have the wrong sign. Some examples of this are provided in Table 1. Under certain parameters, OLS will actually be less biased than both AdjY and AvgE. But, as shown in Table 1, there are also cases in which OLS is less biased than AvgE but more biased than AdjY and other cases in which OLS is less biased than AdjY but more biased than AvgE. There are also cases in which the AvgE estimate actually has the opposite (and incorrect) sign.

The correlation between the independent variable of interest, $X$, and the unobserved group-level

---

14 This is not true, however, when there is more than one independent variable. As shown in Proposition 2, the bias in the AdjY estimate is more complicated when there are two independent variables. With two independent variables, as in Equation (9), AdjY only avoids Type 1 errors if both $\beta = 0$ and $\gamma = 0$.  
15 In the case of just one independent variable, as in Table 1, the AdjY estimator cannot incorrectly flip the sign of $\beta$. But, as seen in Proposition 2 (and our later applications in Section 3), this is no longer true when there is more than one independent variable. In this case, both the AdjY and AvgE estimators can return an inconsistent estimate with the incorrect sign while the OLS estimate (also inconsistent) has the correct sign.
heterogeneity, $f$, has a large effect on the relative performance of each estimator. Figure 1 graphs OLS, $AdjY$, and $AvgE$ estimates of Equation (1) as functions of various parameter values when $\beta = 1$. Each figure panel shows the effect of varying a specific parameter in the data structure, while holding the rest constant. When not being varied, the default parameters values are $\rho_{yy} = 0.25$; $\rho_{x,x_{-i}} = 0.5$; $N_j = 10$; $\sigma_y / \sigma_x = 1$; and $\sigma_f / \sigma_x = 1$. Panel A plots the impact of the correlation between the independent variable, $X$, and the unobserved group-level heterogeneity, $f$, on each estimator. $AdjY$ is less biased than the OLS only when the absolute magnitude of the correlation between $X$ and $f$ is large. This is because the $AdjY$ bias is unaffected by $\rho_{xy}$, whereas the magnitude of the OLS bias increases linearly in the absolute value of $\rho_{xy}$. The $AvgE$ bias, in contrast, is nonlinear in $\rho_{xy}$. Under these parameters, $AvgE$ is extremely biased and can even reverse the sign of the coefficient for low values of $\rho_{xy}$, whereas for high values of $\rho_{xy}$, the $AvgE$ is less biased than both OLS and $AdjY$.

The correlation between observations within a group, $\rho_{x_{-i},X_{-i}}$, also affects the relative performance of the $AdjY$ and $AvgE$ estimators. As shown in Panel B, the OLS estimate is unaffected by $\rho_{x_{-i},X_{-i}}$, whereas the sign and magnitude of the bias for both $AvgE$ and $AdjY$ depend on $\rho_{x_{-i},X_{-i}}$. As shown earlier, both the omitted variable in $AdjY$ and the measurement error in $AvgE$ include $\tilde{X}_{-i,j}$ and are correlated with $X_i$ when $\rho_{x_{-i},x_{-i}} \neq 0$, but the omitted variable in OLS is not. Because of this, both $AdjY$ and $AvgE$ are positively biased when $\rho_{x_{-i},x_{-i}}$ is sufficiently low (when $\rho_{x_{-i},x_{-i}} < 0$ in the case of $AdjY$) and are negatively biased when $\rho_{x_{-i},x_{-i}}$ is high.

An increase in the number of observations per group, $N_j$, does not necessarily improve the performance of the various estimators. As shown in Panel C, the OLS and $AdjY$ bias are unaffected by the number of observations per group, because both estimators suffer from an omitted variable bias that is independent of $N_j$. Bias in the $AvgE$ estimate, however, is affected by $N_j$; but under the parameters displayed in Figure 1, an increase in observations per group actually increases the bias. Although increasing $N_j$ reduces the measurement error’s noise, the measurement error is still correlated with $X_i$. On net, the reduction in noise leads the $AvgE$ estimator to asymptote to a biased estimate. The same
occurs when the variance of the error, $\sigma_{\varepsilon}^2$, approaches zero. As shown in Panel D, the AvgE estimate asymptotes to the same biased coefficient as the noise from measurement error declines. For high values of $\sigma_{\varepsilon}^2$, the AvgE estimate asymptotes to a different biased coefficient, which under these parameters, has a bias of the opposite sign as when $\sigma_{\varepsilon}^2$ is low.

The relative variation of the unobserved group-level heterogeneity to that of the independent variable, $\sigma_f / \sigma_x$, also has different implications for the various estimators. As shown in Panel E, the AdjY estimate is unaffected, whereas the magnitude of the bias is increasing with $\sigma_f / \sigma_x$ for both OLS and AvgE—albeit in opposite directions.

In sum, whether AdjY or AvgE provides an improvement over OLS when $\beta \neq 0$ depends on the exact parameter values, but regardless, all three estimates are inconsistent.

2. Fixed Effects Estimation and General Implications

Comparing the AdjY and AvgE estimators to the well-known fixed effects (FE) estimator provides further insight into why AdjY and AvgE are inconsistent. The comparison also highlights why other related estimation techniques, commonly used in the literature, yield inconsistent estimates.

2.1 Fixed effects estimation

Although the OLS, AdjY, and AvgE estimates are all inconsistent in the presence of unobserved heterogeneity, the FE estimator is consistent. FE estimation inserts an indicator variable for each group directly into the OLS equation, thereby allowing the predicted mean of the dependent variable to vary across each group. This estimation, which is also referred to as least squares dummy variable (LSDV) estimation, is consistent because it controls directly for the unobserved group-level heterogeneity $f_j$ in Equation (1). The FE estimate is consistent even when $\rho_{xy} = 0$ and the original OLS estimate is consistent.

Equivalently, the FE estimator can be implemented by transforming the data to remove the unobserved group-level heterogeneity. This is implemented by demeaning all of the variables—both the dependent and independent variables—with respect to the group and then estimating OLS on the transformed data. Specifically, fixed effects (FE) estimates
\[ y_{i,j} - \bar{y}_j = \alpha^{FE} + \beta^{FE} (X_{i,j} - \bar{X}_j) + u_{i,j}, \]  
(11)

where

\[ \bar{y}_j = \frac{1}{N_j} \sum_{k \in \text{group}_j} (\alpha + \beta X_{k,j} + f_j + \epsilon_{k,j}) \]  
(12)

\[ \bar{X}_j = \frac{1}{N_j} \sum_{k \in \text{group}_j} X_{k,j}. \]

Even in the presence of unobserved heterogeneity, as in Equation (1), the dependent variable in the FE specification can be written as

\[ y_{i,j} - \bar{y}_j = \beta (X_{i,j} - \bar{X}_j) + (\epsilon_{i,j} - \bar{\epsilon}_j), \]  
(13)

where

\[ \bar{\epsilon}_j = \frac{1}{N_j} \sum_{k \in \text{group}_j} \epsilon_{k,j}. \]  
(14)

Because the \( X \)'s covariance with \( \epsilon \) is zero, as assumed in Equation (1), their covariance with the last term in Equation (13) is also zero and the FE estimate of \( \beta \) will be consistent.\(^\text{16}\)

Although the estimates are consistent, the standard errors must be appropriately adjusted to account for the degrees of freedom. Typically, the degrees of freedom is adjusted downward (i.e., the estimated standard errors are increased) to account for the number of fixed effects removed in the within transformation. However, when estimating cluster-robust standard errors (which allows for heteroscedasticity and within-group correlations), this adjustment is not required so long as the fixed effects swept away by the within-group transformation are nested within clusters (meaning all the observations for any given group are in the same cluster), as is commonly the case (e.g., firm fixed effects are nested within firm, industry, or state clusters). Statistical software programs that estimate FE specifications make these adjustments automatically.\(^\text{17}\)

See Wooldridge (2010, Chapters 10 and 20), Arellano (1987), and Stock and Watson (2008), for more details.

\(^{16}\) In panel datasets in which the unit of analysis, \( i \), is time (i.e., the “group” is a set of observations over time), first differencing can also be used to remove the unobserved group-level heterogeneity. Although both are consistent in this case, the first-difference and fixed-effects estimators differ in their assumptions about the idiosyncratic errors. See Wooldridge (2010, pg. 321) for details on the relative efficiency of the two estimators.

\(^{17}\) In the current version of Stata, for example, these standard errors are reported by the XTREG command. The AREG command, however, reports larger cluster-robust standard errors that include a full degrees of freedom adjustment. More information on differences between AREG and XTREG cluster-robust standard errors is available on our website.
2.2 Understanding $AdjY$ and $AvgE$ in context of FE

The within-group transformation of the FE estimator highlights the key problem of the $AdjY$ and $AvgE$ estimators: They fail to control for the unobserved group-level heterogeneity across the independent variables. Comparing the FE estimation in Equation (11) with the true underlying structure of the demeaned dependent variable in Equation (13), we see that the FE estimator correctly controls for the independent variable mean, $\bar{X}_j$, and restricts its coefficient to equal $\beta$. The $AdjY$ estimator, however, implicitly assumes this coefficient is zero (see Equation (5)). The $AvgE$ estimator makes the same mistake and also fails to restrict the coefficient on $\bar{y}_{-i,j}$ to equal one (see Equation (10)).

Another way to understand why $AdjY$ and $AvgE$ provide inconsistent estimates is to compare them to a regression of $Y$ onto two independent variables $X$ and $Z$. As is well known, a researcher interested in the effect of $X$ on $Y$ controlling for $Z$ can identify this effect by regressing the residuals from a regression of $Y$ on $Z$ onto the residuals from a regression of $X$ on $Z$. Partialing out the effect of $Z$ from both $X$ and $Y$ before regressing $Y$ on $X$ is equivalent to regressing $Y$ on $X$ controlling for $Z$ (see Greene 2000, pp. 231–33, for more detail). This is the same reason why the within-group transformation implementation of the FE estimator is equivalent to least squares dummy variable estimation. The within-group transformation is simply the result of partialing out the collection of indicator variables (the $Z$ in this case) from both the independent and dependent variables. The $AdjY$ estimation, however, is equivalent to partialing out the effect of $Z$ from only the dependent variable $Y$, which is not equivalent to regressing $Y$ on $X$ controlling for $Z$. The $AvgE$ approach is equivalent to regressing $Y$ on $X$ and the fitted values from a regression of $Y$ on $Z$, which is also not the same as regressing $Y$ on $X$ and $Z$. By failing to transform the independent variable, $X$, both estimators fail to control for the variation in $X$, which is correlated with the unobserved group-level heterogeneity.

2.3 Other related estimation techniques are also inconsistent

Using the same logic, other ways of demeaning the dependent variable in $AdjY$ estimation, such as subtracting a group median or subtracting a value-weighted group-level mean, similarly result in inconsistent estimates. More complicated adjustments to the dependent variable suffer from a similar
problem. In this section, we discuss three related estimation techniques that are inconsistent for a similar reason and explain how to obtain consistent estimates in these cases.

2.3.1 Conglomerate diversification discount

Subtracting the mean or median of a matched control sample (or a variable constructed based on a set of matched controls) presents a similar issue toAdjY estimation. Demeaning the dependent variable using a matched or constructed control but not accounting for the demeaned component of the independent variable creates an omitted variable bias whenever this demeaned component is correlated with a group member’s own independent variable.

Studies examining conglomerates’ diversification “discount” provide an example of this issue. In this literature, it is common to regress conglomerates’ market value of equity onto various independent variables. But first, the conglomerate’s market value is adjusted using the market value of a constructed control of nonconglomerate firms from similar industries. This estimation, however, will suffer an omitted variable bias if the independent variables of the constructed control are correlated with the independent variables of the conglomerate. For example, consider when the independent variable is investment. Assuming investment affects the market value of equity, the investment of the constructed control will affect the adjusted market value of equity, and if the investment of the conglomerate and matched controls are correlated, there will be an omitted variable bias.

To obtain consistent estimates, the researcher can do one of two things. The first is to adjust all the variables in the estimation, both the dependent and independent variables using the constructed controls. The second, and equivalent approach, would be to estimate the model with both sets of observations, the conglomerates and their constructed controls, and include fixed effects for each conglomerate-control pairing.

2.3.2 Characteristically adjusted stock returns

AdjY-type estimators also can be found in many asset pricing articles. One approach commonly used to remove the influence of common risk factors is to adjust firms’ stock returns using the average return of comparable firms. This method was first proposed by Daniel, Grinblatt, Titman, and Wermers (1997) and has since been adopted widely in diverse settings. A researcher first sorts the data into various
benchmark portfolios based on firm-level characteristics, such as size, book-to-market ratios, and momentum, and then “adjusts” the individual stock returns by demeaning them using the average return of other firms in the same benchmark portfolio. From an econometric perspective, there is nothing wrong with using the adjusted return as a measure of stocks’ performance, as proposed by Daniel et al. (1997); it accurately summarizes a portfolio’s performance relative to a benchmark return.

Problems arise, however, if the adjusted return is then correlated with other (unadjusted) stock or firm characteristics. For example, it is common for researchers to calculate adjusted stock returns, sort them into portfolios based on an independent variable that is thought to affect stock returns, and then compare the adjusted returns across the top and bottom portfolios as a test of whether the independent variable affects stock returns. This sort, however, is equivalent to AdjY estimation where the adjusted returns are regressed onto indicators for each independent variable portfolio while excluding the constant term from the regression. Similar to other AdjY estimators, this approach typically provides inconsistent estimates because the specification fails to control for how the average independent variable of other firms in the portfolio influences the adjusted stock return.

As an example, consider analyses of R&D intensity and stock returns. Although R&D and returns are positively correlated, it is possible that differences in firm size confound this relationship. Larger firms are associated with lower stock returns (for reasons presumably unrelated to R&D intensity). If R&D intensity is also correlated with firm size, then the correlation between R&D and returns may be attributable to firm size rather than a causal relation. Using size-matched benchmark portfolios to adjust returns but not adjust R&D does not adequately control for size, because it does not account for average differences in R&D intensity across the benchmark portfolios. The average R&D of firms in a stock’s benchmark portfolio affects its adjusted stock return but this fact is overlooked when one compares adjusted returns across portfolios sorted on firms’ unadjusted R&D intensity.

To control for unobserved risk factors across portfolios appropriately, one needs to adjust the independent variable being analyzed in addition to adjusting returns. One way to accomplish this is to double-sort returns on the benchmark portfolio and the independent variable and then to compare returns across the independent variable within each benchmark portfolio. In the R&D example, this within
benchmark portfolio comparison properly controls for differences in average R&D intensity across firm size. The average difference between the top and bottom independent variable portfolios across all of the benchmark portfolios summarizes that average effect of the independent variable on stock returns while properly controlling for the characteristics used to construct benchmark portfolios. This quantity is equivalent to a fixed effects estimator; the identical estimate can be obtained by regressing returns onto indicators for each independent variable portfolio (excluding the bottom portfolio) and benchmark portfolio fixed effects. In Section 3.4, we provide an example of this alternative fixed effects approach and show that the more conventional characteristically adjusted stock returns estimator is inconsistent.

2.3.3 Difference-in-differences estimators

Using $AdjY$ in the context of a difference-in-differences estimator also yields inconsistent estimates. A properly specified difference-in-differences estimator compares the differences in means between treated and untreated observations across pre- and post-treatment periods. In many applications in finance, however—such as studies of mergers and acquisitions and leveraged buyouts—researchers instead compare the means of an industry-adjusted dependent variable for treated firms across the two periods. This $AdjY$-type comparison does not reveal the true causal effect of the event being analyzed because $AdjY$ removes a weighted-average of treated and untreated firms in the industry, whereas the difference-in-differences estimator removes the mean of just the untreated firms.

Proposition 5. An $AdjY$-type comparison of pre- versus post-treatment periods means for treated firms does not reveal the causal effect of the treatment. Specifically,

$$
\hat{\beta}^{AdjY} = \theta \beta,
$$

where $\theta$ is the average fraction of untreated firms in a treated firm’s group.

The proof of Proposition 5 is provided in the Appendix.

The $AdjY$-type comparison suffers an attenuation bias because it incorrectly removes part of the treatment’s effect on the dependent variable when it demeans the data using a weighted average of treated

---

18 In practice, the double-sort is cumbersome to report when there are many benchmark portfolios (e.g., 5 size x 5 book-to-market x 5 momentum = 125 portfolios) and systematic patterns may be difficult to eyeball. The FE estimator accurately summarizes these patterns and is easy to implement.
and untreated firms. The source of this bias is the same as the more general \( \text{Adj}Y \) bias described previously. In regression form, the \( \text{Adj}Y \) approach to difference-in-differences suffers an omitted variable problem in that it fails to control for the share of firms in each industry that are treated (see the Appendix for more details).

2.4 \( \text{Adj}Y \) and \( \text{AvgE} \) are also inconsistent under other data structures

Even if the true underlying data structure did not exhibit unobserved group-level heterogeneity and exactly matched the \( \text{Adj}Y \) and \( \text{AvgE} \) specifications, both estimators would still be biased. Suppose the underlying data exhibited the following structure:

\[
y_{i,j} = \alpha + \beta X_{i,j} + \bar{y}_{i,j} + \epsilon_{i,j}.
\]

This data structure contains a peer effect such that each observation within a group influences the other observations in the group. However, as shown in Manski (1993) and Leary and Roberts (2010), OLS estimation of the peer effects model does not reveal causal effects because of a reflection problem. Because the dependent variable has a causal effect on the dependent variable of other group members, using the group mean as an independent variable in OLS introduces an endogeneity problem.\(^{19}\)

3. Comparing Approaches in Common Finance Applications

In this section, we examine how important the differences between the various approaches actually are in practice. We estimate standard empirical finance models using each of the various estimators—OLS, FE, \( \text{Adj}Y \), and \( \text{AvgE} \)—and compare the resulting estimates.

3.1 Capital structure and unobserved heterogeneity across firms

We start by estimating a standard capital structure OLS regression:

\[
(D / A)_{i,t} = \alpha + \beta X_{i,t} + \epsilon_{i,t},
\]

where \((D/A)_{i,t}\) is book leverage (debt divided by assets) for firm \(i\) in year \(t\), and \(X_{i,t}\) is a vector of variables thought to affect leverage. We use data for the period of 1950–2010 from Compustat, and we include five independent variables in \(X_{i,t}\): fixed assets / total assets, Ln(sales), return on assets, modified Altman-Z score, and market-to-book ratio. All variables are winsorized at the 1% level, and the standard

\(^{19}\) For the same reason, both OLS and FE would also yield inconsistent estimates of the peer effects model.
errors are adjusted for clustering at the firm level. To account for possible unobserved firm-specific factors in the residual—$e_{it}$—that covary with $X_{it}$, we estimate the model using the four different techniques analyzed here: OLS, AdjY, AvgE, and FE. The estimates are reported in Table 2.

The various estimation techniques lead to very different estimates for $\beta$, confirming the importance of within-firm correlations that cause OLS, AdjY, and AvgE to yield inconsistent estimates. The OLS estimates, reported in Column (1) of Table 2, differ considerably from the FE estimates in Column (4); this suggests the presence of an unobserved firm characteristic—like the firm’s cost of capital or investment opportunities—that is correlated with both the dependent variable (leverage) and independent variables (e.g., ROA). Such unobserved heterogeneities cause the OLS and AvgE estimates to be inconsistent. The AdjY estimates (Column (2)) also differ from the FE estimates. Based on Propositions 1 and 2, the large differences between AdjY and FE imply that there is considerable correlation among independent variables within groups; for example, both fixed assets/total assets and Ln(sales) are serially correlated across time for firms. Such correlations cause both AdjY and AvgE to be inconsistent.

The estimates in Table 2 illustrate that these within-firm correlations can lead to severe biases and incorrect inferences. For example, for the coefficient on the proportion of fixed assets, the AdjY and AvgE estimates are 73% and 58% smaller in magnitude than the FE estimate, respectively; for the z-score, the AdjY and AvgE are smaller in magnitude than the FE estimate by about 40%. A researcher using either AdjY or AvgE might therefore infer that the role of collateral or bankruptcy risk on leverage is considerably smaller than the truth. AdjY and AvgE can even yield an estimate that has the opposite sign as OLS and FE. As reported in Table 2, the OLS estimate for return on assets is -0.015. Relying on the AdjY and AvgE estimates of 0.051 and 0.039, respectively, one might conclude that unobserved firm-level heterogeneity imposes a large downward bias on the OLS coefficient. But, the FE estimate of -0.028—almost twice the OLS estimate—suggests that the bias actually works the other way.

Trying to understand (or predict) the sign of the bias of AdjY and AvgE is typically impractical. As seen in Propositions 2 and 3, the AvgE bias is very complicated, even with just one independent variable, and AdjY is similarly complicated with just two independent variables. When there are even
more independent variables (as in this example), the bias will depend on even more correlations. The complexity of the bias in \( \text{Adj}Y \) and \( \text{Avg}E \) and their dependence on so many correlation parameters make it difficult to guess the direction of the bias or to infer bounds for the true coefficients.

### 3.2 Executive compensation and unobserved heterogeneity across managers

We next estimate the following model for executive compensation:

\[
\text{Ln(Total Compensation)}_{i,j,t} = \alpha + \beta'X_{i,j,t} + \delta_t + \epsilon_{i,j,t},
\]

where \( \text{Ln(Total Compensation)}_{i,j,t} \) is the natural log of total compensation for manager \( i \), at firm \( j \), in year \( t \), \( X_n \) is a vector of variables thought to affect compensation, and \( \delta_t \) is a year fixed effect. Using data for the period of 1992–2010 from Execucomp, Compustat, and CRSP, we estimate the model using each of the different estimators—OLS, \( \text{Adj}Y \), \( \text{Avg}E \), and FE—to account for unobserved heterogeneity across managers. We include ten commonly included independent variables in our regressions: \( \text{Ln(total assets)} \), market-to-book ratio, contemporaneous and lagged stock returns, contemporaneous and lagged return on assets, volatility of daily log stock returns, and separate indicators for being a CEO, chairman, or female. The estimates are reported in Table 3.

Similar to the capital structure regressions, the various techniques lead to very different estimates of \( \beta \). These differences indicate that within-manager correlations cause OLS, \( \text{Adj}Y \), and \( \text{Avg}E \) to provide inconsistent estimates. Different time observations for a given manager are bound to be correlated—indeed the gender indicator is perfectly correlated—and unobserved factors, like managerial ability, are correlated with both the dependent variable (compensation) and independent variables (e.g., profitability). These unobserved factors and within-manager correlations cause the \( \text{Adj}Y \) and \( \text{Avg}E \) estimates to differ considerably from the FE estimates. A researcher using either \( \text{Adj}Y \) or \( \text{Avg}E \) might incorrectly infer that past returns have no effect on total compensation when the opposite may be true. Without estimating the true FE estimates, this bias is hard to predict because it depends on numerous factors. \( \text{Adj}Y \) and \( \text{Avg}E \) estimates are also considerably smaller than both the OLS and FE estimates for \( \text{Ln(Total Assets)} \), market-to-book ratio, stock returns, return on assets, and the CEO indicator.
3.3 Firm value and unobserved, time-varying heterogeneity across industries

We next estimate a model for firm value, as measured using Tobin’s Q:

$$Q_{i,j,t} = \alpha + \beta' X_{i,j,t} + \delta_t + \epsilon_{i,j,t},$$

(18)

where $Q_{i,j,t}$ is Tobin’s Q for firm $i$, in the 4-digit SIC industry $j$, in year $t$, $X_{i,t}$ is a vector of variables thought to affect firm value, and $\delta_t$ is a year fixed effect. Using Compustat data from the period of 1962–2000, we estimate the model using each of the different estimators—OLS, AdjY, AvgE, and FE—to account for unobserved heterogeneity across industry-year combinations. Using AdjY to account for such heterogeneity is often referred to as analyzing “industry-adjusted” data. We include four commonly included independent variables in our regressions: an indicator to capture being incorporated in Delaware, Ln(sales), R&D expenses / assets, and return on assets. All variables are winsorized at the 1% level, and the standard errors are adjusted for clustering at the firm level. The estimates are reported in Table 4.

As before, the various techniques lead to significantly different estimates of $\beta$. These differences indicate the importance of industry-year factors and correlations that cause OLS, AdjY, and AvgE to provide inconsistent estimates. For example, common shocks lead investment opportunities and profitability to be correlated across firms in an industry. When these shocks affect control variables, like return on assets, the within-industry-year correlation biases the AdjY and AvgE estimates. As these shocks also affect omitted factors, such as investment opportunities, that are correlated with both the dependent variable (Tobin’s Q) and independent variables (like R&D expenses), OLS, AdjY, and AvgE are biased further.

It is apparent that analyzing industry-adjusted data rather than using industry-year FE can distort inference. For the coefficient on Delaware incorporation, the AdjY and AvgE estimates are statistically insignificant and considerably smaller than the statistically significant OLS and FE estimates. Because firms in the same industry experience similar demand and cost shocks, positive within industry-year correlations among the independent variables (like return on assets and sales) bias these estimates. A researcher relying on AdjY or AvgE would conclude that incorporation in Delaware is uncorrelated with firm value after accounting for industry trends, when the opposite is true. The AdjY and AvgE estimates for Ln(sales) and R&D expenses also differ considerably from the FE estimates.
3.4 Stock returns and unobserved, time-varying heterogeneity across industries and size

In a final example, we examine the relationship between firms’ research and development expenses and stock returns. Using CRSP and Compustat data from the period of 1962–2010, we sort stock returns based on firms’ ratio of R&D to market value of equity in the preceding year. In Table 5, Panel A, we report the unadjusted, market-weighted average two-year holding period return and standard error for firms in each R&D quintile and for firms for which information on R&D is not available. Similar to the literature, we find that R&D is positively related to stock returns; firms in the upper quintile have an average two-year stock return that is 750 basis points larger than firms with R&D in the lowest quintile.

But, firms with high or low R&D may also differ in other ways that could affect stock returns. To illustrate a common approach to controlling for such factors, we construct $48 \times 5 = 240$ benchmark portfolios at the end of June in each year based on firms’ 48 Fama-French industry classification and size quintile. We then demean each stock return using the corresponding return on the benchmark portfolio and sort these characteristically adjusted returns based on R&D. The market-weighted adjusted returns, reported in Table 5, Panel B, also show a positive relationship between R&D and stock returns. Specifically, firms in the highest quintile of R&D have an average adjusted return that is 530 basis points higher than the average adjusted return of firms in the lowest quintile of R&D.

However, the comparison of such adjusted returns across quintiles is an $AdjY$ estimation. This can be seen by estimating the following model:

\[
r_{i,j,s,t} = \beta' R&D_{i,j,s,t} + \varepsilon_{i,j,s,t} ,
\]

where $r_{i,j,s,t}$ is the stock return for firm $i$ in 48 Fama-French industry $j$, size quintile $s$, year $t$, and $R&D_{i,j,s,t}$ is a vector of indicators for firms’ R&D quintile in year $t$. We also include an indicator for firms missing information on R&D. We exclude an indicator for the lowest R&D quintile such that the resulting estimates represent the average differences in stock returns between each quintile and the lowest. The regressions are weighted by firms’ market value of equity, and the standard errors are adjusted for clustering at the firm level. In this example, the concern motivating $AdjY$ estimation is that unobserved heterogeneity across the 240 industry-size benchmark portfolios is correlated with R&D and therefore
distorting inference. An alternative approach is to use the benchmark portfolios as groups in \( \text{AvgE} \) or FE estimation. Market-weighted OLS, \( \text{AdjY} \), \( \text{AvgE} \), and FE estimates of Equation (19) are reported in Columns (1)–(4) of Table 5, Panel C.

The OLS and \( \text{AdjY} \) estimates reported in Panel C correspond exactly to the quintile returns reported in Panels A and B, respectively. The OLS estimates in Column (1) exactly equal the difference in average returns between each portfolio return and the lowest quintile portfolio reported in Panel A; for example, the difference in average returns between the highest R&D quintile portfolio and the lowest R&D quintile portfolio is 800 – 50 = 750 basis points. Similarly, the \( \text{AdjY} \) estimates in Column (2) exactly equal the difference in characteristically adjusted returns reported in Panel B. Comparisons of characteristically adjusted returns across quintiles are the same as running an \( \text{AdjY} \) estimation.

Like in the other examples, the various techniques lead to different estimates of \( \beta \). In this case, correlations between R&D, industry, and firm size cause OLS, \( \text{AdjY} \), and \( \text{AvgE} \) to provide inconsistent estimates for the relation between R&D and returns. A researcher relying on characteristic adjustments (\( \text{AdjY} \)) or \( \text{AvgE} \) would conclude that being in the highest quintile of R&D increases stock returns over firms in the lowest quintile by 530 basis points (Columns (2)–(3)). The FE estimation, however, shows that the effect of R&D quintiles is actually almost twice as large at 940 basis points (Column (4)).

From this example, it is apparent that analyzing characteristically adjusted data rather than controlling for benchmark-portfolio fixed effects can distort inference. The problem is that returns are characteristically adjusted but the sorting variable—R&D in this case—is not. This approach can provide inconsistent estimates because the specification fails to control for how the average sorting variable of other firms in the benchmark portfolio influences the adjusted stock return. Whether this failure matters in practice depends on the specific application. The degree of the problem depends on the exact benchmarking method, the sample, and ultimately whether the sorting variable is correlated within the benchmark portfolios. A researcher can diagnose the problem by measuring these correlations directly, or more simply, avoid this concern entirely by estimating a FE model instead, as is done here.

Overall, the large differences between \( \text{AdjY} \), \( \text{AvgE} \), and FE in each of the four different examples presented here suggest that researchers’ choice of which approach to use is important in practice. Within-
group correlations, which as we have shown are common in many settings, will cause both $AdjY$ and $AvgE$ estimates to be inconsistent in the presence of unobserved heterogeneity and can lead to severe biases and incorrect inferences.

4. Limitations of Fixed Effects Estimation and How to Overcome Them

Based on our analysis above, it is clear that $AdjY$ and $AvgE$ should *not* be used to control for unobserved heterogeneity. These ad hoc methods are inconsistent and can lead to estimates that differ substantially from the true underlying parameters. The FE estimator provides consistent estimates in the presence of unobserved group-level heterogeneity and should be used instead.

The FE estimator, however, has limitations. One limitation (which is also a limitation of $AdjY$ and $AvgE$) is that FE is unable to control for unobserved heterogeneity *within* groups. For example, industry-level fixed effects do not control for unobserved geographic differences in demand or costs across local markets within an industry. Likewise, firm-level fixed effects do not control for unobserved firm factors that vary over time, such as the accumulation of organization capital. Given the potential limitations of FE, when should researchers estimate a model with fixed effects? There are four criteria that indicate estimating a FE model:

1. There likely exists unobserved group-level heterogeneity.
2. The heterogeneity is potentially correlated with a variable of interest.
3. There exists within-group variation in the variable of interest.
4. The variable of interest is well measured.

Conditions (1) and (2) represent existence criteria for unobserved group-level heterogeneity that could cause an omitted variable bias. If the factor is observable, then the researcher should control for potential omitted variable directly using OLS. If the heterogeneity is unobservable and uncorrelated with the independent variable of interest, then OLS without the group indicator variables is also consistent. Conditions (1) and (2) provide the motivation for using a FE estimator, which unlike the $AdjY$ and $AvgE$ estimators, successfully controls for the unobserved group-level heterogeneity.

The remaining two conditions refer to potential limitations of FE and its ability to obtain correct inferences. Condition (3) refers to the inability of FE estimation to directly identify the effect of
independent variables that do not vary within groups; Condition (4) refers to the potentially severe attenuation biases that can occur in FE estimations when independent variables are not well measured. In this section, we discuss these two limitations and how they can be overcome. We also describe how to overcome a third limitation of the FE estimator—computational difficulties that arise when estimating FE models with multiple sources of unobserved heterogeneity.

4.1 Independent variables that do not vary within groups

If there is no within-group variation in the variable of interest, it is not possible to disentangle the group component from that of the independent variable in FE estimation, because the fixed effects are perfectly collinear with variables that do not vary within groups. For example, the effect of a manager’s gender on his or her total compensation cannot be estimated while also controlling for manager fixed effects because gender is time-invariant and perfectly collinear with the manager fixed effects. Likewise, the effect of internal governance on firm outcomes cannot be estimated while also controlling for firm fixed effects when internal governance mechanisms do not vary over time.

In some cases, violations of condition (3) can be addressed using IV strategies within the FE estimation framework. As shown in Hausman and Taylor (1981), the coefficients on variables that are constant within groups can be recovered using a two-step procedure when other covariates vary within groups and are uncorrelated with the unobserved heterogeneity. In the first step, the FE estimation is used to estimate the coefficients for variables that vary within groups. In the second step, group-average residuals from the first step are regressed on the covariates that do not vary within groups using as instruments covariates that vary within groups and are not correlated with the unobserved heterogeneity. This estimation can be implemented in Stata using the XTHTAYLOR command.

4.2 Attenuation bias from noisy independent variables

FE estimators are subject to a potentially severe attenuation bias. Although FE estimators successfully remove unobserved heterogeneity that would otherwise bias estimates, they also remove meaningful variation from the variable of interest. If some of the variable’s within-group variation is noise, then the share of variation being analyzed that is noise can rise sharply in FE estimation. This
increase in noise occurs when there is classical measurement error of an independent variable or when the independent variable is measured perfectly but the within-group variation does not capture the relevant variation. For example, a firm-level fixed effects estimation may exhibit a lot of noise if the dependent variable responds to sustained but not transitory changes in the independent variable, because the meaningful, sustained variation is largely removed by the firm fixed effects (McKinnish 2008). This noise can cause severe attenuation bias and lead a researcher to infer that the variable of interest does not affect on the dependent variable when the opposite is true.

For example, consider a researcher who is interested in estimating the effect of a regulatory change on banks’ lending to low-income households. The researcher obtains data from a credit reporting bureau and controls for unobserved shocks to credit demand by including Zip Code-by-quarter fixed effects. If there is measurement error in the credit reporting data (e.g., some loan originations and payoffs are not recorded in a timely manner), then the importance of this error could be magnified by the FE strategy. Even when the FE estimation indicates that the regulatory change had little or no effect on low-income borrowing, the true effect may be large but absorbed by the fixed effects, resulting in severe attenuation bias. This magnification of a measurement error bias is particularly acute when much of the variation in the variable of interest occurs across, rather than within, groups.20

In the presence of such bias, standard techniques to address measurement error can be applied. The typical solution is to identify an instrument for within-group variation in the independent variable and implement standard IV methods. The instrument is most often a variable outside the data structure that meets the standard relevance and exclusion criteria. In panel data, another option is to recover the true $\beta$ from the biased coefficients obtained from OLS estimation on different transformations of the data, such as first differences or within-group transformation. If the measurement errors are serially uncorrelated, the different data transformations affect the estimate in predictable ways that can be used to analytically solve for the bias and recover the true parameter (see Griliches and Hausman 1986; McKinnish 2008).21

---


21 Even if the measurement errors are serially correlated, it may be possible to identify the true parameters if the researcher is willing to make additional assumptions. See Griliches and Hausman (1986) for more details.
4.3 Estimating models with multiple high-dimensional fixed effects

Researchers often face computational hurdles when trying to estimate FE models on large datasets with multiple sources of unobserved heterogeneity. When there are two or more high-dimension group effects (i.e., each set of groups includes many distinct occurrences), many indicator variables must be included in the estimation. The large number of parameters to be estimated can lead to computer memory requirements that exceed the available computer resources. Overcoming these difficulties, in fact, provides a potential motivation for using AdjY or AvgE rather than FE.22

When there is just one type of unobserved group effect, the fixed effects estimation is always computationally feasible if the original OLS estimation is feasible. This is because the data can be transformed by demeaning with respect to the group component and then estimating OLS on the transformed data. In a specification with \( K \) independent variables of interest and \( G \) groups, the within transformation reduces the number of parameters to estimate from \( G + K \) to \( K \)—the same number as OLS. Statistical programs typically use this transformation to estimate models with fixed effects.

However, when there are two unobserved group effects of dimension \( G_1 \) and \( G_2 \), there is generally no such transformation to reduce the number of parameters. If the data is a balanced panel, meaning there is a consistent set of observations for each subgroup, the data can be transformed by demeaning the dependent variable and each independent variable with respect to each group sequentially and then estimating the regression using OLS (see Greene 2000, pp. 564–65, for more detail). This transformation reduces the number of parameters from \( G_1 + G_2 + K \) to \( K \). But if the panel is unbalanced, which is far more common in practice, such a transformation typically does not exist.23

In practice, unbalanced models with two unobserved group effects, such as firm and year group effects, are estimated using a partial transformation. The researcher inserts indicator variables for the

---

22 However, of the published articles we found using either AdjY or AvgE, applying an FE estimator appears feasible in the vast majority of articles. Of the more than sixty articles published in the Journal of Finance, Journal of Financial Economics, and Review of Financial Studies between 2008 and 2010 that use either AdjY or AvgE estimation, there were only three articles in which computational problems arise. In all but two of the other articles, there was only one fixed effect, which is easily handled using the within transformation, and in the other two cases, the amount of required memory to estimate the FE model would not pose a computational problem.

23 Wansbeek and Kapteyn (1989) proposed a transformation for cases in which the data are unbalanced but patterned, such as with individual and time group effects, but this transformation is not typically used in practice because the number of time periods is usually not large enough to cause computational hurdles. See Baltagi (1995, pp. 159–60) for further discussion.
smaller group directly in the specification and performs a within-group transformation in the higher dimension. The indicator variables must be inserted before the transformation is applied. This combination of indicator variables and data transformation yields consistent estimates for the $K$ parameters of interest, and eases computational difficulties by reducing the number of estimated parameters from $G_1 + G_2 + K$ to $G_2 + K$, where $G_2 < G_1$. A common application of this approach inserts year indicators in firm-year panel estimation and then subtracts the firm’s mean from each variable (including the indicators).

However, this partial transformation is only computational feasible if one of the two groups is of low enough dimension that creating a design matrix with $G_2 + K$ parameters is feasible. There are many examples for which this is not the case. For instance, suppose the researcher is working with a large panel of firms and wants to control for both unobserved, time-invariant firm characteristics and time-varying industry shocks. In this case, both group effects are of high-dimension, particularly when industrial classification is at a level similar to 3- or 4-digit SIC and there are many years of data.

Finance researchers increasingly argue for the need to control for multiple, high-dimensional fixed effects. For example, in the analysis of executive compensation, there may be concern about unobserved heterogeneity across managers (such as skill, risk aversion, and personality) and unobserved heterogeneity across firms (such as firm culture and organization capital) that might also correlate with variables of interest (such as size, profitability, and CEO age; Graham, Li, and Qui 2012; Coles and Li 2011a). The inclusion of manager fixed effects, in addition to firm fixed effects, can be used to remove this unobserved heterogeneity and allow the researcher to remove potential omitted variable biases introduced by such unobserved heterogeneity at the manager- or firm-level. The inclusion of firm- and manager-level fixed effects may also be important in other contexts (Coles and Li 2011b). Likewise, researchers that use firm-level data are increasingly concerned about time-varying heterogeneity across industries, such as industry-level shocks to demand (Matsa 2010). Such heterogeneity may warrant the addition of industry-by-time fixed effects to a specification that already includes firm fixed effects. And, in identification strategies that exploit local changes in regulation over time, there is a concern that these changes in regulation may coincide with other time-varying local characteristics (Cetorelli and Strahan...
If not all firms are affected by the regulation equally, then including location-by-time fixed effects can control for such characteristics.

Fortunately, there exist computational techniques that provide consistent estimates for models with more than one high-dimensional group effect without storing large matrices in memory. One approach is to interact the multiple types of unobserved heterogeneity into a one-dimensional set of fixed effects, which is accounted for using a within transformation. A second approach is to use memory-saving procedures and/or iterative algorithms. Each of these approaches has benefits and limitations.

4.3.1 Interacted fixed effects

Memory requirements can be reduced by interacting multiple fixed effects into a one-dimensional set of fixed effects and then applying a within-group transformation. For example, when there are unobserved factors at the firm and industry-year levels, the two types of fixed effects (firm and industry-year) can be replaced with one set of firm-industry-year fixed effects and then controlled for using a within-firm-year transformation. This within transformation removes both firm and industry-year factors, reducing the number of estimated parameters to $K$ (the same as OLS), avoiding the computational problems of trying to estimate a model with separate fixed effects for firm and industry-year.

However, the interacted fixed effects approach has potentially serious limitations. The first limitation is that interacted fixed effects remove more heterogeneity than necessary and may, as a result, severely limit the types of parameters that can be estimated. In the above example with interacted firm-year fixed effects, only variables that vary within firm-years can be identified, whereas the original specification provides estimates for variables that vary within firms and within industry-years. In fact, given that most finance datasets only contain one observation per firm-year, using interacted fixed effects in this case is infeasible because there is no within variation left after including firm-year fixed effects.\(^{24}\) Even when some variation remains after transforming the data, the estimates may suffer serious attenuation bias from even modest measurement error.

---

\(^{24}\) Firm-manager matched data is one example in which interacted fixed effects are feasible. An interacted manager-firm fixed effect can be used to remove unobserved heterogeneity both within managers and within firms. Such estimations are possible when managers work at multiple firms over their careers. See Coles and Li (2011a,b) and Graham, Li, and Qui (2012) for examples of this approach.
A second limitation of interacted fixed effects estimation is that it does not allow the researcher to recover the uninteracted fixed effects. When the researcher seeks to analyze the distribution, correlation, and importance of the fixed effects for specific groups (such as manager/worker fixed effects, as in Abowd, Kramarz, and Margolis 1999; Abowd, Creecy, and Kramarz 2002; Coles and Li 2011a; Graham, Li, and Qui 2012), other estimation techniques that allow the researcher to recover the estimates on the separate fixed effects are required.

4.3.2 Memory-saving procedure and iterative algorithms

Other approaches maintain the unobserved heterogeneity’s multidimensional structure but modify the algorithm used to estimate the original FE model. For example, memory requirements can be reduced by recognizing that the design matrix is a sparse matrix—a matrix with many zeros—because of all the fixed effects. Sparse matrices can be compressed into smaller matrices that require less memory by only storing the nonzero values. By compressing the matrix of indicator variables, the required memory can be significantly reduced (Abowd, Creecy, and Kramarz 2002). Cornelissen (2008) provides a detailed description of how this can be implemented, and provides a program, FELSDVREG, that uses this estimation technique in Stata for a model with two high-dimensional fixed effects.

Memory requirements can be reduced further using an iterative algorithm to estimate the FE model rather than constructing and inverting matrices. The iterative algorithm avoids needing to store any indicator variables in memory, eliminating memory limitations that bind even after implementing memory-saving procedures. Guimarães and Portugal (2010) show how iterative methods can be used to estimate coefficients and standard errors in models with two high-dimensional fixed effects; the program REG2HDFE implements this algorithm in Stata. However, when computer memory is not a binding constraint, a drawback of the iterative algorithm is that it can take longer to compute when a large number of iterations are required.

These alternative estimation techniques for models with more than one high-dimension group effect are computationally feasible and relatively quick. To illustrate this, we attempted to estimate a capital structure regression using the standard FE approach and the alternative techniques. Specifically,

---

25 See Smyth (1996) and Guimarães and Portugal (2010) for more details on how this algorithm is performed, particularly in the presence of two high-dimension fixed effects.
we estimated the following regression:

\[ \frac{D}{A}_{ijt} = \alpha + \beta X_{ijt} + f_i + \delta_{jt} + u_{ijt}, \]

where \( X_{ijt} \) is a vector of five time-varying controls (i.e., \( K = 5 \)) for firm \( i \), in 4-digit SIC industry \( j \), and year \( t \). The five independent variables we use are fixed assets / total assets, \( \ln(\text{sales}) \), return on assets, modified Altman-Z score, and market-to-book ratio. The regression includes two high-dimension group fixed effects: firm fixed effects, \( f_i \), and industry-by-year fixed effects, \( \delta_{jt} \). We use a sample from Compustat that covers the period of 1970–2008, which contains \( N = 318,808 \) firm-year observations, \( G_1 = 28,365 \) unique firms, 450 4-digit SIC industries, and \( G_2 = 16,769 \) unique industry-years.

Estimating this model using a combination of a within transformation and indicator variables is not computationally feasible. Even after applying the within-transformation to remove the firm fixed effects, 16,769 industry-year dummies remain. If the memory required to store each element of the design matrix is eight bytes, the total memory required to store the design matrix of the transformed data is \( N \times (G_2 + K) \times 8 \), or 39.84 gigabytes. Even if we reduce the industry controls to the 3-digit SIC level, there are 10,527 industry-year dummies, requiring 26.86 gigabytes in memory, which still exceeds the memory available to most researchers (including the authors).

The memory-saving and iteration techniques we discuss previously avoid these limitations. (Using interacted fixed effects in this case is not possible because there does not exist any within firm-year variation.) Using the FELSDVREG algorithm to estimate the FE model reduces the required memory to \( (G_2 + K)^2 \times 8 \) = 2.09 gigabytes. Although computational times will obviously vary based on computing resources, FELSDVREG was able to successfully generate the OLS estimates with standard errors clustered at the firm level on the authors’ computer in about 7 hours and 50 minutes. REG2HDFE was considerably faster. Using the iterative approach, REG2HDFE successfully returned estimates with clustered standard errors in less than five minutes.\textsuperscript{26}

\textsuperscript{26} Both of these programs report cluster-robust errors that reduce the degrees of freedom by the number of fixed effects swept away in the within-group transformation (i.e., \( G_1-1 \)), which as noted earlier, may be inappropriate in some applications. To recover the smaller cluster-robust standard errors that do not make this adjustment, one multiplies the reported standard errors by the square root of \( (N - G_1 - G_2 - K + 1) / (N - G_2 - K) \). See our website for further details on how to implement this adjustment and the methods discussed in this section.
5. Conclusion

As empirical researchers, it is well understood that we must overcome the identification challenge posed by unobserved heterogeneity if we hope to infer causal statements from data we analyze. It is less clear, however, how researchers can best account for such heterogeneity. In practice, there are numerous methodologies that are used widely to account for unobserved shocks affecting groups of observations. One approach subtracts the mean of the group from the dependent variable; another controls for correlations with the mean instead. A third approach—fixed-effects estimation—includes indicator variables for each group as additional controls, or equivalently, demeans all of the model variables within groups (not just the dependent variable). This paper explores how these various approaches differ and under which circumstances each provides consistent estimates of the parameters of interest.

We find that only the fixed effects approach yields consistent estimates in the presence of unobserved group-level heterogeneity while the other widely used approaches yield inconsistent estimates. Demeaning the dependent variable suffers an omitted variable problem by failing to account for the within-group mean of the independent variables. Controlling directly for the within-group mean suffers from measurement error bias, as the mean is only a noisy measure of the unobserved factors.

The difference between the various approaches is important in practice. Besides providing inconsistent estimates, the alternative approaches can lead to severe biases and incorrect inferences. They have the potential to generate estimates whose magnitude and sign do not match the true parameter. Estimating textbook finance models using each of the approaches, we confirm that these biases are severe in practice. Compared to the FE estimates, the alternative approaches often result in very different estimates and occasionally return statistically significant estimates of the opposing sign.

Although we show fixed effects estimation to be the best way to account for unobserved common factors, the estimation strategy has limitations. It can neither control for unobserved factors that vary within groups nor identify the effect of independent variables that are constant within groups. The estimates are also subject to especially severe attenuation biases when the independent variable is measured with error. Given these limitations, we provide guidance on when fixed effects estimation is appropriate and how measurement error problems can be overcome in the estimation.
We also address how researchers can overcome computational difficulties when estimating fixed effects models that include multiple sources of unobserved heterogeneity across many groups. This type of estimation is increasingly common as researchers work with large datasets and attempt to account for more sources of unobserved heterogeneity in their analyses. We describe new techniques to estimate the FE model when standard approaches are computationally infeasible. These new techniques are likely to be of increasing importance and of practical use to empirical researchers.

References


Figure 1. Comparative statics on bias of OLS, Adj Y, and Avg E when $\beta = 1$

This figure presents analytical solutions for the OLS, Adj Y, and Avg E estimates of the $\beta$ in Equation (1),
\[ y_{ij} = \alpha + \beta X_{ij} + f_i + \varepsilon_{ij}, \]
as a function of underlying parameters ($\rho_{xf}$, $\rho_{X(i)X(-i)}$, $\sigma_f/\sigma_X$, and $\sigma_f/\sigma_X$) and the number of observations in a group ($N_j$) when $\beta = 1$. Each figure panel shows the effect of varying a specific parameter in the data structure while holding the rest constant. The vertical axis in each graph displays the estimated $\beta$. When not varying along the horizontal axis, the default parameter values are: $\sigma_f/\sigma_X = 1$; $\sigma_f/\sigma_X = 1$; $\rho_{X(i)X(-i)} = 0.5$; $\rho_{xf} = 0.5$; $N_j = 10$.

Panel A. Correlation between $X$ and $f$

Panel B. Correlation between $X_i$ and $X_{-i}$
Figure 1 continued

Panel C. Number of observations per group

Panel D. Relative variation in $\varepsilon$ and $X$

Panel E. Relative variation in $f$ and $X$
Table 1. Analytical examples of bias with OLS, Adj Y, and Avg E

This table reports analytical solutions for the OLS, Adj Y, and Avg E estimators of Equation (1),
\[ y_{ij} = \alpha + \beta x_{ij} + f_j + \varepsilon_{ij}, \]
under different assumptions about the underlying data structure. In all cases, the true coefficient (\( \beta \)) equals 1, the number of observations per group (\( N_j \)) equals 10, and \( \sigma_{\varepsilon}/\sigma \) equals 1. The assumed \( \text{corr}(X_{ij}, X_{-ij}), \text{corr}(X_{ij}, f_j), \) and \( \sigma_f/\sigma_X \) are given in Columns (1)-(3), and the OLS, Adj Y, and Avg E estimates are given in Columns (7)-(9).

<table>
<thead>
<tr>
<th>Underlying Data Structure</th>
<th>Estimates of ( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta_{OL} )</td>
</tr>
<tr>
<td>( \rho_{x_{i,j}, x_{-i,j}} )</td>
<td>( \rho_{x_f} )</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>-0.05</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

**OLS is less biased than Adj Y, Avg E**

| 0.5 | -0.7 | 1 | 1 | 10 | 1 | 0.30 | 0.50 | -3.88 |
| 0.25| -0.5 | 1 | 1 | 10 | 1 | 0.50 | 0.75 | -0.56 |
| 0.1 | -0.4 | 1 | 1 | 10 | 1 | 0.60 | 0.90 | 0.18  |

**Adj Y least biased, Avg E most biased**

| 0.5 | 0.35 | 1 | 1 | 10 | 1 | 1.35 | 0.50 | 0.81  |
| 0.75| 0.5  | 1 | 1 | 10 | 1 | 1.50 | 0.25 | 0.68  |
| 0.25| 0.5  | 0.45| 1 | 10 | 1 | 1.22 | 0.75 | 1.05  |

**Avg E least biased, Adj Y most biased**

| 0.25| 0.5  | 1 | 1 | 10 | 1 | 1.50 | 0.75 | 1.05  |
| 0.5 | 0.6  | 1 | 1 | 10 | 1 | 1.60 | 0.50 | 0.98  |
| -0.05| 0.2 | 1 | 1 | 10 | 1 | 1.20 | 1.05 | 1.09  |
Table 2. Firm heterogeneity and capital structure

This table reports coefficients from firm-panel regressions of book leverage on fixed assets / total assets, Ln(sales), return on assets, modified Altman Z-score, and market-to-book ratio using different methodologies to account for unobserved group-level heterogeneity across firms. The data are from Compustat for the period of 1950–2010 and exclude financial and regulated industries. All variables were winsorized at the 1% tails. Column (1) reports the OLS estimates; Column (2) reports the AdjY estimates; Column (3) reports the Avg estimates; and Column (4) reports the FE estimates. Standard errors, clustered at the firm level, are reported in parentheses. *** significant at 1% level.

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>AdjY</th>
<th>Avg E</th>
<th>FE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Assets/ Total Assets</strong></td>
<td>0.270***</td>
<td>0.066***</td>
<td>0.103***</td>
<td>0.248***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.014)</td>
</tr>
<tr>
<td><strong>Ln(sales)</strong></td>
<td>0.011***</td>
<td>0.011***</td>
<td>0.011***</td>
<td>0.017***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>0.000</td>
<td>0.000</td>
<td>(0.001)</td>
</tr>
<tr>
<td><strong>Return on Assets</strong></td>
<td>-0.015***</td>
<td>0.051***</td>
<td>0.039***</td>
<td>-0.028***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td><strong>Z-score</strong></td>
<td>-0.017***</td>
<td>-0.010***</td>
<td>-0.011***</td>
<td>-0.017***</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.001)</td>
</tr>
<tr>
<td><strong>Market-to-book Ratio</strong></td>
<td>-0.006***</td>
<td>-0.004***</td>
<td>-0.004***</td>
<td>-0.003***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>166,974</td>
<td>166,974</td>
<td>166,974</td>
<td>166,974</td>
</tr>
<tr>
<td><strong>R²</strong></td>
<td>0.29</td>
<td>0.14</td>
<td>0.56</td>
<td>0.66</td>
</tr>
</tbody>
</table>
### Table 3. Manager heterogeneity and executive compensation

This table reports coefficients from manager-level panel regressions of Ln(Total Compensation) on year fixed effects and control variables using different methodologies to account for unobserved group-level heterogeneity across managers. The data are from Execucomp, Compustat, and CRSP for the period of 1992–2010. Column (1) reports the OLS estimates; Column (2) reports the Adj $Y$ estimates; Column (3) reports the Avg $E$ estimates; and Column (4) reports the FE estimates. The included control variables are Log(Total Assets); lagged market-to-book ratio; contemporary and lagged stock returns, which is calculated using a 12 month holding period; contemporary and lagged return on assets, which is calculated using income before extraordinary items / total assets; volatility of annualized daily log stock returns; an indicator for whether the CEO is chairman of the board; an indicator for whether the manager is the CEO; and an indicator for being female. Standard errors, clustered at the firm level, are reported in parentheses. * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

<table>
<thead>
<tr>
<th>Dependent Variable = Ln(Total Compensation)</th>
<th>OLS</th>
<th>Adj $Y$</th>
<th>Avg $E$</th>
<th>FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Total Assets)</td>
<td>0.341***</td>
<td>0.021***</td>
<td>0.066***</td>
<td>0.240***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.002)</td>
<td>(0.005)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Market-to-Book Ratio $[t - 1]$</td>
<td>0.093***</td>
<td>0.009***</td>
<td>0.021***</td>
<td>0.032***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Stock Return $[t]$</td>
<td>0.120***</td>
<td>0.039***</td>
<td>0.050***</td>
<td>0.091***</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Stock Return $[t - 1]$</td>
<td>0.041***</td>
<td>0.053***</td>
<td>0.051***</td>
<td>0.076***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Return on Assets $[t]$</td>
<td>0.287***</td>
<td>0.092***</td>
<td>0.120***</td>
<td>0.268***</td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td>(0.032)</td>
<td>(0.033)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Return on Assets $[t - 1]$</td>
<td>0.135**</td>
<td>0.004</td>
<td>0.023</td>
<td>0.193***</td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td>(0.034)</td>
<td>(0.036)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Volatility of Daily Ln(Returns)</td>
<td>0.132***</td>
<td>0.002</td>
<td>0.021</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.012)</td>
<td>(0.013)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>CEO = Chairman Indicator</td>
<td>0.225***</td>
<td>0.045***</td>
<td>0.071***</td>
<td>0.028*</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>CEO Indicator</td>
<td>0.723***</td>
<td>0.141***</td>
<td>0.224***</td>
<td>0.431***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.012)</td>
<td>(0.015)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Female Indicator</td>
<td>-0.115***</td>
<td>-0.023***</td>
<td>-0.036***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>96,719</td>
<td>96,719</td>
<td>96,719</td>
<td>96,719</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.47</td>
<td>0.08</td>
<td>0.76</td>
<td>0.82</td>
</tr>
</tbody>
</table>
Table 4. Industry heterogeneity and Tobin's Q

This table reports coefficients from firm-level regressions of Tobin's Q on an indicator for being incorporated in Delaware, Ln(sales), R&D expenses / assets, return on assets, and year fixed effects using different methodologies to account for unobserved group-level heterogeneity across industry-years at the 4-digit SIC industry level. The data are from Compustat for the period of 1962–2000 and exclude financial and regulated industries. All variables were winsorized at the 1% tails. Column (1) reports the OLS estimates; Column (2) reports the Adj Y estimates; Column (3) reports the Avg estimates; and Column (4) reports the FE estimates. Standard errors, clustered at the firm level, are reported in parentheses. ** significant at 5% level; *** significant at 1% level.

<table>
<thead>
<tr>
<th>Dependent Variable = Tobin's Q</th>
<th>OLS (1)</th>
<th>Adj Y (2)</th>
<th>Avg E (3)</th>
<th>FE (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delaware Incorporation</td>
<td>0.100***</td>
<td>0.019</td>
<td>0.040</td>
<td>0.086**</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.032)</td>
<td>(0.032)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Ln(sales)</td>
<td>-0.125***</td>
<td>-0.054***</td>
<td>-0.072***</td>
<td>-0.131***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>R&amp;D Expenses / Assets</td>
<td>6.724***</td>
<td>3.022***</td>
<td>3.968***</td>
<td>5.541***</td>
</tr>
<tr>
<td></td>
<td>(0.260)</td>
<td>(0.242)</td>
<td>(0.256)</td>
<td>(0.318)</td>
</tr>
<tr>
<td>Return on Assets</td>
<td>-0.559***</td>
<td>-0.526***</td>
<td>-0.535***</td>
<td>-0.436***</td>
</tr>
<tr>
<td></td>
<td>(0.108)</td>
<td>(0.095)</td>
<td>(0.097)</td>
<td>(0.117)</td>
</tr>
<tr>
<td>Observations</td>
<td>55,792</td>
<td>55,792</td>
<td>55,792</td>
<td>55,792</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.22</td>
<td>0.08</td>
<td>0.34</td>
<td>0.37</td>
</tr>
</tbody>
</table>
Table 5. Industry and size heterogeneity and stock returns

This table compares two-year holding period stock returns as a function of firms' ratio of research and development expenses to market value of equity (R&D) using data from CRSP and Compustat for the period of 1962–2010. Panel A reports the average two-year holding period stock return and standard deviation by R&D quintile, where firms with missing R&D are reported separately. Panel B does a similar comparison but first subtracts the average return on a benchmark portfolio for each stock using 48×5=240 benchmark portfolios (i.e., groups) that capture the 48 Fama-French industries and a firm's size quintile. Panel C reports estimates from a regression of stock returns onto indicators for firms' R&D quintile using different methodologies to account for unobserved group-level heterogeneity across the 240 benchmark portfolios. Column (1) reports the OLS estimates; Column (2) reports the Adj Y estimates; Column (3) reports the Avg E estimates; and Column (4) reports the FE estimates. All regressions are weighted by firms' market value of equity. Standard errors, clustered at the firm level, are reported in parentheses. ** significant at 5% level; *** significant at 1% level.

<table>
<thead>
<tr>
<th>Panel A. Yearly stock returns sorted by R&amp;D Quintile (i.e., OLS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing</td>
</tr>
<tr>
<td>0.080***</td>
</tr>
<tr>
<td>(0.006)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Characteristically adjusted returns by R&amp;D Quintile (i.e., Adj Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing</td>
</tr>
<tr>
<td>-0.012***</td>
</tr>
<tr>
<td>(0.003)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C. Regression analysis of R&amp;D (Differences relative to Quintile 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable = Yearly Stock Return</strong></td>
</tr>
<tr>
<td>OLS</td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>Missing</td>
</tr>
<tr>
<td>(0.015)</td>
</tr>
<tr>
<td>Q2</td>
</tr>
<tr>
<td>(0.016)</td>
</tr>
<tr>
<td>Q3</td>
</tr>
<tr>
<td>(0.017)</td>
</tr>
<tr>
<td>Q4</td>
</tr>
<tr>
<td>(0.021)</td>
</tr>
<tr>
<td>Q5</td>
</tr>
<tr>
<td>(0.021)</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
</tbody>
</table>
Appendix

A1. Proof of Proposition 1

Given $\mu_X = \mu_f = 0$ and Equation (1), when the number of observations, $N$, goes to infinity and the number of observations per group, $N_j$, remains constant, the $AdjY$ estimates $\hat{\alpha}^{AdjY}$ and $\hat{\beta}^{AdjY}$ are given by

$$\lim_{N \to \infty} \hat{b} = \lim_{N \to \infty} \left( \frac{1}{N} X'X \right)^{-1} \frac{1}{N} X'y,$$

where

$$b = \begin{bmatrix} \hat{\alpha}^{AdjY} \\ \hat{\beta}^{AdjY} \end{bmatrix},$$

$$X' = \begin{bmatrix} 1 & \cdots & 1 \\ X_{1,j} & \cdots & X_{N,j} \end{bmatrix},$$

$$y = \begin{bmatrix} y_{1,j} - \bar{y}_{i,j} \\ \vdots \\ y_{N,j} - \bar{y}_{N,j} \end{bmatrix} = \begin{bmatrix} \beta X_{1,j} - \bar{y}_{i,j} + \epsilon_{i,j} \\ \vdots \\ \beta X_{N,j} - \bar{y}_{N,j} + \epsilon_{N,j} \end{bmatrix},$$

and

$$\bar{y}_{i,j} = \alpha + f_j + \bar{f}_{i,j},$$

$$\bar{f}_{i,j} = \frac{1}{N_j - 1} \sum_{k \neq i} \left( \beta X_{k,j} + \epsilon_{k,j} \right).$$

It can then be shown that

$$\lim_{N \to \infty} \left( \frac{1}{N} X'X \right)^{-1} = \lim_{N \to \infty} \left( \frac{1}{N} X_{1,j} \cdots X_{N,j} \right)^{-1} \left( \frac{1}{N} X_{1,j} \cdots X_{N,j} \right)^{-1} = \lim_{N \to \infty} \left( \frac{1}{N} \sum_{j} X_{i,j}^2 \right)^{-1} = \left( \frac{1}{N} \sum_{j} X_{i,j}^2 \right)^{-1} \left( \frac{1}{N} \sum_{j} X_{i,j} \right)^{-1} = \left( \frac{1}{N} \sum_{j} X_{i,j} \right)^{-1} \left( \frac{1}{N} \sum_{j} X_{i,j} \right)^{-1} = \left( \frac{1}{N} \sum_{j} X_{i,j} \right)^{-1}. $$

Because the covariance between $X_{i,j}$ and $\bar{f}_{i,j}$, $\sigma_{Xf}$, is given by

$$A-1$$
\[ \text{cov}(X_{i,j}, \bar{X}_{i,j}) = \text{cov} \left( \frac{1}{N_j - 1} \sum_{k \in \text{group}_j} (\beta X_{k,j} + \epsilon_{k,j}) \right) \]
\[ = \text{cov} \left( \frac{1}{N_j - 1} \sum_{k \in \text{group}_j} X_{k,j} \right) + \text{cov} \left( \frac{1}{N_j - 1} \sum_{k \in \text{group}_j} \epsilon_{k,j} \right), \]  \text{(A5)}
\[ = \beta \sigma_{X_i X_{i,j}} \]

It can also be shown that
\[
\text{plim} \frac{1}{N} X'y = \left[ \begin{array}{c}
1 \\
\vdots \\
1
\end{array} \right] \left[ \begin{array}{c}
X_{i,j} \\
\vdots \\
X_{N,j}
\end{array} \right] \left( \begin{array}{c}
\beta X_{i,j} - \bar{X}_{i,j} + \epsilon_{i,j} \\
\vdots \\
\beta X_{N,j} - \bar{X}_{N,j} + \epsilon_{N,j}
\end{array} \right) N
\]
\[ = \text{plim} \frac{1}{N} \sum_{i} \left( \begin{array}{c}
\frac{\sum \beta X_{i,j} - \bar{X}_{i,j} + \epsilon_{i,j}}{N} \\
\frac{\sum X_{i,j} (\beta X_{i,j} - \bar{X}_{i,j} + \epsilon_{i,j})}{N}
\end{array} \right) = \left( \begin{array}{c}
0 \\
\beta \sigma_{X}^2 - \sigma_{y}^2
\end{array} \right) = \left( \begin{array}{c}
0 \\
\beta \sigma_{X}^2 - \sigma_{x,x_{i,j}}^2
\end{array} \right). \]  \text{(A6)}

Therefore,
\[
\hat{b} = \left[ \hat{\alpha}_{4(8)} \hat{\beta}_{4(8)} \right] = \left( \begin{array}{cc}
1 & 0 \\
0 & 1
\end{array} \right) \left( \begin{array}{cc}
\sigma_{x}^2 - \sigma_{x,x_{i,j}}^2 & 0 \\
0 & \beta - \frac{\sigma_{X_i X_{i,j}}}{\sigma_{X}^2}
\end{array} \right) = \left( \begin{array}{cc}
0 & 0 \\
\beta - \frac{\rho_{X_i X_{i,j}}}{\sigma_{X}^2} & 0
\end{array} \right). \]  \text{(A7)}

QED

**A2. Proof of Proposition 2**

Given \( \mu_x = \mu_z = \mu_y = 0 \) and Equation (13), when \( N \) goes to infinity and the number of observations per group, \( N_j \), remains constant, the \( \hat{Adj}Y \) estimates \( \hat{\alpha}_{4(8)} \), \( \hat{\beta}_{4(8)} \), and \( \hat{\gamma}_{4(8)} \) are given by

\[ \text{plim} \frac{1}{N} X'y = \text{plim} \left( \frac{1}{N} X'X \right)^t \frac{1}{N} X'y, \]  \text{(A8)}

where
\[ b = \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{bmatrix} \]

\[ X' = \begin{pmatrix} 1 & \ldots & 1 \\ X_{1,i} & \ldots & X_{N,i} \\ Z_{1,i} & \ldots & Z_{N,i} \end{pmatrix} \]

\[ y = \begin{pmatrix} y_{1,i} - \bar{y}_{i,j} \\ \vdots \\ y_{N,i} - \bar{y}_{N,j} \end{pmatrix} = \begin{pmatrix} \beta X_{1,i} - \bar{r}_{i,j} + \epsilon_{i,j} \\ \vdots \\ \beta X_{N,i} - \bar{r}_{N,j} + \epsilon_{N,j} \end{pmatrix} \]

and

\[ \bar{y}_{i,j} = \alpha + f_j + \bar{r}_{i,j} \]

\[ \bar{r}_{i,j} = \frac{1}{N_j - 1} \sum_{k \neq i} \left( \beta X_{k,j} + \gamma Z_{k,j} + \epsilon_{k,j} \right). \]

It can then be shown that

\[ \text{plim} \left( \frac{1}{N} X'X \right)^{-1} = \text{plim} \left( \frac{1}{N} \begin{pmatrix} X_{1,i} & \ldots & X_{N,i} \\ Z_{1,i} & \ldots & Z_{N,i} \end{pmatrix} \right)^{-1} \]

\[ = \text{plim} \left( \frac{1}{N} \begin{pmatrix} \sum X_{1,i} \\ \sum Z_{1,i} \end{pmatrix} \right)^{-1} \]

\[ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sigma_x^2 & \sigma_{xz} \\ 0 & \sigma_{xz} & \sigma_z^2 \end{pmatrix} \]

\[ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sigma_x^2} + \frac{\sigma_{xz}^2}{\sigma_x^2} - \frac{\sigma_{xz}^2}{\sigma_x^2} & \frac{\sigma_{xz}}{\sigma_x^2} \\ 0 & \frac{\sigma_{xz}}{\sigma_x^2} & \frac{\sigma_z^2}{\sigma_x^2} \end{pmatrix} \]

where \( \sigma_{xz} \) is the covariance between \( X_{1,i} \) and \( Z_{1,i} \).

Because the covariance between \( X_{1,i} \) and \( \bar{r}_{1,i} \), \( \sigma_{xy} \), is given by

\[ (A11) \]
\[
\text{cov}(X_{i,j}, \tilde{r}_{i,j}) = \text{cov}\left(\frac{X_{i,j}}{N_j - 1 \sum_{k \in \text{group}_j} X_{k,i}}, \sum_{k \in \text{group}_j} \left(\beta X_{k,i} + \gamma Z_{k,i} + \epsilon_{k,i}\right)\right) \\
= \text{cov}\left(\frac{X_{i,j}}{N_j - 1 \sum_{k \in \text{group}_j} X_{k,i}} + \frac{\gamma}{N_j - 1 \sum_{k \in \text{group}_j} Z_{k,j}} \right), \\
= \beta \sigma_{x_{i,j}, x_{i,j}} + \gamma \sigma_{x_{i,j}, z_{i,j}}
\]

where \(\sigma_{x_{i,j}, x_{i,j}}\) is the covariance between \(X_{i,j}\) and \(Z_{i,j}\), and the covariance between \(Z_{i,j}\) and \(\tilde{r}_{i,j}\), \(\sigma_{Zr}\), is given by

\[
\text{cov}(Z_{i,j}, \tilde{r}_{i,j}) = \text{cov}\left(\frac{Z_{i,j}}{N_j - 1 \sum_{k \in \text{group}_j} X_{k,i}}, \sum_{k \in \text{group}_j} \left(\beta X_{k,i} + \gamma Z_{k,i} + \epsilon_{k,i}\right)\right) \\
= \text{cov}\left(\frac{Z_{i,j}}{N_j - 1 \sum_{k \in \text{group}_j} X_{k,i}} + \frac{\gamma}{N_j - 1 \sum_{k \in \text{group}_j} Z_{k,j}} \right), \\
= \beta \sigma_{z_{i,j}, x_{i,j}} + \gamma \sigma_{z_{i,j}, z_{i,j}}
\]

where \(\sigma_{z_{i,j}, z_{i,j}}\) is the covariance between \(Z_{i,j}\) and \(Z_{i,j}\) and \(\sigma_{z_{i,j}, x_{i,j}}\) is the covariance between \(Z_{i,j}\) and \(X_{i,j}\), it can be shown that

\[
\text{plim}_{N \to \infty} \frac{1}{N} \mathbf{X'y} = \frac{1}{N} \left[
\begin{array}{c}
X_{i,j} \\
\vdots \\
X_{N,j} \\
Z_{i,j} \\
Z_{N,j}
\end{array}
\right]
\left[
\begin{array}{c}
\beta X_{i,j} + \gamma Z_{i,j} - \tilde{r}_{i,j} + \epsilon_{i,j} \\
\vdots \\
\beta X_{N,j} + \gamma Z_{N,j} - \tilde{r}_{N,j} + \epsilon_{N,j}
\end{array}
\right]
\]

\[
= \frac{1}{N} \left[
\begin{array}{c}
\sum_{j} \beta X_{i,j} + \gamma Z_{i,j} - \tilde{r}_{i,j} + \epsilon_{i,j} \\
\sum_{j} X_{i,j} \left(\beta X_{i,j} + \gamma Z_{i,j} - \tilde{r}_{i,j} + \epsilon_{i,j}\right) \\
\sum_{j} Z_{i,j} \left(\beta X_{i,j} + \gamma Z_{i,j} - \tilde{r}_{i,j} + \epsilon_{i,j}\right)
\end{array}
\right]
\]

\[
= \left[
\begin{array}{c}
\beta \sigma_X^2 + \gamma \sigma_{XZ} - \sigma_X \\
\beta \sigma_{XZ} + \gamma \sigma_Z^2 - \sigma_Z
\end{array}
\right]
\]

\[
= \left[
\begin{array}{c}
\beta \left[\sigma_X^2 - \sigma_{X_iX_{-i}}\right] + \gamma \left[\sigma_{XZ} - \sigma_{X_iZ_{-i}}\right] \\
\gamma \left[\sigma_Z^2 - \sigma_{Z_iZ_{-i}}\right] + \beta \left[\sigma_{XZ} - \sigma_{Z_iX_{-i}}\right]
\end{array}
\right]
\]

\[(A14)\]
Therefore,

\[ b = \begin{bmatrix} \hat{\alpha}^{Avg\hat{Y}} \\ \hat{\beta}^{Avg\hat{Y}} \\ \hat{\gamma}^{Avg\hat{Y}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{\sigma^2}{\sigma^2_x - \sigma^2_z} & -\frac{\sigma_{xz}}{\sigma^2_x - \sigma^2_z} & 0 \\ 0 & -\frac{\sigma_{xz}}{\sigma^2_x - \sigma^2_z} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \beta \left[ \frac{\sigma^2_z - \sigma_{X_i,i,j}}{\sigma^2_x - \sigma^2_z} \right] + \gamma \left[ \sigma_{xz} - \sigma_{X_i,i,j} \right] \\ \gamma \left[ \frac{\sigma^2_z - \sigma_{Z_i,i,j}}{\sigma^2_x - \sigma^2_z} \right] + \beta \left[ \sigma_{xz} - \sigma_{X_i,i,j} \right] \end{bmatrix} \]

\[ = \begin{bmatrix} \beta + \beta \left( \frac{\rho_{xz} \rho_{Z_i,i,j} - \rho_{X_i,i,j}}{1 - \rho^2_{xz}} \right) + \gamma \left( \frac{\sigma_x \left( \rho_{xz} \rho_{Z_i,i,j} - \rho_{X_i,i,j} \right)}{\sigma^2_x - \sigma^2_z} \right) \\ \gamma + \gamma \left( \frac{\rho_{xz} \rho_{Z_i,i,j} - \rho_{Z_i,i,j}}{1 - \rho^2_{xz}} \right) + \beta \left( \frac{\sigma_x \left( \rho_{xz} \rho_{X_i,i,j} - \rho_{Z_i,i,j} \right)}{\sigma^2_x - \sigma^2_z} \right) \end{bmatrix} \]

(A15)

where \( \rho_{Z_i,i,j} \) is the correlation between \( Z_{i,j} \) and \( Z_{i,j} \), \( \rho_{Z_i,i,j} \) is the correlation between \( Z_{i,j} \) and \( X_{i,j} \), and \( \rho_{X_i,i,j} \) is the correlation between \( X_{i,j} \) and \( Z_{i,j} \). QED

A3. Proof of Proposition 3

Given \( \mu_X = \mu_f = 0 \) and Equation (1), when \( N \) goes to infinity and the number of observations per group, \( N_j \), remains constant, the AvgE estimates \( \hat{\alpha}^{AvgE} \), \( \hat{\beta}^{AvgE} \), and \( \hat{\gamma}^{AvgE} \) are given by

\[ \text{plim}_{N \to \infty} b = \text{plim}_{N \to \infty} \left( \frac{1}{N} X'X \right)^{-1} \frac{1}{N} X'y \]

(A16)

where

\[ b = \begin{bmatrix} \hat{\alpha}^{AvgE} \\ \hat{\beta}^{AvgE} \\ \hat{\gamma}^{AvgE} \end{bmatrix} \]

\[ X' = \begin{pmatrix} 1 & \cdots & 1 \\ X_{i,j} & \cdots & X_{i,j} \\ y_{-i,j} & \cdots & y_{-i,j} \end{pmatrix}, \]

\[ y = \begin{pmatrix} y_{i,j} \\ \vdots \\ y_{N,j} \end{pmatrix} \]

(A17)

and
\( \bar{\alpha}_{i,j} = \alpha + f_j + \bar{\alpha}_{i,j} \)
\( \bar{\alpha}_{i,j} = \frac{1}{N - 1} \sum_{k=\text{group}} (\beta X_{i,j} + \epsilon_{k,i,j}) \).  

(A18)

It can then be shown that

\[
\begin{pmatrix}
1 & \cdots & 1 \\
X_{i,j} & \cdots & X_{N,j} \\
\alpha + f_j + \bar{\alpha}_{i,j} & \cdots & \alpha + f_j + \bar{\alpha}_{N,j}
\end{pmatrix}^{-1} \times
\begin{pmatrix}
1 & \cdots & 1 \\
X_{i,j} & \cdots & X_{N,j} \\
\alpha + f_j + \bar{\alpha}_{i,j} & \cdots & \alpha + f_j + \bar{\alpha}_{N,j}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
1 & 0 & \alpha \\
0 & \sigma_x^2 & \sigma_{xy} \\
\alpha & \sigma_{xy} & \alpha^2 + \sigma_y^2 + 2\sigma_{xy} + \sigma_r^2
\end{pmatrix}^{-1}
\]

where \( \sigma_{xy} \) is the covariance between \( X_{i,j} \) and \( \bar{\alpha}_{i,j} \), \( \sigma_y \) is the covariance between \( f_j \) and \( \bar{\alpha}_{i,j} \), and \( \sigma_r^2 \) is the variance of \( \bar{\alpha}_{i,j} \).

And, it can be shown that

\[
\begin{pmatrix}
1 & \cdots & 1 \\
X_{i,j} & \cdots & X_{N,j} \\
\alpha + f_j + \bar{\alpha}_{i,j} & \cdots & \alpha + f_j + \bar{\alpha}_{N,j}
\end{pmatrix}^{-1} \times
\begin{pmatrix}
1 & \cdots & 1 \\
X_{i,j} & \cdots & X_{N,j} \\
\alpha + f_j + \bar{\alpha}_{i,j} & \cdots & \alpha + f_j + \bar{\alpha}_{N,j}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\sum_i \alpha + \beta X_{i,j} + f_j + \epsilon_{i,j} \\
\sum_i X_{i,j} (\alpha + \beta X_{i,j} + f_j + \epsilon_{i,j}) \\
\sum_i (\alpha + f_j + \bar{\alpha}_{i,j}) (\alpha + \beta X_{i,j} + f_j + \epsilon_{i,j})
\end{pmatrix}^{-1}
\]

\[
= \begin{pmatrix}
\alpha \\
\beta \sigma_x^2 + \sigma_{xy} \\
\alpha^2 + \beta \sigma_{xy} + \sigma_{y}^2
\end{pmatrix}
\]

(A20)

Therefore,
\[
\begin{bmatrix}
\hat{\alpha} \\
\hat{\beta} \\
\hat{\lambda}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & \alpha \\
0 & \sigma_x^2 & \sigma_{xy} + \sigma_{x}\sigma_r \\
\alpha & \sigma_{xy} + \sigma_{x}\sigma_r & \alpha^2 + \sigma_r^2 + 2\sigma_{x} + \sigma_r^2
\end{bmatrix}^{-1}
\begin{bmatrix}
\alpha \\
\beta \sigma_x^2 + \sigma_{xy} \\
\alpha^2 + \beta \sigma_{xy} + \sigma_r^2 + \beta \sigma_{x} + \sigma_r^2
\end{bmatrix}
\]

\[
\begin{bmatrix}
\alpha - \alpha \\
\sigma_{xy} \left( \frac{\sigma_{xy}^2 + \sigma_{x}^2}{\sigma_x^2} \right) - \left( \sigma_{xy}^2 + \sigma_{x} \sigma_r \right) \left[ \sigma_{xy} + \sigma_r \right] - \left[ \frac{\sigma_{xy}^2 + \sigma_{x} \sigma_r}{\sigma_x^2} \right] \\
\sigma_{xy} \left( \frac{\sigma_{xy}^2 + \sigma_{x}^2}{\sigma_x^2} \right) - \left( \sigma_{xy}^2 + \sigma_{x} \sigma_r \right) \left[ \sigma_{xy} + \sigma_r \right] - \left[ \frac{\sigma_{xy}^2 + \sigma_{x} \sigma_r}{\sigma_x^2} \right] \\
\sigma_{xy} \left( \frac{\sigma_{xy}^2 + \sigma_{x}^2}{\sigma_x^2} \right) - \left( \sigma_{xy}^2 + \sigma_{x} \sigma_r \right) \left[ \sigma_{xy} + \sigma_r \right] - \left[ \frac{\sigma_{xy}^2 + \sigma_{x} \sigma_r}{\sigma_x^2} \right]
\end{bmatrix}
\]  

This expression can be simplified by recognizing that

\[
\sigma_{xy} = \text{cov} \left\{ f, \frac{1}{N_j - 1} \sum_{k \in \text{group } j} (\beta X_{k,j} + \epsilon_{k,j}) \right\}
\]

\[
= \text{cov} \left\{ f, \frac{\beta}{N_j - 1} \sum_{k \in \text{group } j} X_{k,j} \right\} + \text{cov} \left\{ f, \frac{1}{N_j - 1} \sum_{k \in \text{group } j} \epsilon_{k,j} \right\}
\]

\[
= \beta \sigma_{xy}
\]

and

\[
\sigma_{x}^2 = \text{var} \left\{ \frac{1}{N_j - 1} \sum_{k \in \text{group } j} (\beta X_{k,j} + \epsilon_{k,j}) \right\}
\]

\[
= \text{var} \left\{ \frac{\beta}{N_j - 1} \sum_{k \in \text{group } j} X_{k,j} + \frac{1}{N_j - 1} \sum_{k \in \text{group } j} \epsilon_{k,j} \right\}
\]

\[
= \left( \frac{\beta}{N_j - 1} \right)^2 \text{var} \left\{ \sum_{k \in \text{group } j} X_{k,j} \right\} + \left( \frac{1}{N_j - 1} \right)^2 \text{var} \left\{ \sum_{k \in \text{group } j} \epsilon_{k,j} \right\}
\]  

(A21)
\[
\frac{1}{N_j - 1} \left[ \frac{1}{N_j - 1} \right] \sigma_i^2 + \frac{1}{N_j - 1} \left[ \frac{1}{N_j - 1} \right] (N_j - 1) \sigma_i^2
\]

\[
= \beta \left( N_j - 1 \right) \sigma_i^2 + \left( N_j - 1 \right) (N_j - 2) \text{cov}(X_y, X_j) + \frac{\sigma_i^2}{N_j - 1}
\]

\[
= \frac{\beta^2 \sigma_i^2}{N_j - 1} + \frac{\beta \left( N_j - 2 \right) \sigma_{X,X_{i,j}}}{N_j - 1} + \frac{\sigma_i^2}{N_j - 1}
\]

\[
= \frac{\beta^2 \sigma_i^2}{N_j - 1} + \left( \frac{N_j - 2}{N_j - 1} \beta^2 \sigma_{X,X_{i,j}} \right) + \frac{\sigma_i^2}{N_j - 1}
\]

\[
= \beta^2 \sigma_{X,X_{i,j}} + \frac{\beta^2 \left( \sigma_i^2 - \sigma_{X,X_{i,j}} \right) + \sigma_i^2}{N_j - 1}
\]

With these results and some algebra, it can be shown that

\[
\frac{1}{\alpha} \begin{bmatrix}
\frac{\sigma_l}{\sigma_x} \\
\frac{\beta \rho_{X,Y} \left( 1 - \rho_{X,Y} \right)}{\rho_{X,Y} \sigma_x^2}
\end{bmatrix}
= \beta \left[ \frac{1}{N_j - 1} \right] \begin{bmatrix}
\left( 1 - \rho_{X,Y} \right) \left( \frac{\sigma_l}{\sigma_x} \right)^2 - \left( \rho_{X,Y} - \rho_{X,X_{i,j}} \right) \left( \beta^2 \left( 1 - \rho_{X,X_{i,j}} \right) + \frac{\sigma_l^2}{\sigma_x^2} \right) + \frac{\beta^2 \left( 1 - \rho_{X,Y} \right) + \sigma_l^2}{N_j - 1}
\end{bmatrix}
\]

\[
\frac{1}{\alpha} \begin{bmatrix}
\tilde{\alpha}_{\text{avg}} \\
\tilde{\beta}_{\text{avg}} \\
\tilde{\lambda}_{\text{avg}}
\end{bmatrix}
= \beta \left[ \frac{1}{N_j - 1} \right] \begin{bmatrix}
\left( 1 - \rho_{X,Y} \right) \left( \frac{\sigma_l}{\sigma_x} \right)^2 - \left( \rho_{X,Y} - \rho_{X,X_{i,j}} \right) \left( \beta^2 \left( 1 - \rho_{X,X_{i,j}} \right) + \frac{\sigma_l^2}{\sigma_x^2} \right) + \frac{\beta^2 \left( 1 - \rho_{X,Y} \right) + \sigma_l^2}{N_j - 1}
\end{bmatrix}
\]

\[
\left( 1 - \rho_{X,Y} \right) \left( \frac{\sigma_l}{\sigma_x} \right)^2 - \left( \rho_{X,Y} - \rho_{X,X_{i,j}} \right) \left( \beta^2 \left( 1 - \rho_{X,X_{i,j}} \right) + \frac{\sigma_l^2}{\sigma_x^2} \right) + \frac{\beta^2 \left( 1 - \rho_{X,Y} \right) + \sigma_l^2}{N_j - 1}
\]

\[
\text{QED}
\]
A4. Proof of Proposition 4

Given Equation (1) and Propositions 1 and 4, the estimates of $\beta$ when $\beta = 0$ are given by

$$
\begin{vmatrix}
\hat{\beta}_{OLS} \\
\hat{\beta}_{Adj} \\
\hat{\beta}_{AvgE}
\end{vmatrix}
=
\frac{
\rho_{yx} \left( \frac{\sigma_f}{\sigma_x} \right) \begin{pmatrix}
\frac{\sigma^2_f}{\sigma^2_x} \\
\frac{\sigma^2_f}{\sigma^2_x} N_j - 1
\end{pmatrix}
}{
(1 - \rho^2_{yx}) \left( \frac{\sigma^2_f}{\sigma^2_x} \right) + \left( \frac{\sigma^2_f}{\sigma^2_x} N_j - 1 \right) - \rho_{yx} \left( \frac{\sigma_f}{\sigma_x} \right) \begin{pmatrix}
\frac{\sigma^2_f}{\sigma^2_x} \\
\frac{\sigma^2_f}{\sigma^2_x} N_j - 1
\end{pmatrix}
}
.$$  \hspace{1cm} (A25)

Therefore, the $AdjY$ estimate is unbiased, and the bias of the $AvgE$ estimate approaches zero as $N_j \to \infty$.

It can also be shown that the $AvgE$ estimate of $\beta$ is less biased than is the OLS estimate of $\beta$. When $\rho_{yx} < 0$, $\hat{\beta}_{OLS} < \hat{\beta}_{AvgE} < 0$ because

$$
\hat{\beta}_{AvgE} = \frac{
\rho_{yx} \left( \frac{\sigma_f}{\sigma_x} \right) \begin{pmatrix}
\frac{\sigma^2_f}{\sigma^2_x} \\
\frac{\sigma^2_f}{\sigma^2_x} N_j - 1
\end{pmatrix}
}{
(1 - \rho^2_{yx}) \left( \frac{\sigma^2_f}{\sigma^2_x} \right) + \left( \frac{\sigma^2_f}{\sigma^2_x} N_j - 1 \right)
- \rho_{yx} \left( \frac{\sigma_f}{\sigma_x} \right) \begin{pmatrix}
\frac{\sigma^2_f}{\sigma^2_x} \\
\frac{\sigma^2_f}{\sigma^2_x} N_j - 1
\end{pmatrix}
}
< 0
$$  \hspace{1cm} (A26)

and

$$
\hat{\beta}_{AvgE} > \hat{\beta}_{OLS}
$$

$$
\rho_{yx} \left( \frac{\sigma_f}{\sigma_x} \right) \begin{pmatrix}
\frac{\sigma^2_f}{\sigma^2_x} \\
\frac{\sigma^2_f}{\sigma^2_x} N_j - 1
\end{pmatrix}
>(1 - \rho^2_{yx}) \left( \frac{\sigma^2_f}{\sigma^2_x} \right) + \left( \frac{\sigma^2_f}{\sigma^2_x} N_j - 1 \right)
$$
\[
\begin{align*}
\left( 1 - \rho^2_{xy} \right) \frac{\sigma^2_f}{\sigma^2_x} + \frac{\frac{\sigma^2_f}{\sigma^2_x}}{N_j - 1} < 1
\end{align*}
\]

When \( \rho_{xy} > 0 \), \( 0 < \hat{\beta}^{\text{avg}} < \hat{\beta}^{\text{OLS}} \) because

\[
\hat{\beta}^{\text{avg}} = \frac{\rho_{xy} \left( \frac{\sigma_f}{\sigma_x} \right) \left( \frac{\sigma^2_f}{\sigma^2_x} \right) + \frac{\frac{\sigma^2_f}{\sigma^2_x}}{N_j - 1}}{1 - \rho^2_{xy}} > 0
\]

and

\[
\rho_{xy} \left( \frac{\sigma_f}{\sigma_x} \right) \left( \frac{\sigma^2_f}{\sigma^2_x} \right) < \rho^2_{xy} \left( \frac{\sigma_f}{\sigma_x} \right)
\]

\[
\left( 1 - \rho^2_{xy} \right) \frac{\sigma^2_f}{\sigma^2_x} + \frac{\frac{\sigma^2_f}{\sigma^2_x}}{N_j - 1}
\]

\[
\left( \frac{\sigma^2_f}{\sigma^2_x} \right) \left( \frac{\sigma^2_f}{\sigma^2_x} \right) < 1
\]

\[
\left( 1 - \rho^2_{xy} \right) \frac{\sigma^2_f}{\sigma^2_x} + \frac{\frac{\sigma^2_f}{\sigma^2_x}}{N_j - 1} < 1
\]

\[
0 < \left( 1 - \rho^2_{xy} \right) \left( \frac{\sigma^2_f}{\sigma^2_x} \right)
\]

QED
A5. Proof of Proposition 5

Suppose that a researcher is interested in determining the effect of a treatment, $T$, on a variable $y$, when the true underlying structure of the data is given by

$$y_{ijt} = \beta_0 + \beta_1 P_i + \beta_2 T_{ij} + \beta_3 (P_i \times T_{ij}) + \varepsilon_{ijt}, \quad \text{(A30)}$$

where $y_{ijt}$ is the outcome for unit $i$, in group $j$, and period $t$; $P_i$ is an indicator equal to one if treatment has occurred by period $t$; and $T_{ij}$ is an indicator equal to one if unit $i$ is treated.

The difference-in-differences estimator, which is a direct estimation of Equation (A30), compares the mean of $y$ for the untreated and treated units in the pre- and post-treatment periods. This estimator provides a consistent estimate of $\beta_3$, which is the causal effect of the treatment on the outcome $y$.

The $AdjY$-style approach of estimating $\beta_3$ is to compare the pre- and post-treatment group-adjusted-$Y$ for the treated units. The $AdjY$ estimator is

$$y_{ijt} - \overline{y}_{jt} = \alpha^{AdjY} + \beta^{AdjY} P_i + \varepsilon^{AdjY}_{ijt}. \quad \text{(A31)}$$

Using Equation (A30), we can see that the group mean is given by

$$\overline{y}_{j,t} = \beta_0 + \beta_1 P_i + \beta_2 \overline{T}_j + \beta_3 (P_i \times \overline{T}_j) + \overline{\varepsilon}_{j,t}, \quad \text{(A32)}$$

and the group-adjusted mean is given by

$$y_{ijt} - \overline{y}_{jt} = \beta_2 (T_{ij} - \overline{T}_j) + \beta_3 (T_{ij} - \overline{T}_j) P_i + (\varepsilon_{ijt} - \overline{\varepsilon}_{jt}). \quad \text{(A33)}$$

Thus, the group mean for treated firms, where $T_{ij} = 1$, is equal to

$$y_{ijt} - \overline{y}_{jt} = \beta_2 (1 - \overline{T}_j) + \beta_3 (1 - \overline{T}_j) P_i + (\varepsilon_{ijt} - \overline{\varepsilon}_{jt}). \quad \text{(A34)}$$

Comparing the $AdjY$ estimator in Equation (A31) and the true underlying data structure in Equation (A34) reveals that the $AdjY$ estimator is not consistent. Specifically, $\hat{\beta}^{AdjY} = \beta_3 (1 - \overline{T})$, where $\overline{T}$ is the average share of untreated firms in a treated firm’s group. Intuitively, the $AdjY$ estimator exhibits an attenuation bias because it incorrectly demeans the data using an average of treated and untreated observations, which removes some of the treatment effect on $y$. The difference-in-differences estimator only removes the mean of untreated observations.
Equivalently, one can describe the bias in the AdjY estimator as an omitted variable bias. Equation (A34) can be written as

\[ y_{ij} - \bar{y}_j = \beta_2 \left( 1 - T_j \right) + \beta_i P_t - \beta_i P_T j + (\epsilon_{ij} - \bar{\epsilon}_j). \quad (A35) \]

Comparing (A35) to Equation (A31), the AdjY estimator fails to control for \( P_T j \), which affects the group-adjusted mean, \( y_{ij} - \bar{y}_j \), and is correlated with the independent variable of interest, \( P_t \). Biased estimates for the causal effect, \( \beta_2 \), result. QED

A6 – Proof That the AdjY and AvgE Estimators Are Still Biased when the Observation at Hand Is Not Excluded when Calculating Group Averages

Given \( \mu_X = \mu_f = 0 \) and Equation (1), the AdjY estimates when the observation at hand is not excluded, \( \tilde{\alpha}^{Adj} \) and \( \tilde{\beta}^{Adj} \), are given by the same calculation as in Appendix A1, except that

\[ \mathbf{y} = \begin{pmatrix} y_{i,j} - \bar{y}_j \\ \vdots \\ y_{N,j} - \bar{y}_j \end{pmatrix} = \begin{pmatrix} \beta X_{i,j} - \bar{r}_j + \epsilon_{i,j} \\ \vdots \\ \beta X_{N,j} - \bar{r}_j + \epsilon_{N,j} \end{pmatrix}, \quad (A36) \]

where

\[ \bar{y}_j = \alpha + f_j + \bar{r}_j \]

\[ \bar{r}_j = \frac{1}{N_j} \sum_{k \text{ group } j} \left( \beta X_{k,j} + \epsilon_{k,j} \right). \quad (A37) \]

The \( \text{plim}_{N \to \infty} \left( \frac{1}{N} \mathbf{X}'\mathbf{X} \right)^{-1} \) remains the same as in Appendix A1, and because the covariance between \( X_{i,j} \) and \( \bar{r}_j \), \( \sigma_{Xr} \), is given by

\[ \text{cov}(X_{i,j}, \bar{r}_j) = \text{cov} \left( X_{i,j}, \frac{1}{N_j} \sum_{k \text{ group } j} \left( \beta X_{k,j} + \epsilon_{k,j} \right) \right) = \text{cov} \left( X_{i,j}, \frac{\beta}{N_j} \sum_{k \text{ group } j} X_{k,j} \right) + \text{cov} \left( X_{i,j}, \frac{1}{N_j} \sum_{k \text{ group } j} \epsilon_{k,j} \right) = \text{cov} \left( X_{i,j}, \frac{\beta}{N_j} \sum_{k \text{ group } j} X_{k,j} \right) + \frac{\beta}{N_j} \sigma^2_X = \beta \left( \frac{N_j - 1}{N_j} \right) \sigma_{X,x,j} + \frac{\beta}{N_j} \sigma^2_X \quad (A38) \]
it can be shown that

\[
\lim_{N \to \infty} \frac{1}{N} \mathbf{X}'\mathbf{y} = \left[ \begin{array}{c}
1 & \cdots & 1 \\
X_{1,j} & \cdots & X_{N,j} \\
\vdots \\
\beta X_{N,j} - \bar{Y}_j + \epsilon_{N,j} \\
\end{array} \right] \left( \beta X_{1,j} - \bar{Y}_j + \epsilon_{1,j} \right) \\
\vdots \\
\frac{\sum_{j} \beta X_{1,j} - \bar{Y}_j + \epsilon_{1,j}}{N} \\
\frac{\sum_{j} X_{1,j} (\beta X_{1,j} - \bar{Y}_j + \epsilon_{1,j})}{N} = \beta \mathbf{\sigma}^2 - \mathbf{\sigma}_\epsilon \\
\end{array} \right].
\]  
(A39)

Therefore,

\[
b = \begin{bmatrix} \bar{\alpha}^{Adj} \\ \bar{\beta}^{Adj} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1/\mathbf{\sigma}^2 \end{bmatrix} \begin{bmatrix} 0 \\ \beta \left( \frac{N_j - 1}{N_j} \right) \left[ \mathbf{\sigma}^2 - \mathbf{\sigma}_{\epsilon,0} \right] \end{bmatrix} = \begin{bmatrix} 0 \\ \beta \left( \frac{N_j - 1}{N_j} \right) \left[ 1 - \rho_{X_{1,j},0} \right] \end{bmatrix}.
\]  
(A40)

Comparing (A40) to (A7), it is apparent that including the observation at hand just multiplies the AdjY estimate by a factor of \((N_j - 1) / N_j\).

Given \(\mu_{X} = \mu_{Y} = 0\) and Equation (1), the AvgE estimates \(\bar{\alpha}^{AvgE}\), \(\bar{\beta}^{AvgE}\), and \(\bar{\lambda}^{AvgE}\) are given by the same calculation as in Appendix A4, except that \(\mathbf{X}'\) is now given by

\[
\mathbf{X}' = \begin{bmatrix} 1 & \cdots & 1 \\
X_{1,j} & \cdots & X_{N,j} \\
\bar{Y}_j & \cdots & \bar{Y}_j \end{bmatrix}.
\]  
(A41)

It can then be shown that

Therefore,

\[
\mathbf{b} = \begin{bmatrix} \bar{\alpha}^{AvgE} \\ \bar{\beta}^{AvgE} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1/\mathbf{\sigma}^2 \end{bmatrix} \begin{bmatrix} 0 \\ \beta \left( \frac{N_j - 1}{N_j} \right) \left[ \mathbf{\sigma}^2 - \mathbf{\sigma}_{X_{1,j}X_{1,j}} \right] \end{bmatrix} = \begin{bmatrix} 0 \\ \beta \left( \frac{N_j - 1}{N_j} \right) \left[ 1 - \rho_{X_{1,j},X_{1,j}} \right] \end{bmatrix}.
\]  
(A40)
\[
\lim_{N \to \infty} \left( \frac{1}{N} \mathbf{X}'\mathbf{X} \right)^{-1} = \lim_{N \to \infty} \left( \frac{1}{N} \begin{pmatrix}
X_{1,j} & \cdots & X_{N,j} \\
\alpha + f_j + \bar{F}_j & \cdots & \alpha + f_j + \bar{F}_j
\end{pmatrix} \right)^{-1},
\]
\[
= \begin{pmatrix}
1 & \alpha \\
0 & \sigma^2_X + \sigma_{\bar{F}}^2 + \sigma_{\bar{F}}^2 + 2\sigma_{\bar{F}}^2 + \sigma_{\bar{F}}^2
\end{pmatrix}^{-1},
\]

where \( \sigma_{\bar{F}}^2 \) is the covariance between \( f_j \) and \( \bar{F}_j \), and \( \sigma_{\bar{F}}^2 \) is the variance of \( \bar{F}_j \).

And, it can be shown that
\[
\lim_{N \to \infty} \frac{1}{N} \mathbf{X}'\mathbf{y} = \lim_{N \to \infty} \left( \frac{1}{N} \begin{pmatrix}
X_{1,j} & \cdots & X_{N,j} \\
\alpha + f_j + \bar{F}_j & \cdots & \alpha + f_j + \bar{F}_j
\end{pmatrix} \right)
\]
\[
= \begin{pmatrix}
\frac{\sum_i (\alpha + \beta X_{i,j} + f_j + \epsilon_{i,j})}{N} \\
\frac{\sum_i X_{i,j} (\alpha + \beta X_{i,j} + f_j + \epsilon_{i,j})}{N} \\
\frac{\sum_i (\alpha + f_j + \bar{F}_j) (\alpha + \beta X_{i,j} + f_j + \epsilon_{i,j})}{N}
\end{pmatrix},
\]
\[
= \begin{pmatrix}
\alpha \\
\beta \sigma^2_X + \sigma_{\bar{F}}^2 \\
\alpha^2 + \beta \sigma_{\bar{F}}^2 + \sigma_{\bar{F}}^2 + \beta \sigma_{\bar{F}}^2 + \sigma_{\bar{F}}^2
\end{pmatrix}.
\]

Therefore,
Comparing (A44) to (A21), we see that, by including the observation at hand, $\sigma_{\tau_j}^2$ replaces $\sigma_{\tau_j}^2$, $\sigma_{x_{\cdot j}}$ replaces $\sigma_{x_{\cdot j}}$, and $\sigma_{\tau_{\cdot j}}$ replaces $\sigma_{\tau_{\cdot j}}$ in the expression for the $\text{AvgE}$ estimate. But as shown in Equations (A5) and (A38),

$$\sigma_{x_{\cdot j}} = \left( \frac{N_j}{N_j - 1} \right) \sigma_{x_{\cdot j}} + \left( \frac{\beta}{N_j} \right) \sigma_{x_{\cdot j}}^2,$$

and it can be shown that $\sigma_{\tau_{\cdot j}} = \sigma_{\tau_{\cdot j}}$. Moreover,
\[ \sigma^2_v = \text{var} \left( \frac{1}{N_j} \sum_{k \in \text{group } j} (\beta X_{kj} + \epsilon_{kj}) \right) \]
\[ = \text{var} \left( \frac{\beta}{N_j} \sum_{k \in \text{group } j} X_{kj} + \frac{1}{N_j} \sum_{k \in \text{group } j} \epsilon_{kj} \right) \]
\[ = \left( \frac{\beta}{N_j} \right)^2 \text{var} \left( \sum_{k \in \text{group } j} X_{kj} \right) + \left( \frac{1}{N_j} \right)^2 \text{var} \left( \sum_{k \in \text{group } j} \epsilon_{kj} \right) \]
\[ = \left( \frac{\beta}{N_j} \right)^2 \left[ \sum_{k \in \text{group } j} \sigma_x^2 + \sum_{k \in \text{group } j} \sum_{m \in \text{group } j} \text{cov}(X_{kj}, X_{mj}) \right] + \left( \frac{1}{N_j} \right)^2 \sum_j \sigma^2_\epsilon \]
\[ = \left( \frac{\beta}{N_j} \right)^2 \left[ N_j \sigma_x^2 + N_j (N_j - 1) \text{cov}(X_{kj}, X_{kj}) \right] + \frac{\sigma^2_\epsilon}{N_j} \]
\[ = \frac{\beta^2 \sigma_x^2}{N_j} + \frac{\beta^2 (N_j - 1) \sigma_{X_{kj}X_{kj}}}{N_j} + \frac{\sigma^2_\epsilon}{N_j} \]
\[ = \frac{\beta^2 \sigma_x^2}{N_j} + \frac{\beta^2 (N_j - 1) \sigma_{X_{kj}X_{kj}}}{N_j} + \frac{\sigma^2_\epsilon}{N_j} \]
\[ = \frac{\beta^2 \sigma_x^2 - \sigma_{X_{kj}X_{kj}}}{N_j} + \frac{\sigma^2_\epsilon}{N_j} \]
\[ = \frac{\beta^2 \sigma_x^2 - \sigma_{X_{kj}X_{kj}}}{N_j} + \frac{\sigma^2_\epsilon}{N_j} \]

which is similar to \( \sigma^2_v \), as defined in Equation (A23); the expression only differs in that the denominator in the second term is \( N_j \) rather than by \( N_j - 1 \).

QED