University of Pennsylvania
ScholarlyCommons

Finance Papers | Wharton Faculty Research

2014

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

Todd A. Gormley
University of Pennsylvania

David A. Matsa

Follow this and additional works at: https://repository.upenn.edu/fnce_papers

Part of the Finance Commons, and the Finance and Financial Management Commons

Recommended Citation

Gormley, T. A., & Matsa, D. A. (2014). Common Errors: How to (and Not to) Control for Unobserved Heterogeneity. *Review of Financial Studies, 27* (2), 617-661. http://dx.doi.org/10.1093/rfs/hht047

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/fnce_papers/318
For more information, please contact repository@pobox.upenn.edu.

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

Abstract

Controlling for unobserved heterogeneity (or “common errors”), such as industry-specific shocks, is a fundamental challenge in empirical research.This paper discusses the limitations of two approaches widely used in corporate finance and asset pricing research: demeaning the dependent variable with respect to the group (e.g., “industry-adjusting”) and adding the mean of the group's dependent variable as a control. We show that these methods produce inconsistent estimates and can distort inference. 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.

Disciplines

Finance | Finance and Financial Management

This journal article is available at ScholarlyCommons: https://repository.upenn.edu/fnce_papers/318

Common Errors: How to (and Not to) Control for Unobserved Heterogeneity[A]

Todd A. Gormley[B] and David A. Matsa[C]

August 3, 2013

Abstract

Controlling for unobserved heterogeneity (or "common errors"), such as industry-specific shocks, is a fundamental challenge in empirical research. This paper discusses the limitations of two approaches widely used in corporate finance and asset pricing research: demeaning the dependent variable with respect to the group (e.g., "industry-adjusting") and adding the mean of the group's dependent variable as a control. We show that these methods produce inconsistent estimates and can distort inference. 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

---

[A] Computer code and additional resources related to the issues discussed in this paper are posted online at http://www.kellogg.northwestern.edu/faculty/matsa/htm/fe.htm. For helpful comments, we thank Michael Anderson, Joshua Angrist, Xavier Giroud, Christian Hansen, Dirk Jenter, Sandy Klasa, Jonathan Lewellen, Alexander Ljungqvist, Mitchell Petersen, Nagpurnanand R. Prabhala, Michael Roberts, Nick Souleles, Michael R. Wagner, Toni Whited, and Jeffrey Wooldridge, as well as the seminar participants at Hong Kong University of Science and Technology, Massachusetts Institute of Technology (Sloan), Nanyang Technological University, National University of Singapore, Singapore Management University, University of Arizona (Eller), University of Pennsylvania (Wharton), the American Finance Association annual meeting, the Financial Intermediation Research Society conference, and the Rothschild Caesarea Center Conference. Matthew Denes, Christine Dobridge, Jingling Guan, Jillian Popadak, and Kanis Saengchote provided 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.

[B] 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.

[C] 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 asset prices and most corporate policies 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. For example, unobserved risk factors, which affect both stock returns and corporate decisions, are often common to firms of similar size. Potential unobserved factors abound: unobserved differences in local economic environments, management quality, and the cost of capital, to name a few.

Although the empirical finance literature uses various estimation strategies to control for unobserved group 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 provide inconsistent estimates and can distort inference.

We focus on two popular estimation strategies that are applied when there are a large number of groups and the number of observations per group is small relative to the number of groups (e.g., firm-panel data that is grouped into industry-years). The first estimation strategy, which we refer to as "adjusted-Y" (AdjY), demeans 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 mean of the group's dependent variable 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 in 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.[1] Among articles published in these three journals in 2008–2010, we found over 60 articles, split about evenly between corporate finance and asset pricing, that employed at least one of the two

[1] The exact origin of the two estimators in finance is unclear; we suspect they were adapted from 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.

Common Errors: How to (and Not to) Control for Unobserved Heterogeneity[A]

Todd A. Gormley[B] and David A. Matsa[C]

Abstract

Controlling for unobserved heterogeneity (or "common errors"), such as industry-specific shocks, is a fundamental challenge in empirical research. This paper discusses the limitations of two approaches widely used in corporate finance and asset pricing research: demeaning the dependent variable with respect to the group (e.g., "industry-adjusting") and adding the mean of the group's dependent variable as a control. We show that these methods produce inconsistent estimates and can distort inference. 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

---

[A] Computer code and additional resources related to the issues discussed in this paper are posted online at http://www.kellogg.northwestern.edu/faculty/matsa/htm/fe.htm. For helpful comments, we thank Michael Anderson, Joshua Angrist, Xavier Giroud, Christian Hansen, Dirk Jenter, Sandy Klasa, Jonathan Lewellen, Alexander Ljungqvist, Mitchell Petersen, Nagpurnanand R. Prabhala, Michael Roberts, Nick Souleles, Michael R. Wagner, Toni Whited, and Jeffrey Wooldridge, as well as the seminar participants at Hong Kong University of Science and Technology, Massachusetts Institute of Technology (Sloan), Nanyang Technological University, National University of Singapore, Singapore Management University, University of Arizona (Eller), University of Pennsylvania (Wharton), the American Finance Association annual meeting, the Financial Intermediation Research Society conference, and the Rothschild Caesarea Center Conference. Matthew Denes, Christine Dobridge, Jingling Guan, Jillian Popadak, and Kanis Saengchote provided 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.

[B] 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.

[C] 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 asset prices and most corporate policies 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. For example, unobserved risk factors, which affect both stock returns and corporate decisions, are often common to firms of similar size. Potential unobserved factors abound: unobserved differences in local economic environments, management quality, and the cost of capital, to name a few.

Although the empirical finance literature uses various estimation strategies to control for unobserved group 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 provide inconsistent estimates and can distort inference.

We focus on two popular estimation strategies that are applied when there are a large number of groups and the number of observations per group is small relative to the number of groups (e.g., firm-panel data that is grouped into industry-years). The first estimation strategy, which we refer to as "adjusted-Y" (AdjY), demeans 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 mean of the group's dependent variable 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 in 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.[1] Among articles published in these three journals in 2008–2010, we found over 60 articles, split about evenly between corporate finance and asset pricing, that employed at least one of the two

[1] The exact origin of the two estimators in finance is unclear; we suspect they were adapted from 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.

techniques. The techniques are used to study a variety of finance topics, including banking, capital structure, corporate boards, governance, executive compensation, and corporate control. Articles using these estimation methods have also been published in economics, including in the *American Economic Review*, *Journal of Political Economy*, and *Quarterly Journal of Economics*, and in accounting, including the *Accounting Review, Journal of Accounting and Economics*, and *Journal of Accounting Research*.

Our paper shows that, despite their popularity, the AdjY and AvgE estimators rarely provide consistent estimates; both estimators can exhibit severe biases (where "bias" here refers to the difference between the probability limit of the estimate and the true parameter). The AdjY estimator suffers from an omitted variable bias because it fails to control for the group average of the independent variables. This omission is problematic when any explanatory variable is correlated with its group average, which is likely in practice. The AvgE estimator suffers from a measurement error bias because the sample mean of the group's dependent variable measures the true unobserved heterogeneity with error. AvgE is inconsistent when AdjY is inconsistent and also when any independent variable is correlated with the unobserved heterogeneity. Even when the underlying data structure exactly matches the AdjY or AvgE specifications, both estimators are inconsistent. 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 heterogeneity. The FE estimator, which instead adds group indicator variables to the OLS estimation, is consistent in the presence of unobserved group heterogeneity. When there is only one source of unobserved group heterogeneity, the FE estimator is equivalent to demeaning *all* of the dependent and independent variables with respect to the group and then estimating the model 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. 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 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, AvgE, and related estimators should *not* be used to control for unobserved group heterogeneity. Any estimation that transforms the dependent variable but not the independent variables typically yields inconsistent estimates. For example, subtracting the group median, or the mean or median of a comparable set of firms, from the dependent variable also yields inconsistent estimates by failing to account for how the corresponding median or mean of the independent variables affects the adjusted dependent variable.

Our findings apply to a diverse set of estimations undertaken in the literature. The practice of industry-adjusting dependent variables is common in many corporate finance papers. Even a simple comparison of industry- or benchmark-adjusted outcomes before and after events—as in many analyses of corporate control transactions, stock issues, and other sets of 0/1 events—does not reveal the true effect of the events. Corporate governance analyses of the effects of business combination laws across U.S. states while controlling for industry-year and state-year averages of the dependent variable are also not properly specified. Our criticism also applies to estimators in some asset pricing studies. The method of characteristically adjusting stock returns in asset pricing subtracts the return of a benchmark portfolio containing stocks with similar characteristics, before sorting and comparing these stock returns across subsamples. This method is problematic because it does not control for how the variable used to sort the adjusted stock returns varies across the benchmark portfolios.

Our analysis also highlights related problems with other estimators that are not designed to control for unobserved heterogeneity. For example, the omitted variable problem of AdjY applies to *any* dependent variable that is constructed using multiple observations. Our analysis also raises concerns about instrumental variable (IV) estimators that instrument for an endogenous independent variable using its group average. This estimator's exclusion restriction is violated whenever an unobserved group factor is correlated with the regressor.

FE estimators should be used instead of AdjY or AvgE to control for unobserved heterogeneity. 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 the independent variables. For example, rather than industry-adjusting a dependent variable or controlling for the industry mean of the dependent variable, researchers should instead estimate a model with industry fixed effects. Likewise, rather than correlating benchmark-portfolio-adjusted stock returns with an explanatory variable of interest, a researcher should instead estimate a model with fixed effects for each benchmark portfolio.[2]

The FE estimator, however, also has limitations. Although the FE estimator controls for unobserved group heterogeneties, 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 attenuation bias in the presence of measurement error. We discuss these limitations, how they can be addressed, and when FE estimation is appropriate.

Finally, we address another limitation of FE that has motivated some researchers to use AdjY or AvgE rather than FE—computational difficulties that can arise when estimating FE models that have 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. For example, executive compensation may be affected by unobserved managerial skill and by unobserved firm quality (Graham, Li, and Qui, 2012; Coles and Li, 2011a). Likewise, researchers who use firm-level data are increasingly concerned about both unobserved firm-level characteristics and time-varying heterogeneity across industries, such as industry-level shocks to demand. 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-

[2] To help interested researchers, we have also posted code and additional resources on our website, http://www.kellogg.northwestern.edu/faculty/matsa/htm/fe.htm, to show how common implementations of AdjY and AvgE can be transformed into consistent FE estimators.

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 software Stata.

Overall, our paper provides practical guidance on empirical estimation in the presence of unobserved group heterogeneity, which is 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; Erickson and Whited (2012) 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. In Section 1, we describe the underlying identification concern of estimating a model with unobserved group heterogeneity and why AdjY and AvgE provide inconsistent estimates. In Section 2, we contrast the AdjY and AvgE estimators with the FE estimator and discuss why other related estimation techniques, which are commonly used in the literature, yield inconsistent estimates. In Section 3, we show that differences between the estimation techniques can be important in practice. In Section 4, we discuss the limitations of the FE estimator and describe when its use is appropriate. We conclude in Section 5 and provide derivations and proofs in the Appendix.

1. Estimation Techniques Used to Control for Unobserved Group Heterogeneity

Consider the case in which an independent variable of interest, X, affects a dependent variable, y, which also has unobserved group heterogeneity that is possibly correlated with X. Specifically, assume the data exhibits the following structure:

$$
\begin{aligned}
&y_{i,j} = \beta X_{i,j} + f_i + \varepsilon_{i,j} \\
&\text{var}(\varepsilon_{i,j}) = \sigma_\varepsilon^2, \mu_\varepsilon = 0 \\
&\text{var}(X_{i,j}) = \sigma_X^2 \\
&\text{var}(f_i) = \sigma_f^2 \\
&\text{cov}(f_i, \varepsilon_{i,j}) = 0 \\
&\text{cov}(X_{i,j}, \varepsilon_{i,j}) = \text{cov}(X_{i,j}, \varepsilon_{i,-j}) = 0 \\
&\text{cov}(X_{i,j}, f_i) = \sigma_{Xf},
\end{aligned}
\tag{1}
$$

where i indexes groups of observations (e.g., industries) and j indexes observations within each group (e.g., firms). There is a random sample of N groups ($y_{i,1},\ldots,y_{i,J}$, $X_{i,1},\ldots, X_{i,J}$) with J observations per group. As is typical in finance regressions, we assume that J is small, N is large, and both the independent variable of interest, X, and the residual, ε, are *i.i.d.* across groups but not necessarily *i.i.d.* within groups, and neither f nor the independent variable of interest, X, covary with the residual, ε. For ease of exposition, we assume that the intercept and the means of both the group term and the independent variable are zero (i.e., $\mu_X = \mu_f = 0$); these assumptions simplify the analysis but have no effect on the estimate of β under the different estimation techniques we analyze.

The model in Equation (11) can be augmented to reflect more complicated sources of unobserved heterogeneity without affecting our subsequent analysis. For example, a model with two types of unobserved heterogeneity, such as firm and year group effects in panel data, can be captured by adding an additional heterogeneity term to Equation (1). 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.

It is well known that using OLS to estimate β, the effect of X on y, yields an inconsistent estimate when there exists a nonzero covariance, σ_{Xf}, between the unobserved heterogeneity and independent variable of interest.[3] OLS estimates the following specification:

$$
y_{i,j} = \beta^{OLS} X_{i,j} + u_{i,j}^{OLS}. \tag{2}
$$

The OLS estimate is

$$
\hat{\beta}^{OLS} = \beta + \frac{\sigma_{Xf}}{\sigma_X^2}. \tag{3}
$$

By failing to control for the group term, $\hat{\beta}^{OLS}$ conflates the effects of X and f on the dependent variable, y.

[3] Throughout the paper, we use the standard large-sample approach to determine an estimate's consistency by taking the number of groups, N, to infinity, while holding group size, J, constant.

The bias, σ_{Xf} / σ_X^2, represents the standard omitted variable bias in a univariate regression; it equals the coefficient from a regression of the omitted variable, *f*, onto the included variable, *X*, multiplied by the coefficient on the omitted variable in the true model, which in this case is 1 [see Angrist and Pischke (2009), Section 3.2.2, for a derivation of the omitted variable bias formula].

Because OLS is inconsistent when $\sigma_{Xf} \neq 0$, researchers must rely on other estimation techniques. Two popular approaches are "adjusted-Y" (AdjY), which demeans the dependent variable with respect to the group before estimating the model with OLS, and "average effects" (AvgE), which uses the group's mean of the dependent variable as a control in an OLS specification. In this section, we describe the AdjY and AvgE estimates and discuss why they typically lead to inconsistent estimates of the coefficient of interest, β. The source of the bias extends to models with more complicated data structures.

1.1 Adjusted-Y estimation

The AdjY estimator attempts to remove the influence of the group term from the dependent variable by demeaning the dependent variable within each group. AdjY estimation is applied, for example, at the industry level in firm-panel datasets by subtracting the industry-mean from the dependent variable. When this adjustment is applied at the industry or industry-year level, researchers typically refer to the dependent variable as being "industry-adjusted."

More specifically, the researcher calculates the group mean, $\overline{y}_i$, as

$$\overline{y}_i = \frac{1}{J} \sum_{k \in group\ i} y_{i,k} = \frac{1}{J} \sum_{k \in group\ i} \left(\beta X_{i,k} + f_i + \varepsilon_{i,k} \right) \tag{4}$$

and estimates the following model using OLS:[4]

$$y_{i,j} - \overline{y}_i = \beta^{AdjY} X_{i,j} + u_{i,j}^{AdjY} . \tag{5}$$

The AdjY estimation, however, does not provide a consistent estimate of β because the estimation suffers from an omitted variable problem. To see this, it is helpful to re-express the dependent variable's sample group mean as

$$\overline{y}_i = \beta \overline{X}_i + f_i + \overline{\varepsilon}_i , \tag{6}$$

[4] Another common implementation of AdjY is to demean the dependent variable using the sample group's mean after excluding the observation at hand. In other cases, the median is used. Redefining the group mean, $\overline{y}_i$, to reflect these other implementations of AdjY does not affect our subsequent propositions; both of these approaches also yield inconsistent estimates. We discuss these estimators in Section 2.

where $\bar{X}_i$ and $\bar{\varepsilon}_i$ are the corresponding sample group means for X and ε. In the presence of unobserved group heterogeneity, as in Equation (11), the dependent variable in AdjY estimation can thus be written as

$$y_{i,j} - \bar{y}_i = \beta X_{i,j} - \beta \bar{X}_i + \varepsilon_{i,j} - \bar{\varepsilon}_i . \quad (7)$$

Comparing the AdjY estimation (Equation (55)) to the true data structure (Equation (77)), we see that AdjY fails to control for $\bar{X}_i$. This leads to a biased estimate for β if $\bar{X}_i$ is correlated with the independent variable, X. In other words, the covariance between $X_{i,j}$ and $u_{i,j}^{AdjY}$ is nonzero when the correlation between $\bar{X}_i$ and $X_{i,j}$ is nonzero. AdjY also fails to control for $\bar{\varepsilon}_i$, but this does not bias the estimate for β under the data structure assumed in Equation (11).

As shown in Proposition 1, letting $\sigma_{X\bar{X}}$ represent the covariance between $X_{i,j}$ and its group mean, $\bar{X}_i$, we can derive the sign and magnitude of the bias in the AdjY estimate for β.

Proposition 1. In the presence of unobserved group heterogeneity, as in Equation (11), the AdjY estimator yields an inconsistent estimate for β. Specifically,

$$\hat{\beta}^{AdjY} = \beta - \beta \frac{\sigma_{X\bar{X}}}{\sigma_X^2}.$$

Similar to the OLS estimate, the bias of AdjY follows the standard omitted variable bias formula. The bias, $\beta\sigma_{X\bar{X}} / \sigma_X^2$, equals the coefficient from a regression of the omitted variable, $\bar{X}_i$, on the included variable, $X_{i,j}$, multiplied by the omitted variable's coefficient in the true underlying model, which is $-\beta$ (see Equation (77)). If X co-varies positively with its group mean, then the AdjY estimate for β exhibits an attenuation bias. And, inversely, if X is negatively correlated with its group mean, then the AdjY estimate for β is biased away from zero.[5]

In practice, positive covariance between X and its group mean is common, causing AdjY

[5] In practice, the bias of $\hat{\beta}^{AdjY}$ is typically attenuating because it is unusual for X to be negatively correlated with its group mean when there is also a group component, f, that has nonzero correlation with X (as in Equation (11)). Specifically, the covariance matrix for the underlying data structure implied by Equation (11) is positive definite only if $\rho_{X_{i,j}X_{i,-j}} > (J\rho_{Xf}^2 - 1)/(J-1)$, where $\rho_{X_{i,j}X_{i,-j}}$ is the correlation between $X_{i,j}$ and $X_{i,-j}$, and ρ_{Xf} is the correlation between $X_{i,j}$ and f_i. While this condition places a lower bound on the $\sigma_{X\bar{X}} / \sigma_X^2$ term of the AdjY bias, it is possible for this bound to be less than zero, in which case $\hat{\beta}^{AdjY}$ overestimates the magnitude of β.

estimates to be inconsistent. For example, consider a standard firm-level capital structure estimation, where leverage is regressed onto multiple independent variables, such as the return on assets, bankruptcy risk, and the market-to-book ratio. Because firms in the same industry are subject to common demand and technology shocks, their leverage, return on assets, bankruptcy risk, and the other regressors typically co-vary with their industry averages. By industry-adjusting only leverage, the AdjY approach removes the unobserved heterogeneity in leverage, but fails to account for the covariance between the independent variables and their industry averages in the transformed data. For this reason, AdjY estimation (including so-called "industry-adjusting") is inconsistent.

The bias in the AdjY 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 eliminate the identification problem—the estimation's error term still contains the omitted group average of the independent variable. Moreover, $X_{i,j}$ and $\bar{X}_i$ are typically correlated even when $X_{i,j}$ and f_i are not, so AdjY is inconsistent even when 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.

As shown in Proposition 2, 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 as follows:

$$
\begin{aligned}
& y_{i,j} = \beta X_{i,j} + \gamma Z_{i,j} + f_i + \varepsilon_{i,j} \\
& \text{var}(\varepsilon_{i,j}) = \sigma_\varepsilon^2, \mu_\varepsilon = 0 \\
& \text{var}(X_{i,j}) = \sigma_X^2 \\
& \text{var}(f_i) = \sigma_f^2 \\
& \text{var}(Z_{i,j}) = \sigma_Z^2 \\
& \text{cov}(f_i, \varepsilon_{i,j}) = 0 \\
& \text{cov}(X_{i,j}, \varepsilon_{i,j}) = \text{cov}(X_{i,j}, \varepsilon_{i,-j}) = 0 \\
& \text{cov}(Z_{i,j}, \varepsilon_{i,j}) = \text{cov}(Z_{i,j}, \varepsilon_{i,-j}) = 0 \\
& \text{cov}(X_{i,j}, f_i) = \sigma_{Xf} \\
& \text{cov}(X_{i,j}, Z_{i,j}) = \sigma_{XZ} \\
& \text{cov}(Z_{i,j}, f_i) = \sigma_{Zf}.
\end{aligned}
\tag{8}
$$

There is still just one type of unobserved group heterogeneity, f_i, 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 effect, f_i, do not co-vary with the error, $\varepsilon_{i,j}$. The two independent variables, however, do co-vary with each other and with the unobserved group heterogeneity, f_i. For ease of exposition, we again assume without loss of generality that the intercept and the means of both the unobserved heterogeneity and the independent variables are zero (i.e., $\mu_X = \mu_Z = \mu_f = 0$).

Proposition 2. In the presence of unobserved group heterogeneity and two independent variables, as in Equation (88), the AdjY estimator yields inconsistent estimates for both β and γ. Specifically,

$$\hat{\beta}^{AdjY} = \beta + \frac{\beta\left(\sigma_{XZ}\sigma_{Z\bar{X}} - \sigma_Z^2\sigma_{X\bar{X}}\right) + \gamma\left(\sigma_{XZ}\sigma_{Z\bar{Z}} - \sigma_Z^2\sigma_{X\bar{Z}}\right)}{\sigma_Z^2\sigma_X^2 - \sigma_{XZ}^2}$$

$$\hat{\gamma}^{AdjY} = \gamma + \frac{\beta\left(\sigma_{XZ}\sigma_{X\bar{X}} - \sigma_X^2\sigma_{Z\bar{X}}\right) + \gamma\left(\sigma_{XZ}\sigma_{X\bar{Z}} - \sigma_X^2\sigma_{Z\bar{Z}}\right)}{\sigma_Z^2\sigma_X^2 - \sigma_{XZ}^2},$$

where $\sigma_{Z\bar{Z}}$ is the covariance between $Z_{i,j}$ and $\bar{Z}_i$, $\sigma_{Z\bar{X}}$ is the covariance between $Z_{i,j}$ and $\bar{X}_i$,, and $\sigma_{X\bar{Z}}$ is the covariance between $X_{i,j}$ and $\bar{Z}_i$.

These expressions show that predicting the direction and magnitude of the bias in the AdjY estimator is not straightforward, and the sign of $\hat{\beta}^{AdjY}$ may not even match the sign of the true β. The sign and magnitude of the bias depends on both of the underlying coefficients, β and γ, the relative variances of *X* and *Z*, and the covariance of *X* and *Z* with each other and the omitted variables, $\bar{X}_i$ and $\bar{Z}_i$.[6]

1.2 Average effects estimation

AvgE estimation approaches the problem of unobserved heterogeneity differently. Instead of adjusting the dependent variable, AvgE uses a proxy, the group's sample mean, $\bar{y}_i$, to control for the unobserved variation, f_i. A common implementation of AvgE uses observations' state-year mean to control for time-varying differences in local economic environments. When there is one independent

[6] The bias in Proposition 2 follows the standard omitted variable bias formula for a multivariate regression. See Section 8.4.2 of Greene (2002) or page 61, footnote 14, of Angrist and Pischke (2009) for details.

variable, the AvgE approach estimates the following regression:

$$y_{i,j} = \beta^{AvgE} X_{i,j} + \lambda^{AvgE} \overline{y}_i + u_{i,j}^{AvgE}. \tag{9}$$

Proposition 3 shows that, in the presence of unobserved group heterogeneity, as in Equation (11), AvgE estimation is inconsistent. The underlying problem is that AvgE suffers from measurement error. As seen from Equation (66), $f_i = \overline{y}_i - \beta \overline{X}_i - \overline{\varepsilon}_i$; thus, the sample dependent variable's group means, $\overline{y}_i$, measure the unobserved variation, f_i, with error $-\beta \overline{X}_i - \overline{\varepsilon}_i$. Measurement error biases the coefficient on the mismeasured regressor, along with the coefficients on *all* other regressors. Moreover, because the measurement error co-varies with both the mismeasured variable and the errors, this is not classical measurement error, and the bias on $\hat{\lambda}^{AvgE}$ is not necessarily attenuating.

Proposition 3. In the presence of unobserved group heterogeneity, as in Equation (11), the AvgE estimator yields an inconsistent estimate for β. Specifically,

$$\hat{\beta}^{AvgE} = \beta + \frac{\sigma_{Xf}\left(\beta\sigma_{f\overline{X}} + \beta^2\sigma_{\overline{X}}^2 + \sigma_{\overline{\varepsilon}}^2 - \sigma_{\varepsilon\overline{\varepsilon}}\right) - \beta\sigma_{X\overline{X}}\left(\sigma_f^2 + \beta\sigma_{f\overline{X}} + \sigma_{\varepsilon\overline{\varepsilon}}\right)}{\sigma_X^2\left(\sigma_f^2 + 2\beta\sigma_{f\overline{X}} + \beta^2\sigma_{\overline{X}}^2 + \sigma_{\overline{\varepsilon}}^2\right) - \left(\sigma_{Xf} + \beta\sigma_{X\overline{X}}\right)^2},$$

where $\sigma_{\overline{X}}^2$ is the variance of $\overline{X}_i$, $\sigma_{\overline{\varepsilon}}^2$ is the variance of $\overline{\varepsilon}_i$, $\sigma_{f\overline{X}}$ is the covariance between f_i and $\overline{X}_i$, and $\sigma_{\varepsilon\overline{\varepsilon}}$ is the covariance between ε_i and $\overline{\varepsilon}_i$.

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, $-\beta \overline{X}_i - \overline{\varepsilon}_i$, can co-vary with both the mismeasured variable, f_i, when $\sigma_{Xf} \neq 0$ and with $X_{i,j}$ when $\sigma_{X\overline{X}} \neq 0$.[7] As shown in Proposition 3, either a non-zero σ_{Xf} *or* a non-zero $\sigma_{X\overline{X}}$ causes the AvgE estimate of β to be inconsistent.

These conditions cause the AvgE estimator to be inconsistent in most applications. For example, in a firm-level AvgE estimation, where a researcher uses the state-year average to control for time-varying differences in local economic environments, any covariance between the independent variables,

[7] It is possible, however, to put bounds on the bias. For example, as we discuss in the Appendix, a researcher can use the methods in Erickson and Whited (2005) and assumptions about the correlation between $\overline{y}_i$ and f_i to determine whether the sign of an AvgE estimate is correct. See also Krasker and Pratt (1986), Erickson (1993), and Hu (2006).

such as the return on assets, and their state-average causes AvgE to be inconsistent. Such correlations are common because firms located in the same state are subject to similar local demand shocks and investment opportunities.

1.3 Relative performance of OLS, AdjY, and AvgE

To better understand the sources of the bias and to compare the relative performance of OLS, AdjY, and AvgE, it is helpful to re-express the key covariance term $\sigma_{X\bar{X}}$ that contributes to each estimate's bias. To this end, it is helpful to separate $X_{i,j}$ into its group and idiosyncratic components. Assume that $X_{i,j} = \bar{x}_i + w_{i,j}$, where the group means $\bar{x}_i$ are *i.i.d.* with mean zero and variance $\sigma^2_{\bar{x}}$, the idiosyncratic components $w_{i,j}$ are distributed with mean 0 and variance σ^2_w, and $\text{cov}(\bar{x}_i, w_{i,j}) = 0$. It can be shown that

$$\sigma_{X\bar{X}} = \sigma^2_{\bar{x}} + \left(\frac{J-1}{J}\right)\sigma_{w_{i,j},w_{i,-j}} + \frac{\sigma^2_w}{J}, \tag{10}$$

where $\sigma_{w_{i,j}w_{i,-j}}$ is the covariance between $w_{i,j}$ and $w_{i,-j}$ (proof in the Appendix).

The expression in Equation (1010) sheds additional light on when the AdjY and AvgE estimators are inconsistent. As shown in Propositions 1, 2, and 3, AdjY and AvgE are inconsistent when $\sigma_{X\bar{X}} \neq 0$. Equation (10) shows that $\sigma_{X\bar{X}} \neq 0$ whenever observations either have different means across groups (such that $\sigma^2_{\bar{x}} \neq 0$) or are not independent within groups (such that $\sigma_{w_{i,j}w_{i,-j}} \neq 0$). In most finance applications, observations have different means across industries, geographies, or other groupings. Even if all groups had the same mean, $\sigma_{X\bar{X}}$ would still be non-zero if the observations within groups are not independent, which is also common in practice. For example, time-series observations are often serially correlated. Finally, even if observations are *i.i.d.* within and across groups, $\sigma_{X\bar{X}}$ is still not equal zero because $\bar{X}_i$ includes $X_{i,j}$, which creates a mechanical, nonzero correlation between an observation and its sample group average for finite *J*.

Defining the group average for each observation to exclude the observation at hand (as is sometimes done in practice), such that

$$\overline{y}_{i,-j} = \frac{1}{J-1} \sum_{\substack{k \in group i \\ k \neq j}} y_{i,k} \,, \tag{11}$$

simplifies the expression for $\sigma_{X\bar{X}}$ but does not resolve the underlying identification problem. Although excluding the observation at hand when calculating group means removes the mechanical correlation between $X_{i,j}$ and its sample group average, the covariance between $X_{i,j}$ and its group average is still non-zero whenever observations have different means across groups or are not independent within groups; specifically, $\sigma_{X\bar{X}} = \sigma^2_{\bar{x}} + \sigma_{w_{i,j}, w_{i,-j}}$ (proof in the Appendix).

In some cases, AvgE and AdjY will provide a less-biased estimate of β 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.[8] Under certain parameters, OLS will actually be less biased then 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 then 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. In the case of just one independent variable, as in Table 1, the AdjY estimator cannot incorrectly flip the sign of β. But as shown 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 has the correct sign.

The correlation between the independent variable of interest, *X*, and the unobserved group heterogeneity, *f*, has a large effect on the relative performance of each estimator. Figure 1 graphs OLS, AdjY, and AvgE estimates of Equation (11) as functions of various parameter values when $\beta = 1$. Each panel shows the effect of varying a specific parameter in the data structure, while holding the rest constant. When not being varied, the default parameter values are as follows: $\rho_{Xf} = 0.25$; $\sigma_{\bar{x}} / \sigma_w = 0.25$; $\rho_{w_{i,j}, w_{i,-j}} = 0.5$; $\sigma_f / \sigma_X = 1$; $\sigma_\varepsilon / \sigma_X = 1$; and $J = 10$. Panel A plots the impact of the correlation between the independent variable *X* and the unobserved group heterogeneity *f* on each estimator. AdjY is less

[8] To focus on the key determinants of the biases, we assume in Table 1 (and later in Figure 1) that groups are defined as in Equation (1111); this eliminates the bias that arises from the mechanical correlation between $X_{i,j}$ and its sample average, which is less interesting. To simplify the analysis, we also assume that the errors are *i.i.d.* Analytical solutions for the AdjY and AvgE estimates for β, expressed in terms of $\rho_{Xf}, \sigma_{\bar{x}} / \sigma_w, \rho_{w_{i,j}, w_{i,-j}}, \sigma_f / \sigma_X, \sigma_\varepsilon / \sigma_X$, and *J*, are provided in the Appendix.

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 ρ_{Xf}, whereas the OLS bias increases linearly ρ_{Xf}. The AvgE bias, in contrast, is nonlinear in ρ_{Xf}. Under these parameters, AvgE is more biased than both OLS and AdjY for low values of ρ_{Xf}, whereas for high values of ρ_{Xf}, AvgE is less biased than both OLS and AdjY.

The relative variation across versus within groups, $\sigma_{\bar{x}} / \sigma_{w}$, and the correlation between observations within a group, $\rho_{w_{i,j}w_{i,-j}}$, also affect the relative performance of the AdjY and AvgE estimators. As shown in Panel B of Figure 1, the sign and magnitude of the bias for both AvgE and AdjY increase in $\sigma_{\bar{x}} / \sigma_{w}$, whereas the OLS estimate is unaffected. This is because AdjY and AvgE biases depend on $\sigma_{X\bar{X}}$, which increases in $\sigma_{\bar{x}}$, whereas the bias in OLS does not. For the same reason, correlation between observations in a group, $\rho_{w_{i,j}w_{i,-j}}$, increases the bias of AdjY and AvgE but not OLS, as shown in Panel C.

The extent of unobserved heterogeneity also has different implications for the various estimators. Panel D of Figure 1 shows the OLS, AdjY, and AvgE estimates as a function of σ_{f} / σ_{X}, the relative variation of the unobserved group heterogeneity to that of the independent variable. The AdjY estimate is unaffected by the extent of unobserved heterogeneity, whereas the magnitude of the bias increases in σ_{f} / σ_{X} for both OLS and AvgE, albeit in opposite directions.

Lastly, a reduction in relative noise, $\sigma_{\varepsilon} / \sigma_{X}$, or an increase in the number of observations per group, *J*, does not necessarily improve the performance of the various estimators. As shown in Panels E and F of Figure 1, the OLS and AdjY biases are unaffected by $\sigma_{\varepsilon} / \sigma_{X}$ and *J* (when demeaning the dependent variable after excluding the observation at hand). Relative noise and the number of observations per group affect AvgE's bias, but the estimate asymptotes to different biased coefficients, which under these parameters have *opposite* signs, for small and large values of $\sigma_{\varepsilon} / \sigma_{X}$ (see Panel E). Moreover, under these parameters, an increase in number of observations per group actually *increases* the bias of AvgE (see Panel F).

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 under most parameters.

2. Fixed Effects Estimation and Additional 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 methods that are 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 the unobserved heterogeneity in Equation (11), 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 heterogeneity, f_i, in Equation (11). The FE estimate is also consistent when $\sigma_{Xf} = 0$ and the original OLS estimate is consistent.

Equivalently, the FE estimator can be implemented by transforming the data to remove the unobserved heterogeneity. This transformation 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

$$y_{i,j} - \bar{y}_i = \beta^{FE}\left(X_{i,j} - \bar{X}_i\right) + u_{i,j}^{FE}. \tag{12}$$

Comparing the FE estimator to the true data structure (see Equation (77)), we can see that the FE estimator is consistent, given our assumption in Equation (11) that X does not covary with ε.[9]

Although the estimates are consistent, the standard errors must be appropriately adjusted to account for the reduced 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 as long as the fixed effects swept away by the within-group transformation are nested within clusters (meaning all the

[9] First differencing is another way to remove the unobserved group heterogeneity. Although both are consistent, the first-difference and fixed-effects estimators differ in their assumptions about the idiosyncratic errors. See Wooldridge (2010, p. 321) for details on the relative efficiency of the two estimators.

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

2.2 Understanding AdjY and AvgE in the context of FE

The within-group transformation of the FE estimator highlights the key problem of the AdjY and AvgE estimators: they fail to account for the relation between the group mean of the independent variable, $\bar{X}_i$, and the group mean of the independent variable, $\bar{y}_i$. Comparing the FE estimation in Equation (1212) with the true underlying structure of the demeaned dependent variable in Equation (77), we see that the FE estimator correctly controls for the independent variable mean, $\bar{X}_i$, and restricts its coefficient to equal β. The AdjY estimator, however, fails to account for $\bar{X}_i$, and the AvgE estimator makes the same mistake and also fails to restrict the coefficient on $\bar{y}_i$ to equal one (see Equation (99)).

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 are inconsistent.

[10] 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, http://www.kellogg.northwestern.edu/faculty/matsa/htm/fe.htm.

2.3 Other related estimation techniques are also inconsistent

Other variants of the AdjY and AvgE estimators are also problematic. In this section, we discuss four related estimation strategies that are inconsistent and suggest alternatives.

2.3.1 Other dependent variable adjustments or controls. Estimating variants of AdjY or AvgE that substitute other group summary statistics for the group mean similarly results in inconsistent estimates. For example, some papers adjust the dependent variable in AdjY-type estimation by subtracting a group median or subtracting a value-weighted group mean. These estimates suffer an omitted variable bias because they fail to account for how the corresponding group median or value-weighted group mean of the independent variables affects the adjusted dependent variable. Likewise, using a group median or a value-weighted group mean in AvgE-type estimation is problematic because both measure the unobserved heterogeneity with error. In both cases, the fixed effects estimator is the proper way to control for unobserved heterogeneity.

The omitted variable problem underlying AdjY also applies to other types of estimations. Any dependent variable that is constructed from multiple observations without similarly adjusting the independent variables is problematic, even when this is not done to control for unobserved heterogeneity. For example, consider the case where a linear combination of other observations is used to adjust the dependent variable. Failing to control for the independent variables of all observations used to construct the dependent variable can cause an omitted variable bias.

Adjusting a dependent variable of interest by first regressing it onto other covariates and then using the residuals as the adjusted dependent variable is similarly problematic. As discussed in Section 2.2, this approach is *not* equivalent to controlling for the variables included as covariates in the first step regression. To control for these variables properly, a researcher must either partial them out from the independent variables or simply control for them directly in the main specification.

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 et al. (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 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.[11]

As an example, consider analyses of research and development (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 (in part 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 to R&D. 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

[11] Other papers regress adjusted stock returns on continuous variables. For example, some studies regress firms' adjusted returns on changes in their cash balances and interpret the sensitivity as the firms' internal value of cash. Such AdjY estimations are also inconsistent as they fail to control for differences in cash changes across the benchmark portfolios.

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 the average effect of the independent variable on stock returns while properly controlling for the characteristics used to construct the 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 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. 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 Comparisons of adjusted variables. Even simple comparisons of an adjusted variable, such as in the time periods surrounding an event of interest, can yield incorrect inferences. In many applications, including studies of mergers and acquisitions and leveraged buyouts, researchers compare the means of an adjusted variable for firms across two time periods. The variable of interest, however, is first adjusted by subtracting the mean of some benchmark group, which sometimes consists of all firms from the same industry, all firms of similar size, and/or all firms of similar past performance. As shown in Proposition 4, this pre versus post comparison, however, does *not* reveal the true effect of the event being analyzed. This is because the comparison incorrectly removes part of the event's effect on the dependent variable when it demeans the outcome of interest using an average that includes both affected and unaffected firms.

> **Proposition 4**. A pre- versus post-event comparison of an adjusted outcome variable does not reveal the effect of the event. Specifically,
>
> $$\hat{\beta}^{AdjY} = \theta\beta ,$$
>
> where θ is the average fraction of affected firms in an affected firm's benchmark group.

The source of this bias is the same as the more general AdjY bias described previously. In regression form, the AdjY comparison suffers an omitted variable problem in that it fails to control for the share of firms in each benchmark group that are treated (see the Appendix for more details).

2.3.4 Using independent variable group averages as instrumental variables. Independent variables' group averages are also sometimes used as instrumental variables. Specifically, the researcher instruments for a potentially endogenous regressor $X_{i,j}$ using the regressor's group average, $\bar{X}_{i,-j}$, calculated excluding the observation at hand. The typical justification for such instruments is that the group average of X is correlated with $X_{i,j}$ but is not otherwise related to the dependent variable, $y_{i,j}$. For example, a researcher estimating the impact of ROA on leverage but concerned that financial constraints introduce a simultaneity bias might propose using industry ROA to instrument for firm ROA.

Using group averages of the independent variables as instrumental variables, however, leads to inconsistent estimates in the presence of unobserved group heterogeneity, as in Equation (11). The instrument violates the exclusion restriction whenever the unobserved heterogeneity, f_i, is correlated with the independent variable, $X_{i,j}$, because f_i is then necessarily also correlated with $\bar{X}_{i,-j}$. As noted earlier, such correlations are pervasive in practice. In this example, unobserved industry investment opportunities likely affect both ROA and leverage, making the proposed IV estimator inconsistent.

Unlike the other applications discussed in this section, the problem with the IV estimation cannot be solved by adding fixed effects to the estimating equation. Although fixed effects control for the unobserved heterogeneity, f_i, in the second stage estimation, the fixed effects reintroduce the endogeneity problem in the first stage estimation. Recall that the instrument, $\bar{X}_{i,-j}$, is just the group mean excluding the observation at hand. After controlling for industry fixed effects, the instrument becomes $\bar{X}_{i,-j} - \bar{X}_i$ which is perfectly correlated with the endogenous regressor, $X_{i,j}$. Put differently, the instrument exploits strictly industry-level variation, which is not well-identified in the presence of industry fixed effects.

For a group average instrument to be valid, the independent variable, $X_{i,j}$, must be correlated with its group mean and the underlying economic source of this correlation must be unrelated to f_i (the part of

the industry variation that affects y_i). Although it is possible that there exist scenarios where these conditions hold, examples are rare. Researchers should not assume these conditions hold absent a strong economic justification.

2.4 AdjY and AvgE are also inconsistent under other data structures

Even when a researcher is interested in studying the determinants of $y_{i,j} - \overline{y}_i$ rather than of $y_{i,j}$, AdjY provides inconsistent estimates. Regressing $y_{i,j} - \overline{y}_i$ on $X_{i,j}$ suffers an omitted variable problem whenever $X_{i,j}$ affects $y_{i,j}$ (i.e., whenever $\beta \neq 0$). If $X_{i,j}$ affects $y_{i,j}$, then $X_{i,-j}$ also affects $y_{i,-j}$, and hence, $\overline{X}_i$ affects $y_{i,j} - \overline{y}_i$. Put differently, it is impossible for $X_{i,j}$, but not $\overline{X}_i$, to affect, $y_{i,j} - \overline{y}_i$. Thus, AdjY not only gives inconsistent estimates of the effect of $X_{i,j}$ on $y_{i,j}$ but also gives inconsistent estimates of the effect of $X_{i,j}$ on $y_{i,j} - \overline{y}_i$.

Furthermore, even if the true underlying data structure did not exhibit unobserved group heterogeneity and exactly matched the AvgE specification, both estimators would also still be biased. Suppose the underlying data exhibit the following structure:

$$y_{i,j} = \beta X_{i,j} + \delta \overline{y}_i + \varepsilon_{i,j}. \tag{13}$$

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 (i.e., AvgE) does not reveal causal effects because of a reflection problem. Because the dependent variable affects the dependent variable of other group members, using the group mean as an independent variable in OLS introduces an endogeneity problem. OLS estimation of Equation (22), AdjY, and FE would yield inconsistent estimates of the peer effects model.

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, AdjY, and 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:

$$(D / A)_{i,t} = \alpha + \boldsymbol{\beta}' \mathbf{X}_{i,t} + \varepsilon_{i,t}, \tag{14}$$

where $(D/A)_{i,t}$ is book leverage (i.e., total debt divided by total assets) for firm i in year t, and $\mathbf{X}_{i,t}$ is a vector of variables thought to affect leverage. We use data for the period 1950–2010 from Compustat, and we include five independent variables in $\mathbf{X}_{i,t}$: *fixed assets/total assets*, *Ln(sales)*, *return on assets*, *modified Altman-Z score*, and *market-to-book ratio*. All of the variables are winsorized at their 1% tails, and the standard errors are adjusted for clustering at the firm level. We estimate the model using the four different methods analyzed here—OLS, AdjY, AvgE, and FE—to account for unobserved heterogeneity across firms. The estimates are reported in Table 2.

The various estimation techniques lead to very different estimates for $\boldsymbol{\beta}$, confirming the importance of 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 (Column (3)) 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 the firms' annual independent variables are correlated with the firms' time-series average levels. The sources of these correlations are intuitive; for example, *Ln*(*sales*) is serially correlated (i.e., $\sigma_{w_{i,j} w_{i,-j}} \neq 0$) and its average assuredly differs across firms (i.e., $\sigma^2_{\bar{x}} \neq 0$). A similar concern also applies to the other independent variables. These correlations between observations of independent variables and their firm-level averages cause both AdjY and AvgE to be inconsistent.

The estimates in Table 2 illustrate that AdjY and AvgE estimators can be severely biased and lead to 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 estimates are smaller in magnitude than the FE estimate by about 40%. A

researcher using either AdjY or AvgE might therefore infer that the roles of collateral and bankruptcy risk on leverage are considerably smaller than they are. AdjY and AvgE can even yield estimates that have 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, which is almost twice the OLS estimate, suggests that the bias actually goes 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 AdjY and AvgE and their dependence on so many correlation parameters make it difficult to predict 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:

$$Ln(total\ compensation)_{i,j,t} = \alpha + \boldsymbol{\beta}' \mathbf{X}_{i,j,t} + \delta_t + \varepsilon_{i,j,t}, \tag{15}$$

where *Ln(total compensation)*$_{i,j,t}$ is the natural log of the total compensation for manager *i* at firm *j* and in year *t*, $\mathbf{X}_{i,t}$ is a vector of variables thought to affect compensation, and δ_t is a year fixed effect. Using data for the period 1992–2010 from Execucomp, Compustat, and CRSP, we estimate the model using each of the different estimators—OLS, AdjY, AvgE, and FE—to account for unobserved heterogeneity across managers. We include ten commonly included independent variables in our regressions: *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 $\boldsymbol{\beta}$. This is not surprising. Annual values of managers' covariates are correlated with their average value, and unobserved factors, such as managerial ability, are likely correlated with both the dependent variable

(compensation) and independent variables (like profitability). These correlations and unobserved factors cause the AdjY and AvgE estimates to differ considerably from the FE estimates. As shown in Table 3, a researcher using either AdjY or AvgE might incorrectly infer that ROA has no effect on total compensation when the opposite may be true. AdjY and AvgE estimates are also considerably smaller than both the OLS and FE estimates for *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:

$$Q_{i,j,t} = \alpha + \boldsymbol{\beta}' \mathbf{X}_{i,j,t} + \delta_t + \varepsilon_{i,j,t}, \tag{16}$$

where $Q_{i,j,t}$ is Tobin's Q for firm i in 4-digit SIC industry j and year t; $\mathbf{X}_{i,t}$ is a vector of variables thought to affect firm value; and δ_t is a year fixed effect. Using Compustat data from the period 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 for being incorporated in Delaware, *Ln*(*sales*), *R&D expenses/assets*, and *return on assets*. All of the variables are winsorized at their 1% tails, 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 $\boldsymbol{\beta}$. These differences indicate the importance of industry-year factors and correlations that cause OLS, AdjY, and AvgE to provide inconsistent estimates. For example, common industry shocks lead the return on assets for a given firm to be correlated with the firm's industry average return on assets. Such a correlation biases the AdjY and AvgE estimates. Because these shocks also affect unobservable factors, such as investment opportunities, that are correlated with both the dependent variable (Tobin's Q) and independent variables (like R&D expenses), OLS and AvgE are biased further.

It is apparent that analyzing industry-adjusted data rather than using industry-year fixed effects 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, their average return on assets, sales, and other characteristics vary across industries; thus the estimates will be biased. 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 firm size

In a final example, we examine the relationship between firms' R&D expenses and their stock returns. Using CRSP and Compustat data from the period 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 (bps) 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 48×5=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 Panel B of Table 5, also show a positive relationship between R&D and stock returns. Specifically, firms in the highest R&D quintile have an average adjusted return that is 530 bps higher than the average adjusted return of firms in the lowest R&D quintile.

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

$$r_{i,j,s,t} = \boldsymbol{\beta}'\mathbf{R\&D}_{i,j,s,t} + \varepsilon_{i,j,s,t}, \tag{17}$$

where $r_{i,j,s,t}$ is the stock return for firm i in 48 Fama-French industry j, size quintile s, and year t; and $\mathbf{R\&D}_{i,j,s,t}$ is a vector of indicators for the 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 the use of 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 AvgE or FE estimation. Market-weighted OLS, AdjY, AvgE, and FE estimates of Equation (1717) are reported in Columns (1)–(4) of Table 5, Panel C.

The OLS and 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 1,430 – 680 = 750 bps. Similarly, the 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 AdjY estimation.

As in the other applications, the various techniques lead to different estimates. In this case, correlations between R&D, industry, and firm size cause OLS, AdjY, and AvgE to provide inconsistent estimates for **β,** the relation between R&D and returns. A researcher relying on characteristic adjustments (AdjY) would conclude that being in the highest quintile of R&D increases stock returns over firms in the lowest quintile by 530 bps (Column (2)). The FE estimation, however, shows that the effect of R&D quintiles is actually almost twice as large at 940 bps (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's average varies across 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 AdjY, AvgE, and FE in each of the four different examples suggest that researchers' choice of which approach to use is important in practice. As shown above, correlations present in many common finance settings cause both AdjY and AvgE estimates to be inconsistent 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 heterogeneity and should be used instead.

The FE estimator, however, has limitations. One limitation (which also applies to 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 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 heterogeneity that could cause an omitted variable bias. If the factor is observable, then the researcher should control for the potential omitted variable directly using OLS. If the heterogeneity is unobservable but uncorrelated with the independent variable of interest, then OLS without the group indicator variables is 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 heterogeneity.

The remaining two conditions refer to potential limitations of FE and its ability to facilitate accurate inference. 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 attenuation bias that can occur in FE estimation 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 by 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, 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 potential 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 in FE estimation. This increase in noise occurs when there is 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 bias coefficient estimates and lead the researcher to make incorrect inferences.

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 if the FE estimation indicates that the regulatory change had little or no effect on low-income borrowing, the opposite may be true.[12]

In practice, addressing this bias is difficult. One approach is to exploit information contained in higher-order moments of the observed variables in a generalized method of moments (GMM) estimation (Erickson and Whited, 2000). Another approach is to identify an instrumental variable for within-group variation in the independent variable and implement IV methods (Biorn 2000). In panel data, a third approach is to recover the true β from the biased coefficients obtained from OLS estimation on different transformations of the data, such as first differences or within-group transformation; the different transformations affect the estimate in predictable ways that can be used to analytically solve for the bias and recover the true parameter (Griliches and Hausman 1986; McKinnish 2008). However, these approaches can be difficult to implement because they require assumptions about the exact source and structure of the measurement error process.

[12] See Griliches and Hausman (1986) and Angrist and Pischke (2009, pp. 225–7) for a more detailed discussion of the measurement error problems associated with FE estimators.

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 computing resources. Overcoming these difficulties, in fact, provides a potential motivation for using AdjY or AvgE rather than FE.[13]

When there is just one type of unobserved group effect, 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 are 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.[14]

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

[13] 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*, or *Review of Financial Studies* between 2008 and 2010 and that use either AdjY or AvgE estimation, there were only three articles in which computational problems might 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.

[14] 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 firms' averages 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, or personality) and unobserved heterogeneity across firms (such as firm culture or organization capital) that might also correlate with variables of interest (such as size, profitability, or CEO age). 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 (Graham, Li, and Qui 2012; Coles and Li 2011a). 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 2006). If not all firms are affected by the regulation equally, then including location-by-time fixed effects can control for such characteristics.

Fortunately, there are 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. Both of these approaches have 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 contain only 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. Firm-manager matched data, where managers work at multiple firms over their careers, is an example in which interacted fixed effects are feasible. An interacted manager-firm fixed effect can be used to remove unobserved heterogeneity both across managers and across firms (e.g., Coles and Li 2011a,b; Graham, Li, and Qui 2012). However, even when some variation remains after transforming the data, caution is warranted; the estimates may suffer an attenuation bias if there is any measurement error.

A second limitation of interacted fixed effects estimation is that it does not allow the researcher to recover the *un*interacted 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); and 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 procedures 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 Stata program, FELSDVREG, that applies this estimation method to models 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 (e.g., see Smyth 1996; Guimarães and Portugal 2010). 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 FE capital structure regression in Stata using the standard algorithm and the alternative approaches. Specifically, we estimated the following regression:

$$(D / A)_{i,j,t} = \alpha + \boldsymbol{\beta}\mathbf{X}_{\mathbf{i,j,t}} + f_i + \delta_{j,t} + u_{i,j,t}, \tag{18}$$

where $\mathbf{X}_{\mathbf{i,j,t}}$ is a vector of five time-varying independent variables (i.e., $K = 5$) for firm i in 4-digit SIC industry j and year t. The independent variables are *fixed assets/total assets*, *Ln*(*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, $\delta_{j,t}$. We obtain a sample of firms from Compustat for the period 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.

The memory-saving and iteration techniques discussed above 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 adjusted for clustering at the firm level on the authors' desktop 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.[15]

5. Conclusion

As empirical researchers, it is well understood that we must address unobserved heterogeneity if we hope to infer causal relations from the 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 that affect groups of observations. One approach subtracts the mean of the

[15] 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, http://www.kellogg.northwestern.edu/faculty/matsa/htm/fe.htm, for further details on how to implement this adjustment and the methods discussed in this section.

group from the dependent variable; another attempts to control for unobserved heterogeneity using the mean of the group's dependent variable. 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 heterogeneity while the other widely used approaches yield inconsistent estimates. Demeaning the dependent variable with respect to groups suffers an omitted variable problem by failing to control for how the mean of the independent variables affects the demeaned dependent variable. Using the group mean of the dependent variable as a control suffers from measurement error bias, as the sample mean is only a noisy measure of the unobserved factor.

The difference between the various approaches is important in practice because 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 can be severe in practice. Compared to the FE estimates, the alternative approaches result in different estimates and occasionally even return statistically significant estimates of the opposing sign.

Although we show that fixed effects estimation is 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 attenuation biases when the independent variable is measured with error. Given these limitations, we describe 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 methods to estimate the fixed effects model when standard approaches are computationally infeasible. These new methods are likely to be of increasing importance and of practical use to empirical researchers.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67:251–333.

Abowd, J. M., R. H. Creecy, and F. Kramarz. 2002. Computing person and firm effects using linked longitudinal employer-employee data. Longitudinal Employer – Household Dynamics, Technical Paper No. TP-2002-06.

Angrist, J. D., and J. Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, Princeton.

Arellano, M. 1987. Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics* 49:431–4.

Baltagi, B. H. 1995. *Econometric analysis of panel data*. John Wiley & Sons, Ltd., Hoboken.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119:249–75.

Biorn, E. 2000. Panel data with measurement errors: Instrumental variables and GMM procedures combining levels and differences. *Econometric Reviews* 19:391-424.

Cetorelli, N., and P. E. Strahan. 2006. Finance as a barrier to entry: Bank competition and industry structure in local U.S. markets. *Journal of Finance* 61:437–61.

Coles, J. L., and Z. F. Li. 2011a. Managerial attributes, incentives, and performance. Working Paper, Arizona State University.

———. 2011b. An empirical assessment of empirical corporate finance. Working Paper, Arizona State University.

Cornelissen, T. 2008. The Stata command felsdvreg to fit a linear model with two high-dimensional fixed effects. *Stata Journal* 8:170–89.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52:1035–58.

Erickson, T. 1993. Restricting regression slopes in the errors-in-variables model by bounding the error correlation. *Econometrica* 61:959–69.

Erickson, T., and T. M. Whited. 2000. Measurement error and the relationship between investment and q. *Journal of Political Economy* 108:1027-57.

Erickson, T., and T. M. Whited. 2005. Proxy-quality thresholds: Theory and applications. *Finance Research Letters* 2:131–51.

Erickson, T., and T. M. Whited. 2012. Treating measurement error in Tobin's Q. *Review of Financial Studies*, 25:1286–1329.

Fee, C. E., C. J. Hadlock, and J. R. Pierce. 2011. Managers who lack style: Evidence from exogenous CEO changes. Working Paper, Michigan State University.

Greene, W. H. 2000. *Econometric analysis*. Prentice-Hall, Inc., Upper Saddle River.

Graham, J. R., S. Li, and J. Qiu. 2012. Managerial attributes and executive compensation. *Review of Financial Studies* 25:144–86.

Griliches, Z., and J. A. Hausman. 1986. Errors in variables in panel data. *Journal of Econometrics* 31:93–118.

Guimarães, P., and P. Portugal. 2010. A simple feasible procedure to fit models with high-dimensional fixed effects. *Stata Journal* 10:628–49.

Hausman, J. A., and W. E. Taylor. 1981. Panel data and unobservable individual effects. *Econometrica* 49:1377–98.

Hu, Y. 2006. Bounding parameters in a linear regression model with a mismeasured regressor using additional information. *Journal of Econometrics* 133: 51–70.

Krasker, W.S., and J.W. Pratt. 1986. Bounding the effects of proxy variables on regression coefficients. *Econometrica* 54:641–55.

Leary, M. T., and M. R. Roberts. 2010. Do peer firms affect corporate financial policy? Working Paper, Washington University in St. Louis.

Manski, C. 1993. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60:531–42.

Matsa, D. A. 2010. Capital structure as a strategic variable: Evidence from collective bargaining. *Journal of Finance* 65:1197–232.

McKinnish, T. 2008. Panel data models and transitory fluctuations in the explanatory variable In *Modeling and evaluating treatment effects in econometrics*, eds. Daniel L. Millimet, Jeffrey A. Smith, and Edward J. Vytlacil, 335–58. Elsevier, Amsterdam.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–80.

Smyth, G. K. 1996. Partitioned algorithms for maximum likelihood and other non-linear wstimation. *Statistics and Computing* 6:201–16.

Stock, J. H., and M. W. Watson. 2008. Heteroskedasticity-robust standard errors for fixed effects panel data regression. *Econometrica* 76:155–74.

Wansbeek, T., and A. Kapteyn. 1989. Estimation of the error-components model with incomplete panels. *Journal of Econometrics* 41:341–61.

Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*. MIT Press, Cambridge.

Appendix

A1. Proof of Proposition 1

Given $\mu_X = \mu_f = 0$ and Equation (1), when the number of groups, N, goes to infinity and the number of observations per group, J, remains constant, the AdjY estimate, $\hat{\beta}^{AdjY}$, is given by:

$$\underset{N\to\infty}{plim}\ \hat{\beta}^{AdjY} = \underset{N\to\infty}{plim}\left(\frac{1}{N}\mathbf{X'X}\right)^{-1}\frac{1}{N}\mathbf{X'y}, \tag{A1}$$

where

$$\mathbf{X'} = \begin{pmatrix} X_{1,1} & \cdots & X_{1,J} & X_{2,1} & \cdots & X_{N,J} \end{pmatrix}$$
$$\mathbf{y} = \begin{pmatrix} y_{1,1}-\overline{y}_1 \\ \vdots \\ y_{1,J}-\overline{y}_1 \\ y_{2,1}-\overline{y}_2 \\ \vdots \\ y_{N,J}-\overline{y}_N \end{pmatrix} = \begin{pmatrix} \beta X_{1,1}-\beta\overline{X}_1+\varepsilon_{1,1}-\overline{\varepsilon}_1 \\ \vdots \\ \beta X_{1,J}-\beta\overline{X}_1+\varepsilon_{1,J}-\overline{\varepsilon}_1 \\ \beta X_{2,1}-\beta\overline{X}_2+\varepsilon_{2,1}-\overline{\varepsilon}_2 \\ \vdots \\ \beta X_{N,J}-\beta\overline{X}_N+\varepsilon_{N,J}-\overline{\varepsilon}_N \end{pmatrix}. \tag{A2}$$

Because groups are *i.i.d.*, it can then be shown that:

$$\underset{N\to\infty}{plim}\left(\frac{1}{N}\mathbf{X'X}\right)^{-1} = \underset{N\to\infty}{plim}\left[\frac{\begin{pmatrix} X_{1,1} & \cdots & X_{1,J} & X_{2,1} & \cdots & X_{N,J} \end{pmatrix}\begin{pmatrix} X_{1,1} \\ \vdots \\ X_{1,J} \\ X_{2,1} \\ \vdots \\ X_{N,J} \end{pmatrix}}{N}\right]^{-1} = \left(J\sigma_X^2\right)^{-1} = \frac{1}{J\sigma_X^2} \tag{A3}$$

and

$$\underset{N\to\infty}{plim}\frac{1}{N}\mathbf{X'y}=\left[\frac{\begin{pmatrix}X_{1,1} & \cdots & X_{1,J} & X_{2,1} & \cdots & X_{N,J}\end{pmatrix}\begin{pmatrix}\beta X_{1,1}-\beta\bar{X}_1+\varepsilon_{1,1}-\bar{\varepsilon}_1\\ \vdots\\ \beta X_{1,J}-\beta\bar{X}_1+\varepsilon_{1,J}-\bar{\varepsilon}_1\\ \beta X_{2,1}-\beta\bar{X}_2+\varepsilon_{2,1}-\bar{\varepsilon}_2\\ \vdots\\ \beta X_{N,J}-\beta\bar{X}_N+\varepsilon_{N,J}-\bar{\varepsilon}_N\end{pmatrix}}{N}\right]=J\left(\beta\sigma_X^2-\beta\sigma_{X\bar{X}}\right), \quad \text{(A4)}$$

where $\sigma_{X\bar{X}}$ is the covariance between $X_{i,j}$ and the sample group average, $\bar{X}_i$. Therefore,

$$\underset{N\to\infty}{plim}\ \hat{\beta}^{AdjY}=\left(\frac{1}{\sigma_X^2}\right)\left(\beta\sigma_X^2-\beta\sigma_{X\bar{X}}\right)=\beta-\beta\frac{\sigma_{X\bar{X}}}{\sigma_X^2}. \quad \text{(A5)}$$

■

A2. Proof of Proposition 2

Given $\mu_X = \mu_Z = \mu_f = 0$ and Equation (8), when the number of groups, *N,* goes to infinity and the number of observations per group, *J*, remains constant, the AdjY estimates $\hat{\beta}^{AdjY}$ and $\hat{\gamma}^{AdjY}$ are given by:

$$\underset{N\to\infty}{plim}\ \mathbf{b}=\underset{N\to\infty}{plim}\left(\frac{1}{N}\mathbf{X'X}\right)^{-1}\frac{1}{N}\mathbf{X'y}, \quad \text{(A6)}$$

where

$$\mathbf{b}=\begin{bmatrix}\hat{\beta}^{AdjY}\\ \hat{\gamma}^{AdjY}\end{bmatrix}$$

$$\mathbf{X'}=\begin{pmatrix}X_{1,1} & \cdots & X_{1,J} & X_{2,1} & \cdots & X_{N,J}\\ Z_{1,1} & \cdots & Z_{1,J} & Z_{2,1} & \cdots & Z_{N,J}\end{pmatrix}, \quad \text{(A7)}$$

$$\mathbf{y}=\begin{pmatrix}y_{1,1}-\bar{y}_1\\ \vdots\\ y_{1,J}-\bar{y}_1\\ y_{2,1}-\bar{y}_2\\ \vdots\\ y_{N,J}-\bar{y}_N\end{pmatrix}=\begin{pmatrix}\beta X_{1,1}+\gamma Z_{1,1}-\beta\bar{X}_1-\gamma\bar{Z}_1+\varepsilon_{1,1}-\bar{\varepsilon}_1\\ \vdots\\ \beta X_{1,J}+\gamma Z_{1,J}-\beta\bar{X}_1-\gamma\bar{Z}_1+\varepsilon_{1,J}-\bar{\varepsilon}_1\\ \beta X_{2,1}+\gamma Z_{2,1}-\beta\bar{X}_2-\gamma\bar{Z}_2+\varepsilon_{2,1}-\bar{\varepsilon}_2\\ \vdots\\ \beta X_{N,J}+\gamma Z_{N,J}-\beta\bar{X}_N-\gamma\bar{Z}_N+\varepsilon_{N,J}-\bar{\varepsilon}_N\end{pmatrix}$$

and

$$\begin{aligned} \bar{y}_i &= f_i + \beta \bar{X}_i + \gamma \bar{Z}_i + \bar{\varepsilon}_i \\ \bar{Z}_i &= \frac{1}{J} \sum_{k \in group\ i} Z_{i,k} \end{aligned} \quad . \tag{A8}$$

Because groups are *i.i.d.*, it can be shown that:

$$\begin{aligned} \underset{N\to\infty}{plim} \left(\frac{1}{N} \mathbf{X'X} \right)^{-1} &= \underset{N\to\infty}{plim} \left[\frac{\begin{pmatrix} X_{1,1} & \cdots & X_{1,J} & X_{2,1} & \cdots & X_{N,J} \\ Z_{1,1} & \cdots & Z_{1,J} & Z_{2,1} & \cdots & Z_{N,J} \end{pmatrix} \begin{pmatrix} X_{1,1} & Z_{1,1} \\ \vdots & \vdots \\ X_{1,J} & Z_{1,J} \\ X_{2,1} & Z_{2,1} \\ \vdots & \vdots \\ X_{N,J} & Z_{N,J} \end{pmatrix}}{N} \right]^{-1} \\ &= \begin{pmatrix} J\sigma_X^2 & J\sigma_{XZ} \\ J\sigma_{XZ} & J\sigma_Z^2 \end{pmatrix}^{-1} = \left(\frac{1}{J} \right) \begin{pmatrix} \dfrac{\sigma_Z^2}{\sigma_Z^2\sigma_X^2 - \sigma_{XZ}^2} & -\dfrac{\sigma_{XZ}}{\sigma_Z^2\sigma_X^2 - \sigma_{XZ}^2} \\ -\dfrac{\sigma_{XZ}}{\sigma_Z^2\sigma_X^2 - \sigma_{XZ}^2} & \dfrac{\sigma_X^2}{\sigma_Z^2\sigma_X^2 - \sigma_{XZ}^2} \end{pmatrix}, \end{aligned} \tag{A9}$$

where σ_{XZ} is the covariance between $X_{i,j}$ and $Z_{i,j}$. It can also be shown that:

$$\begin{aligned} \underset{N\to\infty}{plim} \frac{1}{N} \mathbf{X'y} &= \left[\frac{\begin{pmatrix} X_{1,1} & \cdots & X_{1,J} & X_{2,1} & \cdots & X_{N,J} \\ Z_{1,1} & \cdots & Z_{1,J} & Z_{2,1} & \cdots & Z_{N,J} \end{pmatrix} \begin{pmatrix} \beta X_{1,1} + \gamma Z_{1,1} - \beta \bar{X}_1 - \gamma \bar{Z}_1 + \varepsilon_{1,1} - \bar{\varepsilon}_1 \\ \vdots \\ \beta X_{1,J} + \gamma Z_{1,J} - \beta \bar{X}_1 - \gamma \bar{Z}_1 + \varepsilon_{1,J} - \bar{\varepsilon}_1 \\ \beta X_{2,1} + \gamma Z_{2,1} - \beta \bar{X}_2 - \gamma \bar{Z}_2 + \varepsilon_{2,1} - \bar{\varepsilon}_2 \\ \vdots \\ \beta X_{N,J} + \gamma Z_{N,J} - \beta \bar{X}_N - \gamma \bar{Z}_N + \varepsilon_{N,J} - \bar{\varepsilon}_N \end{pmatrix}}{N} \right] \\ &= J \begin{pmatrix} \beta\sigma_X^2 + \gamma\sigma_{XZ} - \beta\sigma_{X\bar{X}} - \gamma\sigma_{X\bar{Z}} \\ \beta\sigma_{XZ} + \gamma\sigma_Z^2 - \beta\sigma_{Z\bar{X}} - \gamma\sigma_{Z\bar{Z}} \end{pmatrix}, \end{aligned} \tag{A10}$$

where $\sigma_{Z\bar{Z}}$ is the covariance between $Z_{i,j}$ and $\bar{Z}_i$, $\sigma_{Z\bar{X}}$ is the covariance between $Z_{i,j}$ and $\bar{X}_i$, and $\sigma_{X\bar{Z}}$ is the covariance between $X_{i,j}$ and $\bar{Z}_i$.

Therefore,

$$\underset{N\to\infty}{plim}\begin{bmatrix}\hat{\beta}^{AdjY}\\ \hat{\gamma}^{AdjY}\end{bmatrix}=\begin{pmatrix}\dfrac{\sigma_Z^2}{\sigma_Z^2\sigma_X^2-\sigma_{XZ}^2} & -\dfrac{\sigma_{XZ}}{\sigma_Z^2\sigma_X^2-\sigma_{XZ}^2}\\ -\dfrac{\sigma_{XZ}}{\sigma_Z^2\sigma_X^2-\sigma_{XZ}^2} & \dfrac{\sigma_X^2}{\sigma_Z^2\sigma_X^2-\sigma_{XZ}^2}\end{pmatrix}\begin{pmatrix}\beta\sigma_X^2+\gamma\sigma_{XZ}-\beta\sigma_{X\bar{X}}-\gamma\sigma_{X\bar{Z}}\\ \beta\sigma_{XZ}+\gamma\sigma_Z^2-\beta\sigma_{Z\bar{X}}-\gamma\sigma_{Z\bar{Z}}\end{pmatrix}$$
$$=\begin{bmatrix}\beta+\dfrac{\beta\left(\sigma_{XZ}\sigma_{Z\bar{X}}-\sigma_Z^2\sigma_{X\bar{X}}\right)+\gamma\left(\sigma_{XZ}\sigma_{Z\bar{Z}}-\sigma_Z^2\sigma_{X\bar{Z}}\right)}{\sigma_Z^2\sigma_X^2-\sigma_{XZ}^2}\\ \gamma+\dfrac{\beta\left(\sigma_{XZ}\sigma_{X\bar{X}}-\sigma_X^2\sigma_{Z\bar{X}}\right)+\gamma\left(\sigma_{XZ}\sigma_{X\bar{Z}}-\sigma_X^2\sigma_{Z\bar{Z}}\right)}{\sigma_Z^2\sigma_X^2-\sigma_{XZ}^2}\end{bmatrix}. \tag{A11}$$

■

A3. Proof of Proposition 3

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

$$\underset{N\to\infty}{plim}\ \mathbf{b} = \underset{N\to\infty}{plim}\left(\frac{1}{N}\mathbf{X'X}\right)^{-1}\frac{1}{N}\mathbf{X'y}, \tag{A12}$$

where

$$\mathbf{b}=\begin{bmatrix}\hat{\beta}^{AvgE}\\ \hat{\lambda}^{AvgE}\end{bmatrix}$$
$$\mathbf{X'}=\begin{pmatrix}X_{1,1} & \cdots & X_{1,J} & X_{2,1} & \cdots & X_{N,J}\\ \bar{y}_1 & \cdots & \bar{y}_1 & \bar{y}_2 & \cdots & \bar{y}_N\end{pmatrix}. \tag{A13}$$
$$\mathbf{y}=\begin{pmatrix}y_{1,1}\\ \vdots\\ y_{1,J}\\ y_{2,1}\\ \vdots\\ y_{N,J}\end{pmatrix}=\begin{pmatrix}\beta X_{1,1}+f_1+\varepsilon_{1,1}\\ \vdots\\ \beta X_{1,J}+f_1+\varepsilon_{1,J}\\ \beta X_{2,1}+f_2+\varepsilon_{2,1}\\ \vdots\\ \beta X_{N,J}+f_N+\varepsilon_{N,J}\end{pmatrix}$$

Because groups are *i.i.d.*, it can then be shown that:

$$\underset{N\to\infty}{plim}\left(\frac{1}{N}\mathbf{X'X}\right)^{-1}=\underset{N\to\infty}{plim}\left[\frac{\begin{pmatrix}X_{1,1} & \cdots & X_{N,J}\\ f_1+\beta\bar{X}_1+\bar{\varepsilon}_1 & \cdots & f_N+\beta\bar{X}_N+\bar{\varepsilon}_N\end{pmatrix}\begin{pmatrix}X_{1,1} & f_1+\beta\bar{X}_1+\bar{\varepsilon}_1\\ \vdots & \vdots\\ X_{1,J} & f_1+\beta\bar{X}_1+\bar{\varepsilon}_1\\ X_{2,1} & f_2+\beta\bar{X}_2+\bar{\varepsilon}_2\\ \vdots & \vdots\\ X_{N,J} & f_N+\beta\bar{X}_N+\bar{\varepsilon}_N\end{pmatrix}}{N}\right]^{-1}$$
$$=\left(\frac{1}{J}\right)\begin{pmatrix}\sigma_X^2 & \sigma_{Xf}+\beta\sigma_{X\bar{X}}\\ \sigma_{Xf}+\beta\sigma_{X\bar{X}} & \sigma_f^2+2\beta\sigma_{f\bar{X}}+\beta^2\sigma_{\bar{X}}^2+\sigma_{\bar{\varepsilon}}^2\end{pmatrix}^{-1}, \tag{A14}$$

where $\sigma_{X\bar{X}}$ is the covariance between $X_{i,j}$ and $\bar{X}_i$, $\sigma_{f\bar{X}}$ is the covariance between f_i and $\bar{X}_i$, $\sigma_{\bar{X}}^2$ is the variance of $\bar{X}_i$, and $\sigma_{\bar{\varepsilon}}^2$ is the variance of $\bar{\varepsilon}_i$.

And, it can be shown that:

$$\underset{N\to\infty}{plim}\frac{1}{N}\mathbf{X'y}=\left[\frac{\begin{pmatrix}X_{1,1} & \cdots & X_{N,J}\\ f_1+\beta\bar{X}_1+\bar{\varepsilon}_1 & \cdots & f_N+\beta\bar{X}_N+\bar{\varepsilon}_N\end{pmatrix}\begin{pmatrix}\beta X_{1,1}+f_1+\varepsilon_{1,1}\\ \vdots\\ \beta X_{1,J}+f_1+\varepsilon_{1,J}\\ \beta X_{2,1}+f_2+\varepsilon_{2,1}\\ \vdots\\ \beta X_{N,J}+f_N+\varepsilon_{N,J}\end{pmatrix}}{N}\right]$$
$$=J\begin{pmatrix}\beta\sigma_X^2+\sigma_{Xf}\\ \beta\sigma_{Xf}+\sigma_f^2+\beta^2\sigma_{X\bar{X}}+\beta\sigma_{f\bar{X}}+\sigma_{\varepsilon\bar{\varepsilon}}\end{pmatrix}, \tag{A15}$$

where $\sigma_{\varepsilon\bar{\varepsilon}}$ is the covariance between ε_i and $\bar{\varepsilon}_i$. Therefore,

$$\underset{N\to\infty}{plim}\begin{bmatrix}\hat{\beta}^{AvgE}\\ \hat{\lambda}^{AvgE}\end{bmatrix}=\begin{pmatrix}\sigma_X^2 & \sigma_{Xf}+\beta\sigma_{X\bar{X}}\\ \sigma_{Xf}+\beta\sigma_{X\bar{X}} & \sigma_f^2+2\beta\sigma_{f\bar{X}}+\beta^2\sigma_{\bar{X}}^2+\sigma_{\bar{\varepsilon}}^2\end{pmatrix}^{-1}\begin{pmatrix}\beta\sigma_X^2+\sigma_{Xf}\\ \beta\sigma_{Xf}+\sigma_f^2+\beta^2\sigma_{X\bar{X}}+\beta\sigma_{f\bar{X}}+\sigma_{\varepsilon\bar{\varepsilon}}\end{pmatrix}$$
$$=\begin{bmatrix}\beta+\dfrac{\sigma_{Xf}\left(\beta\sigma_{f\bar{X}}+\beta^2\sigma_{\bar{X}}^2+\sigma_{\bar{\varepsilon}}^2-\sigma_{\varepsilon\bar{\varepsilon}}\right)-\beta\sigma_{X\bar{X}}\left(\sigma_f^2+\beta\sigma_{f\bar{X}}+\sigma_{\varepsilon\bar{\varepsilon}}\right)}{\sigma_X^2\left(\sigma_f^2+2\beta\sigma_{f\bar{X}}+\beta^2\sigma_{\bar{X}}^2+\sigma_{\bar{\varepsilon}}^2\right)-\left(\sigma_{Xf}+\beta\sigma_{X\bar{X}}\right)^2}\\ \dfrac{\sigma_X^2\left(\sigma_f^2+\beta\sigma_{f\bar{X}}+\sigma_{\varepsilon\bar{\varepsilon}}\right)-\sigma_{Xf}\left(\sigma_{Xf}+\beta\sigma_{X\bar{X}}\right)}{\sigma_X^2\left(\sigma_f^2+2\beta\sigma_{f\bar{X}}+\beta^2\sigma_{\bar{X}}^2+\sigma_{\bar{\varepsilon}}^2\right)-\left(\sigma_{Xf}+\beta\sigma_{X\bar{X}}\right)^2}\end{bmatrix}. \tag{A16}$$

■

A4. Determining whether the sign of $\hat{\beta}^{AvgE}$ is correct

Because AvgE suffers from measurement error bias, it is possible to use the methods of Erickson and Whited (2005) to determine whether the sign of the inconsistent estimate $\hat{\beta}^{AvgE}$ is correct. Because the AvgE measurement error, $-\beta\overline{X}_i - \overline{\varepsilon}_i$, may be correlated with both the regression disturbance term, $\varepsilon_{i,j}$, and the regressor $X_{i,j}$, the most general case of Erickson and Whited (2005) applies. Let u_i be the measurement error of the AvgE estimator and assume the following conditions hold:

(1) $\left(\varepsilon_{i,j}, u_i, f_i, X_{i,j}\right)$ are *i.i.d.* sequences,

(2) $E\left(\varepsilon_{i,j} f_i\right) = E\left(\varepsilon_{i,j} X_{i,j}\right) = 0$,

(3) the covariance matrix $Var\left(\varepsilon_{i,j}, u_i, f_i, X_{i,j}\right)$ is positive definite,

(4) every element of the coefficient vector from the projection of $y_{i,j}$ on $\left(\overline{y}_i, X_{i,j}\right)$ is positive,

(5) and the coefficient relating $\overline{y}_i$ and $X_{i,j}$ in the projection of $\overline{y}_i$ on $\left(f_i, X_{i,j}\right)$ is positive,[16]

It can then be shown that the AvgE estimate $\hat{\beta}^{AvgE}$ has the correct sign when

$$\rho_{\overline{y}f}^2 > Max\left\{\left(1+\left(\frac{\hat{\beta}^{AvgE}}{m\sqrt{\frac{\sigma_{resid_{yX}}^2}{\sigma_{resid_{\overline{y}X}}^2}-\left(\hat{\lambda}^{AvgE}\right)^2}}\right)^2\right)^{-1}, \left(1+\left(\frac{\hat{\lambda}^{AvgE}}{\sqrt{\frac{\sigma_{resid_{yX}}^2}{\sigma_{resid_{\overline{y}X}}^2}-\left(\hat{\lambda}^{AvgE}\right)^2}}\right)^2\right)^{-1}\right\}, \tag{A17}$$

where $\rho_{\overline{y}f}$ is the correlation between the residuals from the projection of $\overline{y}_i$ on $X_{i,j}$ and the residuals from the projection of f_i on $X_{i,j}$; $\hat{\beta}^{AvgE}$ and $\hat{\lambda}^{AvgE}$ are estimates of Equation (9); *m* is the slope parameter from a linear projection of $\overline{y}_i$ on $X_{i,j}$; $\sigma_{resid_{yX}}^2$ is the variance of the residual from a regression of y_i on $X_{i,j}$; and $\sigma_{resid_{\overline{y}X}}^2$ is the variance of the residuals from a regression of $\overline{y}_i$ on $X_{i,j}$. While $\rho_{\overline{y}f}$ must be assumed or obtained using additional information because f_i is unobservable, all of the other parameters in Equation (A1717) can be estimated. ■

[16] As noted in Erickson and Whited, assumptions (4) and (5) are easily satisfied by making negative coefficients positive by multiplying their regressors by –1.

A5. Proof of Equation (10)

When $\bar{X}_i$ is defined as

$$\bar{X}_i = \frac{1}{J}\sum_{k\in group\ i} X_{i,k}\ , \tag{A18}$$

where $X_{i,j} = \bar{x}_i + w_{i,j}$, $\bar{x}_i$ are *i.i.d.* with mean zero and variance $\sigma^2_{\bar{x}}$, $w_{i,j}$ are distributed with mean zero and variance σ^2_W, and cov($\bar{x}_i$, $w_{i,j}$) is equal to zero, then:

$$\sigma^2_X = Var(\bar{x}_i) + Var(w_{i,j}) = \sigma^2_{\bar{x}} + \sigma^2_w \tag{A19}$$

and

$$\begin{aligned}\operatorname{cov}\left(X_{i,j}, X_{i,-j}\right) &= \operatorname{cov}\left(\bar{x}_i + w_{i,j}, \bar{x}_i + w_{i,-j}\right)\\ &= \operatorname{var}\left(\bar{x}_i\right) + \operatorname{cov}\left(w_{i,j}, w_{i,-j}\right).\\ &= \sigma^2_{\bar{x}} + \sigma_{w_{i,j}, w_{i,-j}}\end{aligned} \tag{A20}$$

Thus,

$$\begin{aligned}\sigma_{X\bar{X}} &= \operatorname{cov}(X_{i,j}, \bar{X}_i)\\ &= \operatorname{cov}\left(X_{i,j}, \frac{1}{J}\sum_{k\in group\ i} X_{i,k}\right)\\ &= \operatorname{cov}\left(X_{i,j}, \frac{1}{J}\sum_{\substack{k\in group\ i\\ k\neq j}} X_{i,k}\right) + \frac{1}{J}\sigma^2_X\\ &= \left(\frac{J-1}{J}\right)\operatorname{cov}\left(X_{i,j}, X_{i,-j}\right) + \frac{1}{J}\sigma^2_X\\ &= \left(\frac{J-1}{J}\right)\left(\sigma^2_{\bar{x}} + \sigma_{w_{i,j}, w_{i,-j}}\right) + \frac{1}{J}\left(\sigma^2_{\bar{x}} + \sigma^2_w\right)\\ &= \sigma^2_{\bar{x}} + \left(\frac{J-1}{J}\right)\sigma_{w_{i,j}, w_{i,-j}} + \frac{\sigma^2_w}{J}\end{aligned}\ . \tag{A21}$$

■

A6. Derivation of AdjY and AvgE estimates for Table 1 and Figure 1

When the group mean, $\bar{y}_i$, and the corresponding group means of the independent variable are defined to exclude the observation at hand, as in Equation (11), errors $\varepsilon_{i,j}$ are *i.i.d.*, and $X_{i,j} = \bar{x}_i + w_{i,j}$, where the group means $\bar{x}_i$ are *i.i.d.* with mean zero and variance $\sigma_{\bar{x}}^2$, $w_{i,j}$ are distributed with mean zero and variance σ_w^2, and cov($\bar{x}_i$, $w_{i,j}$) = 0, then $\sigma_X^2 = \sigma_{\bar{x}}^2 + \sigma_w^2$ (as shown in Appendix A5), and

$$\begin{aligned}
\sigma_{X\bar{X}} &= \mathrm{cov}(X_{i,j}, \bar{X}_{i,-j}) \\
&= \mathrm{cov}\left(X_{i,j}, \frac{1}{J-1} \sum_{\substack{k \in group\ i \\ k \neq j}} X_{i,k} \right) \\
&= \frac{1}{J-1} \mathrm{cov}\left(X_{i,j}, \sum_{\substack{k \in group\ i \\ k \neq j}} X_{i,k} \right) \\
&= \sigma_{\bar{x}}^2 + \sigma_{w_{i,j}, w_{i,-j}}
\end{aligned} \tag{A22}$$

where $\sigma_{w_{i,j} w_{i,-j}}$ is the covariance between $w_{i,j}$ and $w_{i,-j}$. Likewise,

$$\sigma_{\varepsilon\bar{\varepsilon}} = \mathrm{cov}(\varepsilon_{i,j}, \bar{\varepsilon}_i) = \mathrm{cov}\left(\varepsilon_{i,j}, \frac{1}{J-1} \sum_{\substack{k \in group\ i \\ k \neq j}} \varepsilon_{i,k} \right) = \sigma_{\varepsilon_{i,j} \varepsilon_{i,-j}} = 0\,, \tag{A23}$$

$$\sigma_{f\bar{X}} = \mathrm{cov}\left(f_i, \frac{1}{J-1} \sum_{\substack{k \in group\ i \\ k \neq j}} X_{i,k} \right) = \sigma_{Xf}\,, \tag{A24}$$

$$
\begin{aligned}
\sigma^2_{\bar{X}} &= \operatorname{var}\left(\frac{1}{J-1}\sum_{\substack{k\in group\ i\\ k\neq j}} X_{i,k}\right)\\
&= \left(\frac{1}{J-1}\right)^2\left[\sum_{\substack{k\in group\ i\\ k\neq j}}\sum_{\substack{m\in group\ i\\ m\neq j}} \operatorname{cov}\left(X_{i,k}, X_{i,m}\right)\right]\\
&= \left(\frac{1}{J-1}\right)^2\left[\sum_{\substack{k\in group\ i\\ k\neq j}} \sigma^2_X + \sum_{\substack{k\in group\ i\\ k\neq j}}\sum_{\substack{m\in group\ i\\ m\neq j,k}} \operatorname{cov}\left(X_{i,k}, X_{i,m}\right)\right], \qquad \text{(A25)}\\
&= \left(\frac{1}{J-1}\right)^2\left[(J-1)\sigma^2_X + (J-1)(J-2)\operatorname{cov}\left(X_{i,j}, X_{i,-j}\right)\right]\\
&= \frac{\sigma^2_{\bar{x}} + \sigma^2_w}{J-1} + \frac{(J-2)\left(\sigma^2_{\bar{x}} + \sigma_{w_{i,j}, w_{i,-j}}\right)}{J-1}\\
&= \sigma^2_{\bar{x}} + \frac{\sigma^2_w}{J-1} + \left(\frac{J-2}{J-1}\right)\sigma_{w_{i,j}, w_{i,-j}}
\end{aligned}
$$

and

$$
\sigma^2_{\bar{\varepsilon}} = \operatorname{var}\left(\frac{1}{J-1}\sum_{\substack{k\in group\ i\\ k\neq j}} \varepsilon_{i,k}\right) = \sigma_{\varepsilon_{i,j}\varepsilon_{i,-j}} + \frac{\sigma^2_\varepsilon - \sigma_{\varepsilon_{i,j}\varepsilon_{i,-j}}}{J-1} = \frac{\sigma^2_\varepsilon}{J-1} \qquad \text{(A26)}
$$

By recognizing that

$$
\rho_{w_{i,j}w_{i,-j}} = \frac{\sigma_{w_{i,j}w_{i,-j}}}{\sigma^2_w}, \qquad \text{(A27)}
$$

one can plug Equations (A2222), (A2323), (A2424), (A2525), and (A2626) into the $\hat{\beta}^{AdjY}$ and $\hat{\beta}^{AvgE}$ from Propositions 1 and 3 to obtain:

$$
\hat{\beta}^{AdjY} = \beta\left(1 - \frac{\rho_{w_{i,j}, w_{i,-j}} + \frac{\sigma^2_{\bar{x}}}{\sigma^2_w}}{1 + \frac{\sigma^2_{\bar{x}}}{\sigma^2_w}}\right) \qquad \text{(A28)}
$$

and

$$\hat{\beta}^{AvgE} = \beta + \frac{\mathrm{A}+\mathrm{B}}{\Gamma+\Delta}, \tag{A29}$$

where

$$\mathrm{A} = \rho_{Xf}\left(\frac{\sigma_f}{\sigma_X}\right)\left(\frac{\beta^2\left(\frac{1-\rho_{w_{i,j},w_{i,-j}}}{1+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}}\right)+\frac{\sigma^2_\varepsilon}{\sigma^2_X}}{J-1}\right),$$

$$\mathrm{B} = \beta\left(\frac{\sigma^2_f}{\sigma^2_X}\right)\left(\rho^2_{Xf}-\frac{\left(\rho_{w_{i,j},w_{i,-j}}+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}\right)}{1+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}}\right), \tag{A30}$$

$$\Gamma = \frac{\left(\rho_{w_{i,j},w_{i,-j}}+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}\right)}{1+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}}\left(\beta^2\left(1-\frac{\left(\rho_{w_{i,j},w_{i,-j}}+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}\right)}{1+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}}\right)-2\beta\rho_{Xf}\left(\frac{\sigma_f}{\sigma_X}\right)\right)+\frac{\beta^2\left(\frac{1-\rho_{w_{i,j},w_{i,-j}}}{J-1}\right)}{1+\frac{\sigma^2_{\bar{x}}}{\sigma^2_w}}+\frac{\frac{\sigma^2_\varepsilon}{\sigma^2_X}}{J-1},$$

$$\Delta = \left(\frac{\sigma_f}{\sigma_X}\right)\left(\left(\frac{\sigma_f}{\sigma_X}\right)\left(1-\rho^2_{Xf}\right)+2\beta\rho_{Xf}\right).$$

■

A7. Proof of Proposition 4

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:

$$y_{i,j,t} = \beta_0 + \beta_1 P_t + \beta_2 T_{i,j} + \beta_3\left(P_t \times T_{i,j}\right) + \varepsilon_{i,j,t}, \tag{A31}$$

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

The difference-in-differences estimator, which is a direct estimation of Equation (A3131), 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 β_3, which is the causal effect of the treatment on the outcome *y*.

The AdjY-style approach to estimating β_3 instead compares the group-adjusted mean of *y* for the treated units in the pre- and post-treatment periods. The AdjY estimator is:

$$y_{i,j,t} - \overline{y}_{j,t} = \alpha^{AdjY} + \beta^{AdjY} P_t + u_{i,j,t}^{AdjY} . \tag{A32}$$

Using Equation (A3131), we can see that the group mean is given by:

$$\overline{y}_{i,t} = \beta_0 + \beta_1 P_t + \beta_2 \overline{T}_i + \beta_3 \left(P_t \times \overline{T}_i \right) + \overline{\varepsilon}_{i,t} , \tag{A33}$$

and the group-adjusted mean is given by:

$$y_{i,j,t} - \overline{y}_{i,t} = \beta_2 \left(T_{i,j} - \overline{T}_i \right) + \beta_3 \left(T_{i,j} - \overline{T}_i \right) P_t + \left(\varepsilon_{i,j,t} - \overline{\varepsilon}_{i,t} \right) . \tag{A34}$$

Thus, the group-adjusted mean for treated firms, where $T_{i,j} = 1$, is equal to:

$$y_{i,j,t} - \overline{y}_{i,t} = \beta_2 \left(1 - \overline{T}_i \right) + \beta_3 \left(1 - \overline{T}_i \right) P_t + \left(\varepsilon_{i,j,t} - \overline{\varepsilon}_{i,t} \right) . \tag{A35}$$

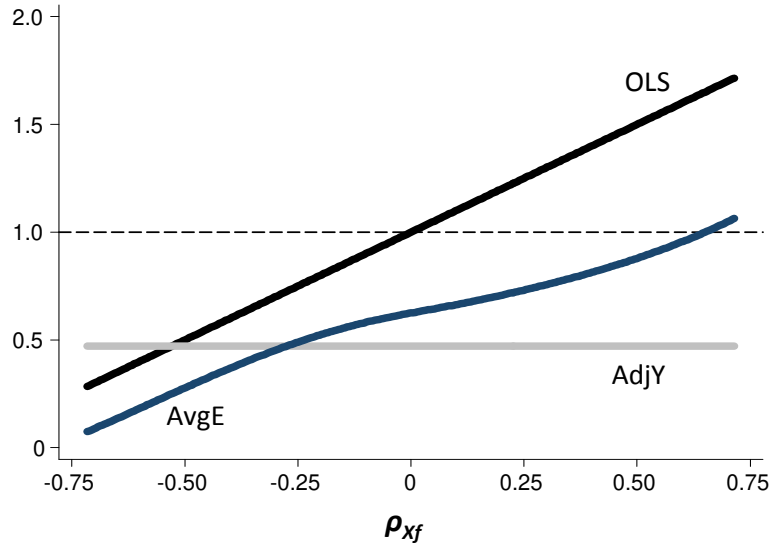
Comparing the AdjY estimator in Equation (A3232) and the true underlying data structure in Equation (A3535) reveals that the AdjY estimator is not consistent. Specifically, $\hat{\beta}^{AdjY} = \beta_3 (1 - \overline{T})$, where $\overline{T}$ is the average share of untreated firms in a treated firm's 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 (A3535) can be written as:

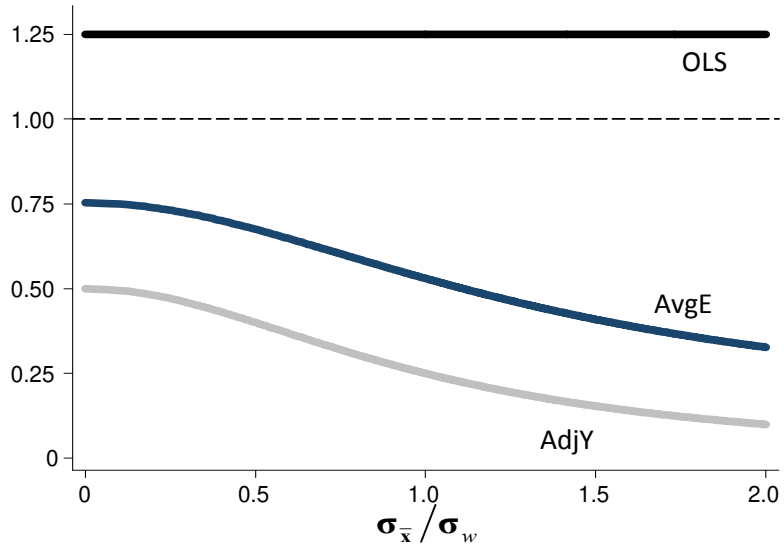
$$y_{i,j,t} - \bar{y}_{i,t} = \beta_2 \left(1 - \bar{T}_i\right) + \beta_3 P_t - \beta_3 P_t \bar{T}_i + \left(\varepsilon_{i,j,t} - \bar{\varepsilon}_{i,t}\right). \quad \text{(A36)}$$

Comparing (A3636) to Equation (A3232), the AdjY estimator fails to control for $P_t \bar{T}_j$, which is correlated with the independent variable of interest, P_t. Biased estimates for the causal effect, β_3, result. ■

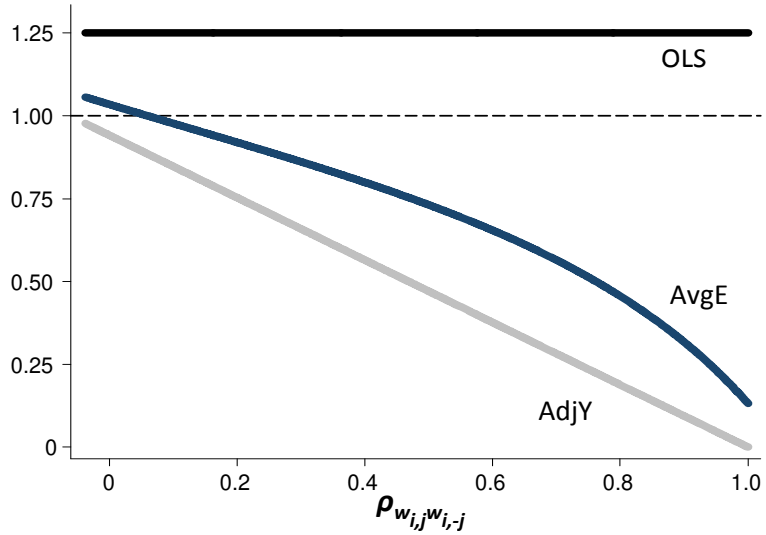
Panel A. Correlation between *X* and *f*



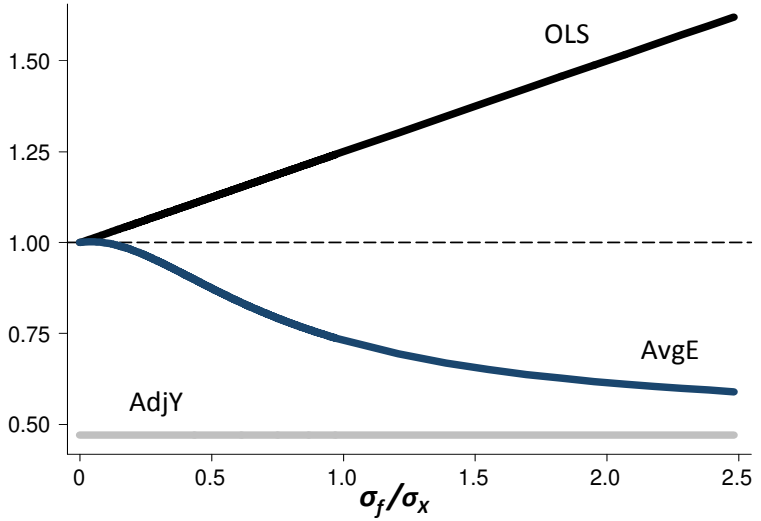
Panel B. Relative variation across versus within groups



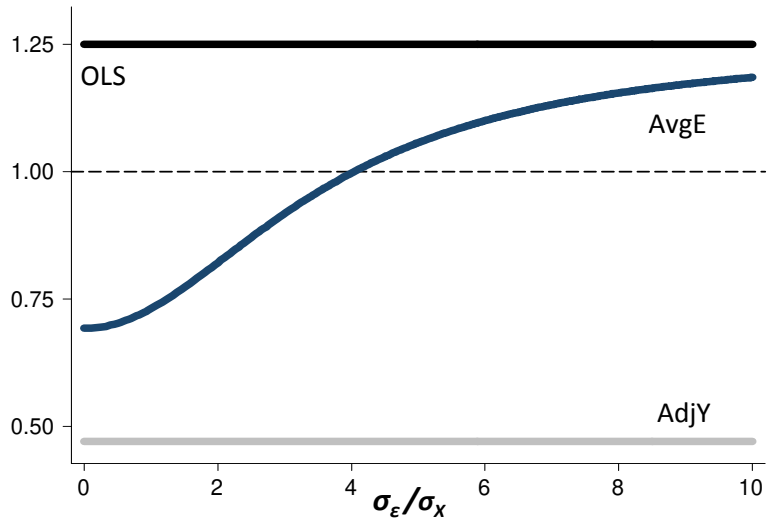
Panel C. Correlation between $w_{i,j}$ and $w_{i,-j}$



Panel D. Relative variation in *f* versus *X*



Panel E. Relative variation in ε versus *X*



Panel F. Number of observations per group

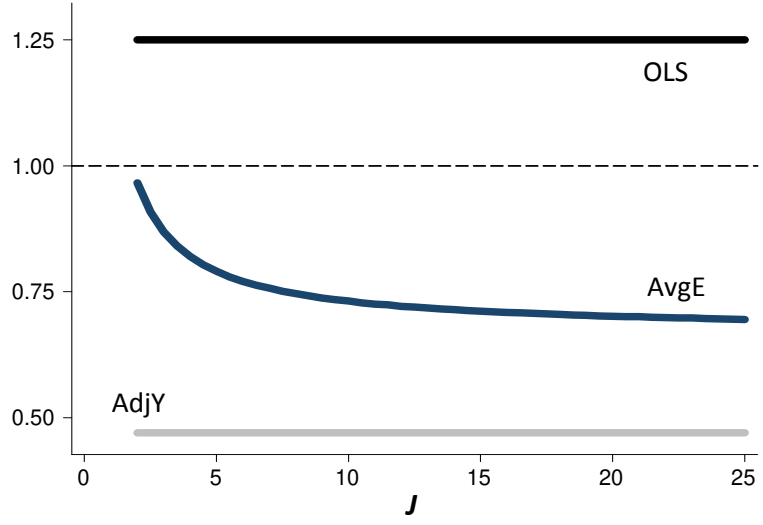


Figure 1. Comparative statics of the bias for OLS, AdjY, and AvgE

This figure presents analytical solutions for the OLS, AdjY, and AvgE estimates of the β in Equation (1), $y_{i,j} = \beta X_{i,j} + f_i + \varepsilon_{i,j}$, as a function of underlying parameters ($\sigma_{\bar{x}}/\sigma_w$, σ_f/σ_X, $\sigma_\varepsilon/\sigma_X$, ρ_{Xf}, $\rho_{w_{i,j}w_{i,-j}}$) and the number of observations in each group (J) when $\beta = 1$. Each panel shows the effect of varying a specific parameter in the data structure while holding the rest constant. When not varying along the horizontal axis, the default parameter values are: $\sigma_f/\sigma_X = \sigma_\varepsilon/\sigma_X = 1$, $\sigma_{\bar{x}}/\sigma_w = \rho_{Xf} = 0.25, \rho_{w_{i,j}w_{i,-j}} = 0.5$, and $J = 10$.

Table 1. Analytical examples of the bias for OLS, AdjY, and AvgE

Underlying Data Structure							Estimates of β		
ρ_{Xf}	$\sigma_{\bar{x}}/\sigma_{w}$	$\rho_{wj,w-j}$	σ_f/σ_X	$\sigma_\varepsilon/\sigma_X$	J	β	β^{OLS}	β^{AdjY}	β^{AvgE}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OLS is less biased than AdjY, AvgE									
0.2	0.5	0.3	1	1	10	1	1.20	0.56	0.77
-0.4	0.4	0.4	1	1	10	1	0.60	0.52	0.39
0.1	0.3	0.5	1	1	10	1	1.10	0.46	0.65
AdjY least biased, AvgE most biased									
-0.6	0.2	0.5	1	1	10	1	0.40	0.48	0.18
-0.5	0.4	0.3	1	1	10	1	0.50	0.60	0.31
-0.3	0.5	0.1	1	1	10	1	0.70	0.72	0.59
AvgE least biased, AdjY most biased									
0.4	0.5	0.5	1	1	10	1	1.40	0.40	0.76
0.3	0.3	0.3	1	1	10	1	1.30	0.64	0.88
0.5	0.4	0.5	1	1	10	1	1.50	0.43	0.85
OLS is more biased than AvgE, AdjY									
0.4	0.5	0.2	1	1	10	1	1.40	0.64	0.93
0.6	0.2	0.4	1	1	10	1	1.60	0.58	1.02
0.2	0.1	0.1	1	1	10	1	1.20	0.89	0.98

This table reports analytical solutions for the OLS, AdjY, and AvgE estimators of Equation (1), $y_{i,j} = \beta X_{i,j} + f_i + \varepsilon_{i,j}$, under different assumptions about the underlying data structure. In all cases, the true coefficient (β) equals 1, the number of observations per group (J) equals 10, and $\sigma_\varepsilon/\sigma_X = \sigma_f/\sigma_X = 1$. The corr($X_{i,j}$, f_i), $\sigma_{\bar{x}}/\sigma_{w}$, and corr($w_{i,j}$, $w_{i,-j}$) are given in Columns (1)-(3), and the OLS, *Adj* Y, and *Avg* E estimates are given in Columns (8)-(10).

Table 2. Firm heterogeneity and capital structure

	Dependent Variable = Total Debt / Total Assets			
	OLS	**AdjY**	**AvgE**	**FE**
	(1)	(2)	(3)	(4)
Fixed Assets/ Total Assets	0.270***	0.066***	0.103***	0.248***
	(0.008)	(0.004)	(0.004)	(0.014)
Ln(sales)	0.011***	0.011***	0.011***	0.017***
	(0.001)	0.000	0.000	(0.001)
Return on Assets	-0.015***	0.051***	0.039***	-0.028***
	(0.005)	(0.004)	(0.004)	(0.005)
Z-score	-0.017***	-0.010***	-0.011***	-0.017***
	0.000	(0.000)	(0.000)	(0.001)
Market-to-book Ratio	-0.006***	-0.004***	-0.004***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	166,974	166,974	166,974	166,974
R^2	0.29	0.14	0.56	0.66

This table summarizes results from firm-panel regressions of *total debt / total assets* 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 heterogeneity across firms. The data are from Compustat for the period 1950–2010 and exclude financial and regulated industries. All variables are winsorized at their 1% tails. Column (1) reports OLS estimates; Column (2) reports AdjY estimates; Column (3) reports AvgE estimates; and Column (4) reports FE estimates. The group averages used to estimate AdjY and AvgE are constructed excluding the observation at hand. Standard errors, adjusted for clustering at the firm level, are reported in parentheses. *** significant at 1% level.

Table 3. Manager heterogeneity and executive compensation

	Dependent Variable = Ln(Total Compensation)			
	OLS	**AdjY**	**AvgE**	**FE**
	(1)	(2)	(3)	(4)
Ln(Total Assets)	0.341***	0.021***	0.066***	0.240***
	(0.007)	(0.002)	(0.005)	(0.014)
Market-to-Book Ratio *[t - 1]*	0.093***	0.009***	0.021***	0.032***
	(0.011)	(0.002)	(0.003)	(0.008)
Stock Return *[t]*	0.120***	0.039***	0.050***	0.091***
	(0.015)	(0.008)	(0.009)	(0.014)
Stock Return *[t - 1]*	0.041***	0.053***	0.051***	0.076***
	(0.011)	(0.007)	(0.007)	(0.010)
Return on Assets *[t]*	0.287***	0.092***	0.120***	0.268***
	(0.062)	(0.032)	(0.033)	(0.083)
Return on Assets *[t - 1]*	0.135**	0.004	0.023	0.193***
	(0.068)	(0.034)	(0.036)	(0.067)
Volatility of Daily Ln(Returns)	0.132***	0.002	0.021	0.002
	(0.030)	(0.012)	(0.013)	(0.023)
CEO = Chairman Indicator	0.225***	0.045***	0.071***	0.028*
	(0.017)	(0.012)	(0.012)	(0.016)
CEO Indicator	0.723***	0.141***	0.224***	0.431***
	(0.017)	(0.012)	(0.015)	(0.017)
Female Indicator	-0.115***	-0.023***	-0.036***	
	(0.026)	(0.005)	(0.006)	
Observations	96,719	96,719	96,719	96,719
R^2	0.47	0.08	0.76	0.82

This table summarizes results from manager-panel regressions of *Ln(Total Compensation)* on year fixed effects and control variables using different methodologies to account for unobserved group heterogeneity across managers. The data are from Execucomp, Compustat, and CRSP for the period 1992–2010. Column (1) reports OLS estimates; Column (2) reports AdjY estimates; Column (3) reports AvgE estimates; and Column (4) reports FE estimates. The group averages used to estimate AdjY and AvgE are constructed excluding the observation at hand. The included control variables are *Log(Total Assets)* ; lagged *market-to-book ratio* ; contemporary and lagged *stock returns* , which is caculated 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 the chairman of the board; an indicator for whether the manager is the CEO; and an indicator for being female. Standard errors, adjusted for clustering at the firm level, are reported in parentheses. * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Table 4. Industry heterogeneity and Tobin's Q

	Dependent Variable = Tobin's Q			
	OLS	**AdjY**	**AvgE**	**FE**
	(1)	(2)	(3)	(4)
Delaware Incorporation	0.100***	0.019	0.040	0.086**
	(0.036)	(0.032)	(0.032)	(0.039)
Ln(sales)	-0.125***	-0.054***	-0.072***	-0.131***
	(0.009)	(0.008)	(0.008)	(0.011)
R&D Expenses / Assets	6.724***	3.022***	3.968***	5.541***
	(0.260)	(0.242)	(0.256)	(0.318)
Return on Assets	-0.559***	-0.526***	-0.535***	-0.436***
	(0.108)	(0.095)	(0.097)	(0.117)
Observations	55,792	55,792	55,792	55,792
R^2	0.22	0.08	0.34	0.37

This table summarizes results from firm-panel 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 heterogeneity across industry-years at the 4-digit SIC industry level. The data are from Compustat for the period 1962–2000 and exclude financial and regulated industries. *Tobin's Q* is measured as (*Total assets + Share price * Shares outstanding – Book equity*) / *Total assets* , and all variables are winsorized at their 1% tails. Column (1) reports OLS estimates; Column (2) reports AdjY estimates; Column (3) reports AvgE estimates; and Column (4) reports FE estimates. The group averages used to estimate *Adj* Y and *Avg* E are constructed excluding the observation at hand. Standard errors, adjusted for clustering at the firm level, are reported in parentheses. ** significant at 5% level; *** significant at 1% level.

Table 5. Industry and size heterogeneity and stock returns

Panel A. Yearly stock returns sorted by R&D Quintile (i.e., OLS)

Missing	Q1	Q2	Q3	Q4	Q5
0.080***	0.068***	0.050***	0.087***	0.124***	0.143***
(0.006)	(0.014)	(0.010)	(0.011)	(0.016)	(0.016)

Panel B. Characteristically adjusted returns by R&D Quintile (i.e., AdjY)

Missing	Q1	Q2	Q3	Q4	Q5
-0.012***	-0.033***	-0.023***	-0.002	0.008	0.020***
(0.003)	(0.009)	(0.008)	(0.007)	(0.013)	(0.006)

Panel C. Regression analysis of R&D (Differences relative to Quintile 1)

	Dependent Variable = Yearly Stock Return			
	OLS	**AdjY**	**AvgE**	**FE**
	(1)	(2)	(3)	(4)
Missing	0.012	0.021**	0.021**	0.030***
	(0.015)	(0.009)	(0.008)	(0.010)
Q2	-0.018	0.01	0.008	0.019
	(0.016)	(0.013)	(0.012)	(0.014)
Q3	0.02	0.032***	0.031***	0.051***
	(0.017)	(0.012)	(0.011)	(0.018)
Q4	0.056***	0.041***	0.042***	0.068***
	(0.021)	(0.015)	(0.015)	(0.020)
Q5	0.075***	0.053***	0.055***	0.094***
	(0.021)	(0.011)	(0.010)	(0.019)
Observations	144,592	144,592	144,592	144,592
R^2	0.00	0.00	0.43	0.47

This table compares two-year holding period stock returns across firms as a function of firms' ratio of research and development expenses to market value of equity ("R&D"). The data are from CRSP and Compustat for the period 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 performs 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 OLS estimates; Column (2) reports AdjY estimates; Column (3) reports AvgE estimates; and Column (4) reports FE estimates. All regressions are weighted by firms' market value of equity, and the group averages used to estimate *Adj* Y and *Avg* E are constructed excluding the observation at hand. Standard errors, adjusted for clustering at the firm level, are reported in parentheses. ** significant at 5% level; *** significant at 1% level.