Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches

Mitchell A. Petersen
Northwestern University

In corporate finance and asset pricing empirical work, researchers are often confronted with panel data. In these data sets, the residuals may be correlated across firms or across time, and OLS standard errors can be biased. Historically, researchers in the two literatures have used different solutions to this problem. This paper examines the different methods used in the literature and explains when the different methods yield the same (and correct) standard errors and when they diverge. The intent is to provide intuition as to why the different approaches sometimes give different answers and give researchers guidance for their use. (*JEL* G12, G3, C01, C15)

It is well known that OLS standard errors are unbiased when the residuals are independent and identically distributed. When the residuals are correlated across observations, OLS standard errors can be biased and either over or underestimate the true variability of the coefficient estimates. Although the use of panel data sets (e.g., data sets that contain observations on multiple firms in multiple years) is common in finance, the ways that researchers have addressed possible biases in the standard errors varies widely and in many cases is incorrect. In recently published finance papers, which include a regression on panel data, 42% of the papers did not adjust the standard errors for possible dependence in the residuals.[1] Approaches for estimating the coefficients and standard errors in the presence of the within-cluster correlation varied among the remaining

I thank the Center for Financial Institutions and Markets at Northwestern University's Kellogg School for support. In writing this paper, I have benefited greatly from discussions with John Ammer, Robert Chirinko, Toby Daglish, Kent Daniel, Joey Engelberg, Gene Fama, Michael Faulkender, Wayne Ferson, Mariassunta Giannetti, John Graham, William Greene, Chris Hansen, Wei Jiang, Toby Moskowitz, Chris Polk, Joshua Rauh, Michael Roberts, Paola Sapienza, Georgios Skoulakis, Doug Staiger, Jeff Wooldridge, Annette Vissing-Jorgensen, the editor, and referees, as well as the comments of seminar participants at the American Finance Association Meetings, Arizona State University, Boston College, Case Western Reserve University, Cornell University, Duke University, Federal Reserve Bank of Chicago, Financial Management Association Meetings, Barclay's Global Investors, Goldman Sachs Asset Management, Harvard Business School, Hong Kong University of Science and Technology, National University of Singapore, Northwestern University, Singapore Management University, Rice University, Stanford University, Stockholm School of Economics, and the Universities of California at Berkeley, Chicago, Columbia, Florida, Iowa, Michigan, Pennsylvania (Wharton), Texas at Dallas, and Washington. The research assistance of Marie Grabinski, Nick Halpern, Casey Liang, Matt Withey, Sungjoon Park, Amit Patel, and Calvin Zhang is greatly appreciated. Send correspondence to Mitchell A. Petersen, Kellogg School of Management, Northwestern University, and NBER, 2001 Sheridan Road, Evanston, IL 60208; telephone: 847-467-1281; E-mail: mpetersen@northwestern.edu.

[1] I searched papers published in the Journal of Finance, the Journal of Financial Economics, and the Review of Financial Studies in the years 2001–2004 for a description of how the coefficients and standard errors were estimated in a panel data set. Panel data sets are data sets that contain multiple observations on a given unit. This can be multiple observations per firm, per industry, per year, or per country. I refer to the unit (e.g., firm

doi:10.1093/rfs/hhn053 Advance Access publication June 3, 2008

papers. Thirty-four percent of the remaining papers estimated both the coefficients and the standard errors using the Fama-MacBeth procedure (Fama and MacBeth, 1973). Twenty-nine percent of the papers included dummy variables for each cluster (e.g., fixed effects or within estimation). The next two most common methods used OLS (or an analogous method) to estimate the coefficients but reported standard errors adjusted for the correlation within a cluster (e.g., within a firm or industry). Seven percent of the papers adjusted the standard errors using the Newey-West procedure (Newey and West, 1987) modified for use in a panel data set, while 23% of the papers reported clustered standard errors (Liang and Zeger, 1986; Moulton, 1986; Arellano, 1987; Moulton, 1990; Andrews, 1991; Rogers, 1993; and Williams, 2000), which are White standard errors adjusted to account for the possible correlation within a cluster. These are also called Rogers standard errors in the finance literature.

Although the literature has used an assortment of methods to estimate standard errors in panel data sets, the chosen method is often incorrect and the literature provides little guidance to researchers as to which method should be used. In addition, some of the advice in the literature is simply wrong. Since the methods sometimes produce incorrect estimates, it is important to understand how the methods compare and how to select the correct one. This is the paper's objective.[2]

There are two general forms of dependence that are most common in finance applications. They will serve as the basis for the analysis. The residuals of a given firm may be correlated across years for a given firm (time-series dependence). I will call this an unobserved firm effect (Wooldridge, 2007). Alternatively, the residuals of a given year may be correlated across different firms (cross-sectional dependence). I will call this a time effect. I will simulate panel data with both forms of dependence, first individually and then jointly. With the simulated data, the coefficients and standard errors will be estimated using each of the methods, and their relative performance will be compared.

Section 1 examines the sensitivity of standard error estimates to the presence of a firm fixed effect, a feature common among many variables including

or industry) as a cluster. I included both linear regressions, as well as nonlinear techniques such as logits and tobits in the survey. I included only papers that report at least five observations in each dimension (e.g., firms and years). Two hundred seven papers met the selection criteria. Papers that did not report the method for estimating the standard errors, or reported correcting the standard errors only for heteroscedasticity (i.e., White standard errors that are not robust to within-cluster dependence) are coded as not having corrected the standard errors for within-cluster dependence. Where the paper's description was ambiguous, I contacted the authors. Although White or OLS standard errors may be correct, many of the published papers report regressions where I would expect the residuals to be correlated across observations on the same firm in different years (e.g., bid-ask spread regressed on exchange dummies, stock price, volatility, and average daily volume or leverage regressed on the market-to-book ratio and firm size) or correlated across observations on different firms in the same year (e.g., equity returns regressed on earnings surprises). In these cases, the bias in the standard errors can be quite large. See Section 5 for two illustrations.

[2] To make it easier for researchers to implement the techniques discussed in this paper, I have posted the code I used for each of the estimation methods discussed in the paper on my web page. I have also posted the basic program that I used to simulate the data and estimate the coefficients and standard errors. This should allow researchers to simulate data sets with their own customized data structure and size, and thus determine for their data sets the magnitude of the biases that I have highlighted.

financial leverage, dividends, and investment. The results show that both OLS and the Fama-MacBeth standard errors are biased downward. The Newey-West standard errors, as modified for panel data, are also biased but the bias is small. Of the most common approaches used in the literature and examined in this paper, only clustered standard errors are unbiased as they account for the residual dependence created by the firm effect. In Section 2, the same analysis is conducted with an unobserved time effect instead of a firm effect. A time effect may be found in equity returns and earnings surprises, for example. Since the Fama-MacBeth procedure is designed to address a time effect, the Fama-MacBeth standard errors are unbiased. The intuition of these first two sections carries over to Section 3, where I simulate data with both a firm and a time effect. I examine estimating standard errors, which are clustered on more than one dimension in this section.

The firm effect was initially specified as a constant (e.g., it does not decay over time). In practice, the firm effect may decay and so the correlation between residuals changes as the time between them grows. In Section 4, I simulate data with a more general correlation structure. This allows a comparison of the OLS, clustered, and Fama-MacBeth standard errors in a more general setting. Simulating the temporary firm effect also allows examination of the relative accuracy of three additional methods for adjusting standard errors (and possibly improving the efficiency of the coefficient estimates): fixed effects (firm dummies), generalized least-squares (GLS) estimation of a random effects model, and adjusted Fama-MacBeth standard errors. I show that including firm dummies or estimating a random effects model with GLS eliminates the bias in the ordinary standard errors only when the firm effect is fixed. I also show that even after adjusting Fama-MacBeth standard errors, as suggested by some authors (e.g., Cochrane, 2001), they are still biased in many, but not all, cases.

Most papers do not report standard errors estimated by multiple methods. Thus in Section 5, I apply the various estimation techniques to two real data sets and compare their relative performance. This serves two purposes. First, it demonstrates that the methods used in some published papers may produce biases in the standard errors and *t*-statistics that are very large. This is why using the correct method to estimate standard errors is important. Examining actual data also allows me to show how differences in standard error estimates can provide information about the deficiency in a model and directions for improving it.

1. Estimating Standard Errors in the Presence of a Fixed Firm Effect

1.1 Clustered standard error estimates

To provide intuition on why the standard errors produced by OLS are incorrect and how alternative estimation methods correct this problem, it is helpful to very briefly review the expression for the variance of the estimated coefficients.

The standard regression for a panel data set is

$$Y_{it} = X_{it}\,\beta + \varepsilon_{it}, \tag{1}$$

where there are observations on firms (i) across years (t). X and ε are assumed to be independent of each other and to have a zero mean and finite variance.[3] I have made the assumption that the model is correctly specified. The zero mean is without loss of generality and allows calculation of variances as sums of the squares of the variable. The estimated coefficient is

$$\begin{aligned}\hat{\beta}_{OLS} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}Y_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}(X_{it}\beta + \varepsilon_{it})}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} \\ &= \beta + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2},\end{aligned} \tag{2}$$

and the asymptotic variance of the estimated coefficient is

$$\begin{aligned}AVar[\hat{\beta}_{OLS} - \beta] &= \underset{\substack{N\to\infty \\ T\,fixed}}{\text{plim}} \left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\varepsilon_{it}\right)^2 \left(\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N}\right)^{-2}\right] \\ &= \underset{\substack{N\to\infty \\ T\,fixed}}{\text{plim}} \left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2\right) \left(\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N}\right)^{-2}\right]. \\ &= \frac{1}{N}\left(T\sigma_X^2\sigma_\varepsilon^2\right)\left(T\sigma_X^2\right)^{-2} \\ &= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT}.\end{aligned} \tag{3}$$

This is the standard OLS formula and is correct when the errors are independent and identically distributed (Greene, 2002). The independence assumption is used to move from the first to the second equality in Equation (3) (i.e., the covariance between residuals is zero). The identical distribution assumption (e.g., homoscedastic errors) is used to move from the second to the third equality.[4] The independence assumption is often violated in panel data and is the focus of the paper.

In relaxing the assumption of independent errors, I initially assume that the data have an unobserved firm effect that is fixed. Thus the residuals consist of

[3] The simulations in the paper are based on linear regressions. However, the results generalize to nonlinear models such as probit and tobit. In simulated results, the clustered standard errors are unbiased and the regular ("OLS") standard errors are biased in these nonlinear models. The magnitude of the biases are similar to what I report for linear models (results available from the author).

[4] Clustered standard errors are robust to heteroscedasticity. Since this is not my focus, I assume the errors are homoscedastic in the simulations and derivations. I use White standard errors as my baseline estimates when analyzing actual data in Section 5, since the residuals are not homoscedastic in those data sets (White, 1984).

a firm-specific component (γ_i) and an idiosyncratic component that is unique to each observation (η_{it}).[5] The residuals can be specified as

$$\varepsilon_{it} = \gamma_i + \eta_{it}. \tag{4}$$

I assume that the independent variable X also has a firm-specific component:

$$X_{it} = \mu_i + \upsilon_{it}. \tag{5}$$

The components of X (μ and ν) and ε (γ and η) have zero mean, finite variance, and are independent of each other. This is necessary for the coefficient estimates to be consistent.[6] Both the independent variable and the residual are correlated across observations of the same firm, but are independent across firms:

$$\begin{aligned} corr(X_{it}, X_{js}) &= 1 \quad \text{for } i = j \text{ and } t = s \\ &= \rho_X = \sigma_\mu^2 / \sigma_X^2 \quad \text{for } i = j \text{ and all } t \neq s \\ &= 0 \quad \text{for all } i \neq j, \\ corr(\varepsilon_{it}, \varepsilon_{js}) &= 1 \quad \text{for } i = j \text{ and } t = s \\ &= \rho_\varepsilon = \sigma_\gamma^2 / \sigma_\varepsilon^2 \quad \text{for } i = j \text{ and all } t \neq s \\ &= 0 \quad \text{for all } i \neq j. \end{aligned} \tag{6}$$

Given this data structure (Equations (1), (4), and (5)), the true standard error of the OLS coefficient can be determined. Since the residuals are no longer independent within a cluster, the square of the summed residuals is not equal to the sum of the squared residuals. The same statement can be made about the independent variable. The covariances must be included as well. The asymptotic variance of the OLS coefficient estimate is

$$\begin{aligned} AVar[\hat{\beta}_{OLS} - \beta] &= \underset{\substack{N\to\infty \\ T\ fixed}}{\text{plim}} \left[\frac{1}{N^2} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}\varepsilon_{it} \right)^2 \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\ &= \underset{\substack{N\to\infty \\ T\ fixed}}{\text{plim}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}\varepsilon_{it} \right)^2 \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\ &= \underset{\substack{N\to\infty \\ T\ fixed}}{\text{plim}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}^2 \varepsilon_{it}^2 + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} X_{it} X_{is} \varepsilon_{it} \varepsilon_{is} \right) \right. \end{aligned}$$

[5] This language follows Wooldridge (2007). When I use the word fixed it means that the unobserved firm effect does not die away over time. Wooldridge calls this a time-constant unobserved effect. This means the correlation between $\varepsilon_{i\,t}$ and $\varepsilon_{i\,t-k}$ is a constant with respect to k. In Section 4, I will examine cases where this correlation dies away (decline) as k increases.

[6] I am assuming that the model is correctly specified. I do this to focus on estimating the standard errors. In actual data sets, this assumption does not necessarily hold and would need to be considered and tested.

$$\times \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \Bigg]$$
$$= \frac{1}{N} \left(T\sigma_X^2 \sigma_\varepsilon^2 + T(T-1)\rho_X \sigma_X^2 \rho_e \sigma_e^2 \right) \left(T\sigma_X^2 \right)^{-2}$$
$$= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT} (1 + (T-1)\rho_X \rho_\varepsilon). \tag{7}$$

I use the assumption that residuals are independent across firms in deriving the second equality. Given the assumed data structure, the within-cluster correlations of both X and ε are positive and are equal to the fraction of the variance that is attributable to the firm effect. When the data have a fixed firm effect, the OLS standard errors will understate the true standard error if and only if both ρ_X and ρ_ε are nonzero.[7] The magnitude of the error is also increasing in the number of years. To understand this intuition, consider the extreme case where the independent variables and residuals are perfectly correlated across time (i.e., $\rho_X = 1$ and $\rho_\varepsilon = 1$). In this case, each additional year provides no additional information and will have no effect on the true standard error. However, the OLS standard errors will assume that each additional year provides N additional observations, and the estimated standard error will shrink accordingly and incorrectly.

The correlation of the residuals within a cluster is the problem the clustered standard errors are designed to correct.[8] By squaring the sum of $X_{it}\varepsilon_{it}$ within each cluster, the covariance between residuals within the cluster is estimated (Figure 1). This correlation can be of any form; no parametric structure is assumed. However, the squared sum of $X_{it}\varepsilon_{it}$ is assumed to have the same distribution across the clusters. Thus these standard errors are consistent as the number of clusters grows (Donald and Lang, 2007; and Wooldridge, 2007). I will return to this issue in Section 2.

---

[7] If the firm effect is not fixed, the variance of the coefficient estimate is a weighted sum of the correlations between ε_t and $\varepsilon_{t\text{-}k}$ times the correlation between X_t and X_{t-k}, for all $k < T$ (Wooldridge, 2007). It is equal to

$$Var[\hat{\beta}_{OLS} - \beta] = \frac{\sigma_\varepsilon^2}{NT\sigma_X^2} \left(1 + \frac{2}{T} \sum_{k=1}^{T} (T-k)\rho_{x,k}\rho_{\varepsilon,k} \right).$$

Since the autocorrelations can be positive or negative, it is possible for the OLS standard error to under- or overestimate the true standard error. If the panel is unbalanced (different T for each i), the true standard error and the bias in the OLS standard errors are even larger than specified by Equation (7) (see Moulton, 1986). Results available from the author.

[8] The exact formula for the clustered standard error is

$$AVar(\beta) = \frac{N(NT-1)\sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}\varepsilon_{it} \right)^2}{(NT-k)(N-1)\left(\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2 \right)^2}.$$

	Firm 1			Firm 2			Firm 3		
Firm 1	ε_{11}^2	$\varepsilon_{11}\,\varepsilon_{12}$	$\varepsilon_{11}\,\varepsilon_{13}$	0	0	0	0	0	0
	$\varepsilon_{12}\,\varepsilon_{11}$	ε_{12}^2	$\varepsilon_{12}\,\varepsilon_{13}$	0	0	0	0	0	0
	$\varepsilon_{13}\,\varepsilon_{11}$	$\varepsilon_{13}\,\varepsilon_{12}$	ε_{13}^2	0	0	0	0	0	0
Firm 2	0	0	0	ε_{21}^2	$\varepsilon_{21}\,\varepsilon_{22}$	$\varepsilon_{21}\,\varepsilon_{23}$	0	0	0
	0	0	0	$\varepsilon_{22}\,\varepsilon_{21}$	ε_{22}^2	$\varepsilon_{22}\,\varepsilon_{23}$	0	0	0
	0	0	0	$\varepsilon_{23}\,\varepsilon_{21}$	$\varepsilon_{23}\,\varepsilon_{22}$	ε_{23}^2	0	0	0
Firm 3	0	0	0	0	0	0	ε_{31}^2	$\varepsilon_{31}\,\varepsilon_{32}$	$\varepsilon_{31}\,\varepsilon_{33}$
	0	0	0	0	0	0	$\varepsilon_{32}\,\varepsilon_{31}$	ε_{32}^2	$\varepsilon_{32}\,\varepsilon_{33}$
	0	0	0	0	0	0	$\varepsilon_{33}\,\varepsilon_{31}$	$\varepsilon_{33}\,\varepsilon_{32}$	ε_{33}^2

Figure 1
Residual cross product matrix: Assumptions about zero covariances
The figure shows a sample covariance matrix of the residuals. Assumptions about the elements of this matrix and which are zero is the source of difference in the various standard error estimates. The standard OLS assumption is that only the diagonal terms are nonzero. Standard errors clustered by firm assume that the correlation of the residuals within the cluster may be nonzero (these elements are shaded). This cluster assumption assumes that residuals across clusters are uncorrelated. These are recorded as zero in the matrix.

1.2 Testing the standard error estimates by simulation

I simulated a panel data set and then estimated the slope coefficient and its standard error. By doing this multiple times, the true standard error can be observed, as well as the average estimated standard errors.[9] In the first version of the simulation, an unobserved firm effect that is fixed is included, but no time effect in the independent variable or the residual. Thus the data are simulated as described in Equations (4) and (5). Across simulations, the standard deviation of the independent variable and the residual are both assumed to be constant at 1 and 2, respectively. This will produce an R^2 of 20%. Across different simulations, I altered the fraction of the variance in the independent variable that is due to the firm effect. This fraction ranges from 0% to 75% in 25% increments (Table 1). The same thing was done for the residual. This demonstrates how the magnitude of the bias in the OLS standard errors varies with the strength of the firm effect in both the independent variable and the residual.

The results of the simulations are reported in Table 1. The first two entries in each cell are the average value of the slope coefficient and the standard deviation of the coefficient estimate. The standard deviation of the coefficient estimate is the true standard error of the coefficient, and ideally the estimated standard error will be close to this number. The average standard error estimated by OLS is the third entry in each cell and is the same as the true standard error

[9] Each simulated data set contains five thousand observations (five hundred firms and ten years per firm). The components of the independent variable (μ ν) and the residual (γ η) are mutually independent and normally distributed with zero means. For each data set, I estimated the coefficients and standard errors using each method described below. The means and standard deviations reported in the tables are based on five thousand simulations.

Table 1
Estimating standard errors with a firm effect OLS and clustered standard errors

		Source of independent variable volatility			
Avg(β_{OLS}) Std(β_{OLS}) Avg(SE_{OLS}) % Sig(T_{OLS}) Avg(SE_C) Avg(SE_C)		0%	25%	50%	75%
Source of residual volatility	0%	1.0004	1.0006	1.0002	1.0001
		0.0286	0.0288	0.0279	0.0283
		0.0283	0.0283	0.0283	0.0283
		[0.0098]	[0.0088]	[0.0094]	[0.0094]
		0.0283	0.0282	0.0282	0.0282
		[0.0108]	[0.0092]	[0.0096]	[0.0098]
	25%	1.0004	0.9997	0.9999	0.9997
		0.0287	0.0353	0.0403	0.0468
		0.0283	0.0283	0.0283	0.0283
		[0.0116]	[0.0348]	[0.0678]	[0.1174]
		0.0283	0.0353	0.0411	0.0463
		[0.0120]	[0.0064]	[0.0112]	[0.0118]
	50%	1.0001	1.0002	1.0007	0.9993
		0.0289	0.0414	0.0508	0.0577
		0.0283	0.0283	0.0283	0.0283
		[0.0124]	[0.0770]	[0.1534]	[0.2076]
		0.0282	0.0411	0.0508	0.0590
		[0.0128]	[0.0114]	[0.0088]	[0.0102]
	75%	1.0000	1.0004	0.9995	1.0016
		0.0285	0.0459	0.0594	0.0698
		0.0283	0.0283	0.0283	0.0283
		[0.0128]	[0.1090]	[0.2230]	[0.2906]
		0.0282	0.0462	0.0589	0.0693
		[0.0128]	[0.0114]	[0.0094]	[0.0112]

The table contains estimates of the coefficient and standard errors based on five thousand simulated panel data sets, each of which contains five hundred firms and ten years per firm. The true slope coefficient is 1, the standard deviation of the independent variable is 1, and the standard deviation of the error term is 2 (see Equation (1)). The independent variable X and the residual are specified as

$$X_{it} = \mu_i + \upsilon_{it}$$
$$\varepsilon_{it} = \gamma_i + \eta_{it}.$$

The fraction of X's variance, which is due to a firm-specific component [Var(μ)/Var(X)], varies across the columns of the table from 0% (no firm effect) to 75% and the fraction of the residual variance, which is due to a firm-specific component [Var(γ)/Var(ε)], varies across the rows of the table from 0% (no firm effect) to 75%. Each cell contains the average slope coefficient estimated by OLS and the standard deviation of this estimate. This is the true standard error of the estimated coefficient. The third entry is the average standard error estimated by OLS. The percentage of OLS t-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets. The fifth entry is the average standard error clustered by firm (i.e., accounts for the possible correlation between observations of the same firm in different years). The percentage of clustered t-statistics that are significant at the 1% level is reported in square brackets. For example, when 50% of the variability in both the residual and the independent variable is due to the fixed firm effect ($\rho_X = \rho_\varepsilon = 0.50$), the true standard error of the OLS coefficient is 0.0508. The OLS standard error estimate is 0.0283 and the clustered standard error is 0.0508; 15.3% of the OLS t-statistics are greater than 2.58 in absolute value (only 1% should be), while 0.9% of the clustered t-statistics are greater than 2.58 in absolute value.

in the first row of the table. When there is no firm effect in the residual (i.e., the residuals are independent across observations), the standard error estimated by OLS is correct (Table 1, row 1). When there is no firm effect in the independent variable (i.e., the independent variable is independent across observations), the standard errors estimated by OLS are also unbiased, even if the residuals are highly correlated (Table 1, column 1). This follows from the intuition in Equation (7). The bias in the OLS standard errors is a product of the dependence in the independent variable (ρ_X) and the residual (ρ_ε). When either correlation is zero, OLS standard errors are unbiased.

When there is a fixed firm effect in both the independent variable and the residual, then the OLS standard errors underestimate the true standard errors, and the magnitude of the underestimation can be large. For example, when 50% of the variability in both the residual and the independent variable is due to the firm effect ($\rho_X = \rho_\varepsilon = 0.50$), the OLS estimated standard error is one half of the true standard error (0.557 = 0.0283/0.0508).[10] The standard errors estimated by OLS do not rise as the firm effect increases across either the columns in Table 1 (i.e., in the independent variable) or across the rows (i.e., in the residual). The true standard error does rise.

When the standard error of the coefficient is estimated using clustered standard errors, the estimates (the fifth entry in each cell) are very close to the true standard error (Table 1). These estimates rise along with the true standard error as the fraction of variability arising from the firm effect increases. The clustered standard errors correctly account for the dependence in the data common in a panel data set and produce unbiased estimates.

An alternative way to examine the magnitude of the bias is to examine the empirical distribution of the simulated *t*-statistics (see Skoulakis, 2006). The fraction of OLS *t*-statistics that are statistically significant at the 1% level (i.e., greater than 2.58) are reported as the fourth entry in each cell of Table 1. The *t*-statistics based on the OLS standard errors are too large in absolute value (Figure 2A and Table 1). As you move down the diagonal in Table 1, the percentage of *t*-statistics that are statistically significant at the 1% level rises. For example, 15.3% of the OLS *t*-statistics are statistically significant at the 1% level when $\rho_X = \rho_\varepsilon = 0.50$. The clustered standard errors are unbiased (Table 1) and the empirical distribution of the *t*-statistics is also correct (Figure 2B); 0.9% of the clustered *t*-statistics are significant at the 1% level. The reason the *t*-statistics give the same intuition as the standard errors is that the standard errors are estimated very precisely. For example, the mean OLS standard error

[10] All of the regressions contain a constant whose true value is zero. The paper's intuition carries over to the intercept estimation. The estimated intercept averages −0.0003 with a standard deviation of 0.0669, when $\rho_X = \rho_\varepsilon = 0.50$. The OLS standard errors are biased (0.0283) and the clustered standard errors are unbiased (0.0663). The simulated residuals are homoscedastic, so calculating standard errors that are robust to heteroscedasticity is unnecessary. When I estimated White standard errors in the simulation they have the same bias as the OLS standard errors. For example, the average White standard error of the slope is 0.0283 compared to the OLS estimate of 0.0283 and a true standard error of 0.0508 when $\rho_X = \rho_\varepsilon = 0.50$. These results are available from the author.

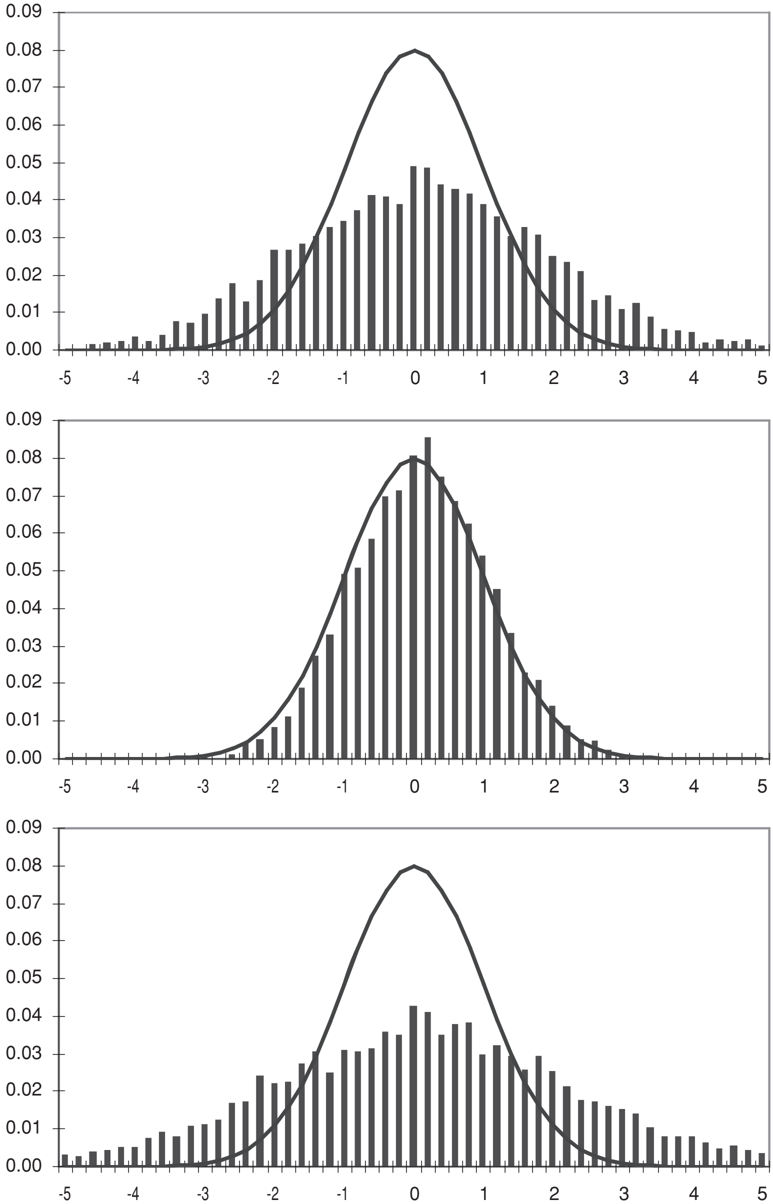

Figure 2
Distribution of simulated *t*-statistics
The figures contain the theoretical *t*-distribution (the line), and the distribution of *t*-statistics produced by the simulation (the bars) when 50% of the variability in the independent variable and the residual is due to a fixed firm effect. The top figure is the distribution of the *t*-statistics based on the OLS standard errors, the middle figure is the distribution of *t*-statistics based on the standard errors clustered by firm, and the bottom figure is the distribution of *t*-statistics based on Fama-MacBeth standard errors. When the standard errors estimates are too small (as with OLS and Fama-MacBeth), there are too many *t*-statistics that are large in absolute value.

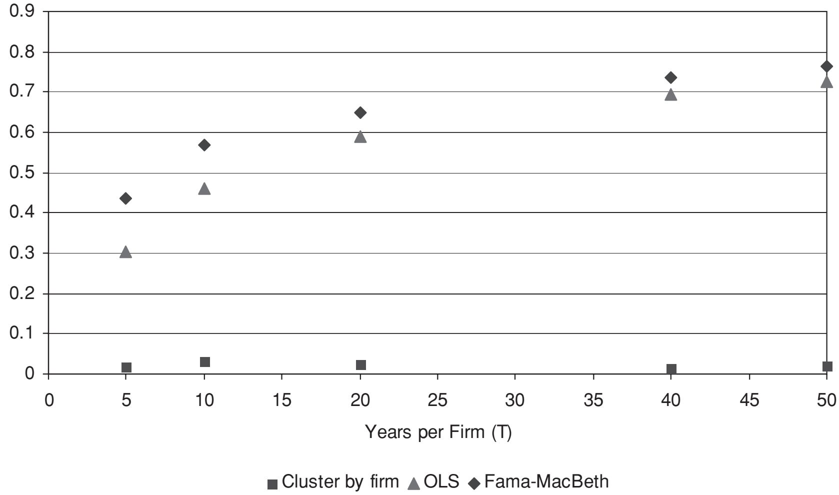

Figure 3
Bias in estimated standard errors as a function of years per cluster
The figure graphs the percentage by which OLS (triangles), clustered by firm (squares) and Fama-MacBeth (diamonds) standard errors, underestimate the true standard error in the presence of a fixed firm effect. The results are based on five thousand simulations of a data set with five thousand observations. The number of years per firm ranges from 5 to 50. The firm effect is assumed to comprise 50% of the variability in both the independent variable and the residual. The underestimates are calculated as one minus the average estimated standard error divided by the true standard deviation of the coefficient estimate. For example, the standard deviation of the coefficient estimate was 0.0406 in the simulation with five years of data ($T = 5$). The average of the OLS estimated standard errors is 0.0283. Thus the OLS underestimated the true standard error by 30% (1 − 0.0283/0.0406).

is 0.0283 with a standard deviation of 0.0007, and the mean clustered standard error is 0.0508 with a standard deviation of 0.0027 (when $\rho_X = \rho_\varepsilon = 0.50$).[11]

The bias in OLS standard errors is highly sensitive to the number of time periods (years) used in the estimation as well. As the number of years doubles, OLS assumes a doubling of the information. However, if both the independent variable and the residual are serially correlated within the cluster, the amount of information increases by less than a factor of 2. The bias rises from about 30% when there are five years of data per firm to 73% when there are fifty years (when $\rho_X = \rho_\varepsilon = 0.50$, see Figure 3). The robust standard errors are consistently close to the true standard errors independent of the number of time periods (Figure 3).[12]

[11] The mean squared error (MSE) of the standard error estimates are not reported, since they add no additional information beyond what is reported in the tables and the figures in most instances. This is because the variances of the standard error estimates are extremely small. The MSEs are essentially equal to the bias squared. The one exception I found was the adjusted Fama-MacBeth standard errors, which are discussed in Section 4.4. Tables of MSEs are available from the author.

[12] Although the bootstrap method of estimating standard errors was rarely used in the articles I surveyed, it is another alternative for estimating standard errors in a panel data set (see, for example, Kayhan and Titman, 2007; and Efron and Tibshirani, 1986). To test its relative performance, I drew 100 samples with replacement and re-estimated the regression for each simulated data set. When I drew observations independently (e.g., I drew

1.3 Fama-MacBeth standard errors: the equations

An alternative way to estimate the regression coefficients and standard errors when the residuals are not independent is the Fama-MacBeth approach (Fama and MacBeth, 1973). In this approach, the researcher runs T cross-sectional regressions. The average of the T estimates is the coefficient estimate:

$$\begin{aligned}\hat{\beta}_{FM} &= \sum_{t=1}^{T} \frac{\hat{\beta}_t}{T} \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N} X_{it}Y_{it}}{\sum_{i=1}^{N} X_{it}^2}\right) = \beta + \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N} X_{it}^2}\right),\end{aligned} \tag{8}$$

and the estimated variance of the Fama-MacBeth estimate is calculated as

$$S^2(\hat{\beta}_{FM}) = \frac{1}{T}\sum_{t=1}^{T}\frac{(\hat{\beta}_t - \hat{\beta}_{FM})^2}{T-1}. \tag{9}$$

The variance formula assumes that the yearly estimates of the coefficient (β_t) are independent of each other. This is only correct if X_{it} ε_{it} is independent of X_{is} ε_{is} for $t \neq s$. As discussed above, this is not true when there is a firm effect in the data (i.e., ρ_X $\rho_\varepsilon \neq 0$). Thus, the Fama-MacBeth variance estimate is too small in the presence of a firm effect. In this case, the asymptotic variance of the Fama-MacBeth estimate is

$$\begin{aligned}AVar(\hat{\beta}_{FM}) &= \frac{1}{T^2}AVar\left(\sum_{t=1}^{T}\hat{\beta}_t\right) \\ &= \frac{AVar(\hat{\beta}_t)}{T} + \frac{2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} ACov(\hat{\beta}_t, \hat{\beta}_s)}{T^2} \\ &= \frac{AVar(\hat{\beta}_t)}{T} + \frac{T(T-1)}{T^2}ACov(\hat{\beta}_t, \hat{\beta}_s).\end{aligned} \tag{10}$$

Given the data generating process (Equations (4) and (5)), the covariance between the coefficient estimates of different years is independent of $t - s$ (which justifies the simplification in the last line of Equation (10)) and can be

five thousand firm-years), the estimated standard errors were the same as the OLS standard errors reported in Table 1 (e.g., 0.0282 for the bootstrap versus 0.0283 for OLS when $\rho_X = \rho_\varepsilon = 0.50$). When I drew observations as a cluster (e.g., I drew five hundred firms with replacement and took all ten years for any firm that was drawn), the estimated standard errors were the same as the clustered standard errors (e.g., 0.0505 for bootstrap versus 0.0508 for clustered). An example with real data can be found in Cheng, Nagar, Rajan (2005). These authors find that bootstrapped standard errors (when a state opposed to a single observation is drawn) are almost identical to the standard errors clustered by state.

calculated as follows for $t \neq s$:

$$
\begin{aligned}
ACov(\hat{\beta}_t, \hat{\beta}_s) &= \operatorname*{plim}_{N\to\infty} \left[\left(\frac{\sum_{i=1}^{N} X_{it}^2}{N} \right)^{-1} \left(\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{N} \right) \right. \\
&\qquad \left. \times \left(\frac{\sum_{i=1}^{N} X_{is}\varepsilon_{is}}{N} \right) \left(\frac{\sum_{i=1}^{N} X_{is}^2}{N} \right)^{-1} \right] \\
&= \left(\sigma_X^2\right)^{-2} \operatorname*{plim}_{N\to\infty} \left[\left(\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{N} \right) \left(\frac{\sum_{i=1}^{N} X_{is}\varepsilon_{is}}{N} \right) \right] \\
&= \left(\sigma_X^2\right)^{-2} \operatorname*{plim}_{N\to\infty} \left[\frac{\sum_{i=1}^{N} X_{it}X_{is}\varepsilon_{it}\varepsilon_{is}}{N^2} \right] \\
&= \left(\sigma_X^2\right)^{-2} \frac{N\rho_X\sigma_X^2\rho_\varepsilon\sigma_\varepsilon^2}{N^2} \\
&= \frac{\rho_X\rho_\varepsilon\sigma_\varepsilon}{N\sigma_X^2}. \qquad (11)
\end{aligned}
$$

Combining Equations (10) and (11) gives the expression for the asymptotic variance of the Fama-MacBeth coefficient estimates:

$$
\begin{aligned}
AVar(\hat{\beta}_{FM}) &= \frac{AVar(\hat{\beta}_t)}{T} + \frac{T(T-1)}{T^2} ACov(\hat{\beta}_t, \hat{\beta}_s) \\
&= \frac{1}{T}\left(\frac{\sigma_\varepsilon^2}{N\sigma_X^2}\right) + \frac{T(T-1)}{T^2}\left(\frac{\rho_X\rho_\varepsilon\sigma_\varepsilon^2}{N\sigma_X^2}\right) \\
&= \frac{\sigma_\varepsilon^2}{NT\sigma_X^2}(1 + (T-1)\rho_X\rho_\varepsilon). \qquad (12)
\end{aligned}
$$

This result is the same as the expression for the variance of the OLS coefficient (Equation (7)). The Fama-MacBeth standard errors are biased in exactly the same way as the OLS estimates. In both cases, the magnitude of the bias is a function of the serial correlation of both the independent variable and the residual within a cluster and the number of time periods per firm (or cluster).

1.4 Simulating Fama-MacBeth standard errors

To document the bias of the Fama-MacBeth standard error estimates, I calculate the Fama-MacBeth estimate of the slope coefficient and the standard error in each of the five thousand simulated data sets that were used in Table 1. The results are reported in Table 2. The Fama-MacBeth estimates are consistent and as efficient as OLS (the correlation between the two is consistently above 0.99). The standard deviation of the two coefficient estimates is also the same (compare the second entry in each cell of Tables 1 and 2). These results demonstrate that both OLS and Fama-MacBeth standard errors are biased downward

Table 2
Estimating standard errors with a firm effect Fama-MacBeth standard errors

$\text{Avg}(\beta_{FM})$ $\text{Std}(\beta_{FM})$ $\text{Avg}(SE_{FM})$ $\%\ \text{Sig}(T_{FM})$		Source of independent variable volatility: 0%	25%	50%	75%
Source of residual volatility	0%	1.0004 0.0287 0.0276 [0.0288]	1.0006 0.0288 0.0276 [0.0304]	1.0002 0.0280 0.0277 [0.0236]	1.0001 0.0283 0.0275 [0.0294]
	25%	1.0004 0.0288 0.0275 [0.0336]	0.9997 0.0354 0.0268 [0.0758]	0.9998 0.0403 0.0259 [0.1202]	0.9997 0.0469 0.0250 [0.1918]
	50%	1.0000 0.0289 0.0276 [0.0330]	1.0002 0.0415 0.0259 [0.1264]	1.0007 0.0509 0.0238 [0.2460]	0.9993 0.0578 0.0219 [0.3388]
	75%	1.0000 0.0286 0.0277 [0.0310]	1.0004 0.0460 0.0248 [0.1778]	0.9995 0.0595 0.0218 [0.3654]	1.0016 0.0699 0.0183 [0.4994]

The table contains estimates of the coefficient and standard errors based on the same simulated panel data sets that are used in Table 1. Each data set contains five hundred firms and ten years per firm. The true slope coefficient is 1, the standard deviation of the independent variable is 1, and the standard deviation of the error term is 2 (see Equation (1)). The independent variable X and the residual are specified as

$$\begin{aligned} X_{it} &= \mu_i + \upsilon_{it} \\ \varepsilon_{it} &= \gamma_i + \eta_{it}. \end{aligned}$$

The fraction of X's variance, which is due to a firm-specific component [Var(μ)/Var(X)], varies across the columns of the table from 0% (no firm effect) to 75% and the fraction of the residual variance, which is due to a firm-specific component [Var(γ)/Var(ε)], varies across the rows of the table from 0% (no firm effect) to 75%. The first entry is the average slope coefficient based on a Fama-MacBeth estimation (e.g., a regression is run for each of the ten years and the estimate is the average of the ten estimated slope coefficients). The second entry is the standard deviation of the coefficient estimated by Fama-MacBeth. This is the true standard error of the Fama-MacBeth coefficient. The third entry is the average standard error estimated by Fama-MacBeth (see Equation (9)). The percentage of Fama-MacBeth t-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets. For example, when 50% of the variability in both the residual and the independent variable is due to the firm effect ($\rho_X = \rho_\varepsilon = 0.50$), the true standard error of the Fama-MacBeth coefficient is 0.0509. The Fama-MacBeth standard error estimate is 0.0238, and 24.6% of t-statistics are greater than 2.58 in absolute value.

(Table 2). However, the Fama-MacBeth standard errors have a larger bias than the OLS standard errors. For example, when both ρ_X and ρ_ε are equal to 75%, the OLS standard error has a bias of 60% ($0.595 = 1 - 0.0283/0.0698$, see Table 1) and the Fama-MacBeth standard error has a bias of 74% ($0.738 = 1 - 0.0183/0.0699$, see Table 2). Moving down the diagonal of Table 2 from upper left to bottom right, the true standard error increases but the standard error estimated by Fama-MacBeth actually shrinks. Remember, the estimated OLS standard errors did not change as we moved down the diagonal of Table 1. As the firm effect becomes larger (ρ_X ρ_ε increases), the OLS bias grows and the Fama-MacBeth bias grows even faster.[13] The incremental bias of the Fama-MacBeth standard errors is due to the way in which the estimated variance is calculated. To understand why, the expression of the estimated variance needs to be expanded (Equation (9)):

$$S^2(\hat{\beta}_{FM}) = \frac{1}{T(T-1)}\sum_{t=1}^{T}\left[\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N} X_{it}^2} - \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N} X_{it}^2}\right)\right]^2$$
$$= \frac{1}{T(T-1)}\sum_{t=1}^{T}\left[\frac{\sum_{i=1}^{N}(\mu_i+\upsilon_{it})(\gamma_i+\eta_{it})}{\sum_{i=1}^{N}(\mu_i+\upsilon_{it})^2} - \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N}(\mu_i+\upsilon_{it})(\gamma_i+\eta_{it})}{\sum_{i=1}^{N}(\mu_i+\upsilon_{it})^2}\right)\right]^2. \quad (13)$$

The true variance of the Fama-MacBeth coefficients is a measure of how far each yearly coefficient estimate deviates from the true coefficient (one in the simulations). The estimated variance, however, measures how far each yearly estimate deviates from the sample average. Since the firm effect influences both the yearly coefficient estimate and the sample average of the yearly coefficient estimates, it does not appear in the estimated variance. Thus increases in the firm effect (increases in $\rho_X\ \rho_\varepsilon$) actually reduce the estimated Fama-MacBeth standard error at the same time it increases the true standard error of the estimated coefficients. To make this concrete, take the extreme example where $\rho_X\ \rho_\varepsilon$ is equal to one; the true standard error is $(\sigma_\varepsilon/N\sigma_X)^{1/2}$ while the estimated Fama-MacBeth standard error is zero. This additional source of bias shrinks as the number of years increases since the estimated slope coefficient will converge to the true coefficient (Figure 3).

1.5 Standard error estimates in published papers

While the above discussion demonstrates that the Fama-MacBeth standard errors are biased in the presence of a firm effect, they are often used to measure statistical significance in published papers when the underlying regression likely contains a firm effect. As part of the literature survey, a search was conducted for papers that ran a regression of one persistent firm characteristic on other persistent firm characteristics (i.e., the serial correlation of the variables is large and dies away slowly as the lag between observations increases). This is the type of data where Fama-MacBeth (and OLS) standard errors may be biased. Since it is not feasible for me to replicate each of the studies in the survey, an alternative approach will be taken. There will first be a discussion of several examples of where the literature has regressed persistent dependent

[13] The distribution of empirical *t*-statistics is even wider for the Fama-MacBeth than for OLS (compare Figures 2A and C). Twenty-five percent of the Fama-MacBeth *t*-statistics are statistically significant at the 1% level compared to 15% of the OLS *t*-statistics when $\rho_X = \rho_\varepsilon = 0.50$ (Tables 1 and 2).

variables on persistent independent variables. This is the data structure where biases in Fama-MacBeth and OLS/White standard errors are most likely.[14] I will then examine a real data set in Section 5.2, and estimate White, Fama-MacBeth, and clustered standard errors. It can then be shown that in real data sets that contain a firm effect, White and Fama-MacBeth standard errors are too small and that the magnitude of the bias can be quite large.

The first example of a regression with a persistent dependent and independent variable is the regression of whether a firm pays a dividend on firm characteristics such as the firm's market-to-book ratio, the earnings-to-assets ratio, and the relative firm size (e.g., Fama and French, 2001). A second example where both the dependent and independent variables are persistent are papers that examine how the market values firms by regressing a firm's market-to-book ratio on firm characteristics such as the firm's age, a dummy for whether it pays a dividend, leverage, and firm size (e.g., Pastor and Veronesi, 2003; and Kemsley and Nissim, 2002).[15] The capital structure literature is a third example of where both the dependent variable and the independent variable are persistent and thus both contain a firm effect. In these papers, authors try to explain a firm's use of leverage by regressing the firm's debt-to-assets ratio on firm characteristics such as the firm's market-to-book ratio, the ratio of property, plant, and equipment to total assets, the earnings to assets, depreciation-to-asset ratio, R&D-to-assets ratio, and firm size (e.g., Baker and Wurgler, 2002; Fama and French, 2002; and Johnson, 2003).[16] As will be seen in Section 5, the serial correlation among these variables is quite large (usually greater than 0.95 after ten years). Since both the left- and right-hand-side variables in these three regressions are highly persistent, this is the kind of data where Fama-MacBeth standard errors may be biased. In Section 5.2, a capital structure regression is estimated and the magnitude of the bias is indeed shown to be large in the data set.

[14] The data in the simulations have a fixed effect in the independent variable, the residual, and thus the dependent variable. The first two are the source of the bias in the standard errors, which I discussed above. It is possible, however, for the dependent variable to contain a firm effect that comes exclusively from the independent variables. In this case, the residual would be serially uncorrelated and both OLS and Fama-MacBeth standard errors would be unbiased (Tables 1 and 2, row 1). Therefore, the only way for me to verify that the standard errors in any given paper are biased (e.g., standard errors clustered by firm would be larger) would be to recalculate the standard errors using a more robust method (e.g., standard errors clustered by firm).

[15] Both of these papers correct the Fama-MacBeth standard errors for first-order autocorrelation of the estimated slopes. I will discuss this method in Section 4.4. Pastor and Veronesi (2003) report that this correction does not change the estimated standard error. As I will show in Section 4.4, Fama-MacBeth standard errors adjusted for first-order autocorrelation can still produce biased standard errors. I will also explain why this adjustment may have very little effect on the estimated standard errors.

[16] Baker and Wurgler (2002) estimate both White and Fama-MacBeth standard errors but do not report the Fama-MacBeth standard errors since they are the same as the White standard errors. This is not surprising given the results of Section 1. The fact that the OLS and Fama-MacBeth standard errors are the same is consistent with no bias in the standard errors (e.g., no firm effect), as well as biased standard errors. In the presence of a firm effect, the bias in White and Fama-MacBeth standard errors will be very similar with longer panel data sets (Figure 3). Fama and French (2002) acknowledge that Fama-MacBeth standard errors may understate the true standard errors and so report adjusted Fama-MacBeth standard errors ("We use a less formal approach. We assume the standard errors of the average slopes ... should be inflated by a factor of 2.5"). I will discuss this method in Section 4.4 and show that it can generate biased standard errors as well.

The literature is a teaching tool. Authors read published papers to learn which econometric methods are appropriate in which situations. Thus when readers see published papers using Fama-MacBeth (or OLS/White) standard errors in the kinds of regressions that are listed here, they believe (incorrectly) that this approach is correct. The problem is actually worse. The published finance literature has not only used incorrect methods but also gone on to provide incorrect advice that states that the Fama-MacBeth approach corrects the standard errors for the residual correlation in the presence of a firm effect (e.g., $\rho_X \neq 0$ and $\rho_\varepsilon \neq 0$). Wu (2004, p. 111) uses "... the Fama and MacBeth (1973) method to account for the lack of independence because of multiple yearly observations per company." Denis, Denis, and Yost (2002, p. 1969) argue that the

> ... pooling of cross-sectional and time-series data in our tests creates a lack of independence in the regression models. This results in the deflated standard errors and, therefore, inflated *t*-statistics. To address the importance of this bias, we estimate the regression model separately for each of the 14 calendar years in our sample ... The coefficients and statistical significance of the other control variables are similar to those in the pooled cross-sectional, time series data.

Finally, Choe, Kho, and Stulz (2005, p. 814) explain that "The Fama-MacBeth regressions take into account the cross-correlations and the serial correlation in the error term, so that the *t*-statistics are much more conservative."

Fama-MacBeth standard errors do account for cross-correlation (e.g., correlations between ε_{it} and ε_{kt}), but they are not robust to serial correlation (e.g., correlation between ε_{it} and ε_{is}). In the presence of a firm effect, Fama-MacBeth and OLS/White standard errors are both biased, and as discussed above, the estimates can be quite close to each other even when the bias is large (compare Equations (7) and (12)). The problem is not with the Fama-MacBeth method, only with its use. It was developed to account for the correlation between observations on different firms in the same year, not to account for the correlation between observations on the same firm in different years. It is now being used and recommended in cases where it may produce biased estimates and overstated significance levels. Given the Fama-MacBeth approach was designed to deal with time effects in a panel data set, not firm effects, I will turn to this data structure in Section 2.[17]

1.6 Newey-West standard errors

An alternative approach for addressing the correlation of errors across observations is the Newey-West procedure (Newey and West, 1987). This procedure

[17] It may be the case that Fama and MacBeth understood the problem of applying their method to data that contain a firm effect in the residual and the independent variables. In their 1973 paper, Fama and MacBeth examine the serial correlation in the residuals and report that it is close to zero. This is consistent with no firm effect in the residuals and, as we will see in the next section, Fama-MacBeth standard errors being unbiased.

was initially designed to account for a serial correlation of unknown form in the residuals of a single time series. To be able to estimate the autocorrelations with a single time series, the Newey-West approach assumes that the correlation between residuals approaches zero as the distance between observations goes to infinity. In addition, Newey-West multiplies the covariance of lag j (e.g., $\varepsilon_t\ \varepsilon_{t-j}$) by the weight $[1 - j/(M + 1)]$, where M is the specified maximum lag. This weight is largest for adjacent observations, declines as the distance between observations increases, and grows to one asymptotically. Since the Newey-West procedure was originally designed for a single time series, the weighting function was necessary to make the estimate of this matrix positive semidefinite.

The Newey-West method for estimating standard errors has been modified for use in a panel data set by estimating only correlations between lagged residuals in the same cluster (see Brockman and Chung, 2001; MacKay, 2003; Bertrand, Duflo, and Mullainathan, 2004; and Doidge, 2004). The problem of choosing a lag length is simplified in a panel data set, since the maximum lag length is one less than the maximum number of years per firm.[18] By setting the maximum lag equal to $T - 1$, the central matrix in the variance equation of the Newey-West standard error is

$$\begin{aligned}\sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}\varepsilon_{it}\right)^2 &= \sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2 + 2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} w(t-s)X_{it}X_{is}\varepsilon_{it}\varepsilon_{is}\right)\\ &= \sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2 + 2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t} w(j)X_{it}X_{it-j}\varepsilon_{it}\varepsilon_{it-j}\right)\\ &= \sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2 + 2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t}\left(1-\frac{j}{T}\right) X_{it}X_{it-j}\varepsilon_{it}\varepsilon_{it-j}\right).\end{aligned} \tag{14}$$

To examine the relative performance of the Newey-West standard error estimates, I simulated five thousand data sets where the firm effect is fixed and assumed to account for 25% of the variability of both the independent variable and the residual. The standard error estimated by the Newey-West procedure is an increasing function of the lag length in the simulation since the autocorrelations do not die away with increasing lag length in the simulation. When the lag length is set to zero, the estimated standard error is numerically identical to the

[18] In the standard application of Newey-West, a lag length of M implies that the correlation between ε_t and ε_{t-k} is included for k running from $-M$ to M. When Newey-West has been applied to panel data sets, correlations between past and future values are only included when they are drawn from the same cluster. Thus a cluster that contains T years of data per firm uses a maximum lag length of $T - 1$ and would include $t - 1$ lags and $T - t$ leads for the t^{th} observation where t runs from 1 to T.

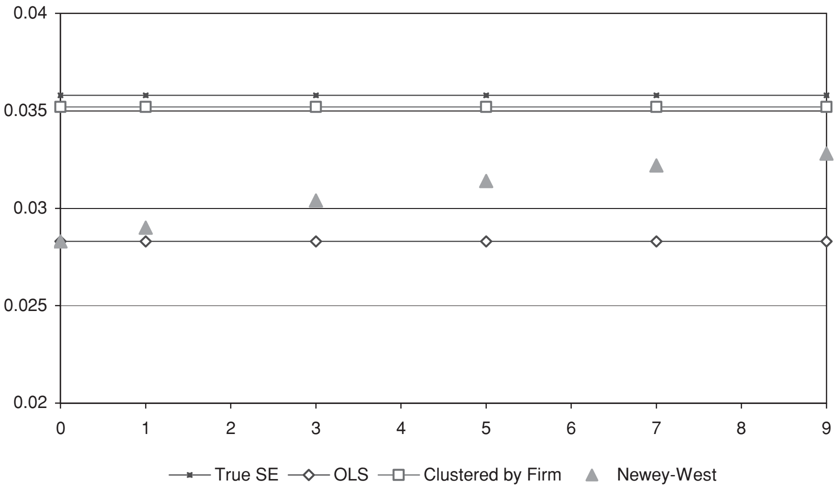

Figure 4
Relative performance of OLS, clustered, and Newey-West standard errors
The figure contains OLS standard errors, standard errors clustered by firm, and Newey-West standard errors, as well as the true standard error. Estimates are based on five thousand simulated data sets. Each data set contains five thousand observations (five hundrerd firms and ten years per firm). In each simulation, 25% of the variability in both the independent variable and the residual is due to a firm effect (i.e., $\rho_X = \rho_\varepsilon = 0.25$). The true standard error (shaded squares), the OLS standard error estimates (empty diamonds), and the clustered standard errors (empty squares) are plotted as straight lines as they do not depend upon the assumed lag length. The Newey-West standard error estimates, which rise with the assumed lag length, are plotted as triangles.

White standard error, which is only robust to heteroscedasticity (White, 1984) and equal to the OLS standard error in the simulation (Figure 4). As the lag length increases from 0 to 9, the standard error estimated by the Newey-West rises from the OLS/White estimate of 0.0283 to 0.0328 when the lag length is 9. In the presence of a fixed firm effect, an observation of a given firm is correlated with all observations for the same firm no matter how far apart in time the observations are spaced. Thus having a lag length of less than the maximum ($T-1$) will cause the Newey-West standard errors to underestimate the true standard error when the firm effect is fixed.

More interestingly is the fact that the Newey-West method underestimates the true standard error even when the maximum lag length is set to $T-1$. The bias in the Newey-West estimates can be traced to the weighting function. In a panel setting, the Newey-West standard error formula is identical to the clustered standard error formula except for the weighting function (compare Equations (7) and (14)). Since the Newey-West procedure weights the covariances by less than one, the estimated standard error is shrunk toward zero, which is what generates the bias. The weighting function is necessary to make the estimated variance matrix positive semi-definite when there is a single time series of data. However, in a panel data setting with multiple time series, the weighting function is not necessary and leads to a small bias in the estimated standard errors (the bias is 8% in the simulations and would have been smaller if the firm effect were not fixed).

2. Estimating Standard Errors in the Presence of a Time Effect

To demonstrate how the techniques work in the presence of a time effect, data sets containing only a time effect were generated (observations on different firms within the same year are correlated). This is the data structure for which the Fama-MacBeth approach was designed (see Fama and MacBeth, 1973). If it is assumed that the panel data structure contains only a time effect, the equations derived above are essentially unchanged. The expressions for the standard errors in the presence of only a time effect are correct once N and T are exchanged.

2.1 Clustered standard error estimates

Simulating the data with only a fixed time effect means that the dependent variable will still be specified by Equation (1), but now the error term and independent variable are specified as

$$\begin{aligned} \varepsilon_{it} &= \delta_t + \eta_{it} \\ X_{it} &= \zeta_t + \upsilon_{it}. \end{aligned} \tag{15}$$

As before, five thousand data sets of five thousand observations each are simulated. The fraction of variability in both the residual and the independent variable which is due to the time effect ranges from 0 to 75% in 25% increments. The OLS coefficient, the true standard error, the OLS and clustered standard errors, as well as the fraction of OLS and clustered t-statistics that are greater than 2.58 are reported in Table 3. There are several interesting findings to note. First, as with the firm effect results, the OLS standard errors are correct when there is no time effect in either the independent variable ($\text{Var}(\zeta) = 0$) or the residual ($\text{Var}(\delta) = 0$). As the time effect in the independent variable and the residual rise, so does the magnitude by which the OLS standard errors underestimate the true standard errors. When half of the variability in both comes from the time effect, the true standard error is eleven times the OLS estimate [10.7 = 0.3015/0.0282, see Table 3] and 81% of t-statistics are significant at the 1% level.

The clustered standard errors are much more accurate, but unlike the results with the firm effect, they underestimate the true standard error. The magnitude of the underestimate is small, ranging from 13% [1 − 0.1297/0.1490] when the time effect accounts for 25% of the variability to 19% [1 − 0.3986/0.4927] when the time effect accounts for 75% of the variability. The problem arises due to the limited number of clusters (e.g., years). When the standard errors were estimated in the presence of the firm effects, there were five hundred

Table 3
Estimating standard errors with a time effect: OLS and clustered standard errors

		Source of independent variable volatility			
Avg(β_{OLS}) Std(β_{OLS}) Avg(SE_{OLS}) % Sig(T_{OLS}) Avg(SE_C) % Sig(T_C)		0%	25%	50%	75%
Source of residual volatility	0%	1.0004 0.0286 0.0283 [0.0098] 0.0277 [0.0330]	1.0002 0.0291 0.0288 [0.0094] 0.0276 [0.0304]	1.0006 0.0293 0.0295 [0.0088] 0.0275 [0.0348]	0.9994 0.0314 0.0306 [0.0114] 0.0270 [0.0520]
	25%	1.0006 0.0284 0.0279 [0.0114] 0.0268 [0.0320]	1.0043 0.1490 0.0284 [0.6064] 0.1297 [0.0360]	0.9962 0.2148 0.0289 [0.7270] 0.1812 [0.0506]	0.9996 0.2874 0.0300 [0.7874] 0.2305 [0.0736]
	50%	0.9996 0.0276 0.0274 [0.0100] 0.0258 [0.0294]	0.9997 0.2138 0.0278 [0.7298] 0.1812 [0.0458]	0.9919 0.3015 0.0282 [0.8096] 0.2546 [0.0596]	1.0079 0.3986 0.0292 [0.8536] 0.3248 [0.0756]
	75%	1.0002 0.0273 0.0267 [0.0110] 0.0244 [0.0322]	0.9963 0.2620 0.0271 [0.7994] 0.2215 [0.0402]	0.9970 0.3816 0.0276 [0.8586] 0.3141 [0.0588]	0.9908 0.4927 0.0284 [0.8790] 0.3986 [0.0768]

The table contains estimates of the coefficient and standard errors based on five thousand simulated panel data sets, each of which contains five hundred firms and ten years per firm. The true slope coefficient is 1, the standard deviation of the independent variable is 1, and the standard deviation of the error term is 2 (see Equation (1)). The independent variable X and the residual are specified as

$$\begin{aligned} X_{it} &= \zeta_t + \upsilon_{it} \\ \varepsilon_{it} &= \delta_t + \eta_{it}. \end{aligned}$$

The fraction of X's variance, which is due to a time-specific component [Var(ζ)/Var(X)], varies across the columns of the table from 0% (no time effect) to 75% and the fraction of the residual variance, which is due to a time-specific component [Var(δ)/Var(ε)], varies across the rows of the table from 0% (no time effect) to 75%. Each cell contains the average estimated slope coefficient from OLS and the standard deviation of this estimate. This is the true standard error of the estimated coefficient. The third entry is the average standard error estimated by OLS. The percentage of OLS t-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets. The fifth entry is the average standard error clustered by year (i.e., accounts for the possible correlation between observations on different firms in the same year). The percentage of clustered t-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets.

firms (clusters). When the standard errors were estimated in the presence of a time effect, there were only ten years (clusters). Since the clustered standard error places no restriction on the correlation structure of the residuals within a cluster, its consistency depends on having a sufficient number of clusters.

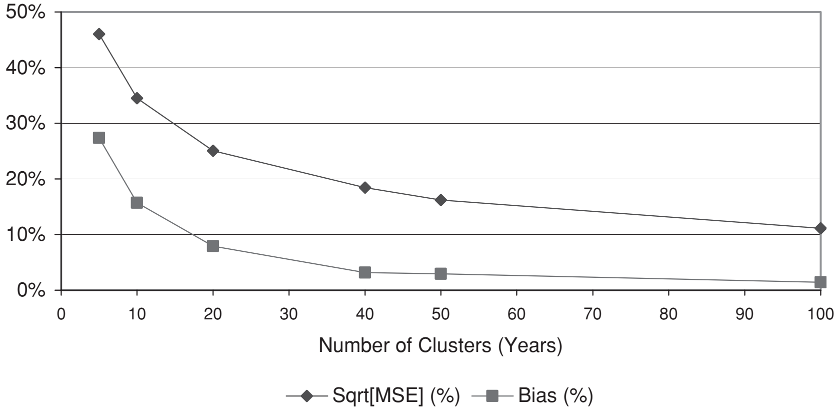

Figure 5
True standard errors and clustered standard errors as a function of cluster size (*T*)
The figure graphs the bias (squares) and mean squared error (*MSE*) (diamonds) as a function of the number of years (clusters) used in each simulation. The bias is the estimated clustered standard error minus the true standard error. The *MSE* is the average value of the squared difference between the estimated standard error and the true standard error:

$$\begin{aligned} MSE &= E[(\widehat{SE} - SE_{true})^2] \\ &= E[(\widehat{SE} - \overline{SE})^2 + (\overline{SE} - SE_{\mathrm{true}})^2] \\ &= Var(\widehat{SE}) + Bias(\widehat{SE})^2. \end{aligned}$$

The *MSE* is equal to the variance of the standard error plus the bias squared. Both the bias and the *MSE* are divided by the true standard error and thus are expressed as a percentage. Each simulated data set has five thousand observations. In each simulation, 25% of the variability in both the independent variable and the residual is due to the time effect (i.e., $\rho_X = \rho_\varepsilon = 0.25$). The standard errors are averaged across five thousand simulations. In these simulations, underestimation of the standard errors ranges from 27% when there are five years in the simulated data set to 1% when there are one hundred years in the simulated data set.

Based on these results, ten clusters is too small and five hundred is sufficient (see Kezdi, 2004; and Hansen, 2007).

To explore this issue, data sets of five thousand observations were simulated with the number of years (or clusters) ranging from 5 to 100. In all of the simulations, 25% of the variability in both the independent variable and the residual is due to the time effect [i.e., $\rho_X = \rho_\varepsilon = 0.25$]. The bias in the clustered standard error estimates declines with the number of clusters, dropping from 27% when there are five years (or clusters) to 3% when there are forty years to 1% when there are one hundred years (Figure 5). The standard deviation of the standard error estimates also declines as the number of clusters increases (holding the total sample size constant). Thus the mean squared error (MSE), which is a sum of the variance of the standard error estimate and the bias squared, declines with cluster size for both reasons.

2.2 Fama-MacBeth standard errors

When there is only a time effect, the correlation of the estimated slope coefficients across years is zero and the standard errors estimated by Fama-MacBeth

Table 4
Estimating standard errors with a time effect: Fama-MacBeth standard errors

		Source of independent variable volatility			
Avg(β_{FM}) Std(β_{FM}) Avg(SE_{FM}) % Sig(T_{FM})		0%	25%	50%	75%
Source of residual volatility	0%	1.0004 0.0287 0.0278 [0.0310]	1.0004 0.0331 0.0318 [0.0312]	1.0007 0.0396 0.0390 [0.0252]	0.9991 0.0573 0.0553 [0.0338]
	25%	1.0005 0.0252 0.0239 [0.0376]	1.0003 0.0284 0.0276 [0.0296]	1.0006 0.0343 0.0338 [0.0284]	0.9999 0.0496 0.0480 [0.0294]
	50%	1.0000 0.0200 0.0195 [0.0254]	1.0002 0.0231 0.0227 [0.0304]	1.0006 0.0280 0.0276 [0.0272]	1.0007 0.0394 0.0387 [0.0278]
	75%	1.0001 0.0142 0.0138 [0.0308]	0.9996 0.0161 0.0159 [0.0302]	1.0000 0.0200 0.0196 [0.0284]	0.9999 0.0285 0.0276 [0.0300]

The table contains estimates of the coefficient and standard errors based on the same five thousand simulated panel data sets that are used in Table 3. Each data set contains five hundred firms and ten years per firm. The true slope coefficient is 1, the standard deviation of the independent variable is 1, and the standard deviation of the error term is 2 (see Equation (1)). The independent variable X and the residual are specified as

$$\begin{aligned} X_{it} &= \zeta_t + \upsilon_{it} \\ \varepsilon_{it} &= \delta_t + \eta_{it}. \end{aligned}$$

The fraction of X's variance, which is due to a time-specific component [Var(ζ)/Var(X)], varies across the columns of the table from 0% (no time effect) to 75% and the fraction of the residual variance, which is due to a time-specific component [Var(δ)/Var(ε)], varies across the rows of the table from 0% (no time effect) to 75%. The first entry is the average slope coefficient based on a Fama-MacBeth estimation (e.g., the regression is run for each of the ten years and the estimate is the average of the ten estimated slope coefficients). The second entry is the standard deviation of the coefficient estimated by Fama-MacBeth. This is the true standard error of the Fama-MacBeth coefficient. The third entry is the average standard error estimated by the Fama-MacBeth procedure (e.g., Equation (9)). The percentage of Fama-MacBeth t-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets.

are unbiased (Equation (12)). This is what is found in the simulation (Table 4). The estimated standard errors are extremely close to the true standard errors and the number of statistically significant t-statistics is close to 3% across the simulations (using a 1% critical value).

3. Estimating Standard Errors in the Presence of a Fixed Firm and Time Effect

The best method for estimating standard errors in a panel data set depends on the source of dependence in the data. For panel data sets with only a firm effect, standard errors clustered by firm produce unbiased standard errors. If the data have only a time effect, the Fama-MacBeth estimates are better than standard errors clustered by time when there are few years (clusters) and equally good when the number of years (clusters) is sufficiently large. These methods allow us to be agnostic about the form of the correlation within a cluster. The cost, however, is that the residuals must be uncorrelated across clusters. For example, if we cluster by firm, we must assume there is no cross-sectional correlation (no time effect). As this assumption may be incorrect in some situations, I next consider a data structure with both a firm and a time effect.

One way empirical finance researchers can address two sources of correlation is to parametrically estimate one of the dimensions (e.g., by including dummy variables). Since many panel data sets have more firms than years, a common approach is to include dummy variables for each time period (to absorb the time effect) and then cluster by firm (Lamont and Polk, 2001; Anderson and Reeb, 2004; Gross and Souleles, 2004; Sapienza, 2004; and Faulkender and Petersen, 2006). If the time effect is fixed (e.g., Equation (15)), the time dummies completely remove the correlation between observations in the same time period. In this case, there is only a firm effect left in the data. As seen in Section 1, OLS and Fama-MacBeth standard errors are biased in this case, while standard errors clustered by firm are unbiased (results available from the author).

The parametric approach only works when the dependence is correctly specified. If the time effect is not fixed, then time dummies will not remove the dependence completely and even standard errors clustered by firm can be biased. I will return to this issue in more detail in Sections 4 and 5. Since researchers do not always know the precise form of the dependence, a less parametric approach may be preferred. A solution is to cluster on two dimensions simultaneously (e.g., firm and time). Cameron, Gelbach, and Miller (2006), and Thompson (2006) proposed the following estimate of the variance-covariance matrix:

$$V_{\text{Firm\&Time}} = V_{\text{Firm}} + V_{\text{Time}} - V_{\text{White}}, \tag{16}$$

which combines the standard errors clustered by firm with the standard errors clustered by time. The standard errors clustered by firm (the first term) capture the unspecified correlation between observations on the same firm in different years (e.g., correlations between ε_{it} and ε_{is}). The standard errors clustered by time (the second term) capture the unspecified correlation between observations on different firms in the same year (e.g., correlations between ε_{it} and ε_{kt}). Since both the firm- and time-clustered variance-covariance matrix include the diagonal of the variance-covariance matrix, the White variance-covariance matrix is subtracted off to avoid double counting these terms.[19]

This method allows for both a firm and a time effect, although observations on different firms in different years are assumed to be uncorrelated (Figure 6). As with standard errors clustered on one dimension, this approach is unbiased as

[19] In some settings (e.g., clustering by industry and year), there can be multiple observations (firms) per industry-year. In this case, the third matrix that is subtracted off in Equation (16) is the variance-covariance matrix clustered by industry-year (see Cameron, Gelbach, and Miller, 2006).

	Firm 1			Firm 2			Firm 3		
Firm 1	ε_{11}^2	$\varepsilon_{11}\,\varepsilon_{12}$	$\varepsilon_{11}\,\varepsilon_{13}$	$\varepsilon_{11}\,\varepsilon_{21}$	0	0	$\varepsilon_{11}\,\varepsilon_{31}$	0	0
	$\varepsilon_{12}\,\varepsilon_{11}$	ε_{12}^2	$\varepsilon_{12}\,\varepsilon_{13}$	0	$\varepsilon_{12}\,\varepsilon_{22}$	0	0	$\varepsilon_{12}\,\varepsilon_{32}$	0
	$\varepsilon_{13}\,\varepsilon_{11}$	$\varepsilon_{13}\,\varepsilon_{12}$	ε_{13}^2	0	0	$\varepsilon_{13}\,\varepsilon_{23}$	0	0	$\varepsilon_{13}\,\varepsilon_{33}$
Firm 2	$\varepsilon_{21}\,\varepsilon_{11}$	0	0	ε_{21}^2	$\varepsilon_{21}\,\varepsilon_{22}$	$\varepsilon_{21}\,\varepsilon_{23}$	$\varepsilon_{21}\,\varepsilon_{31}$	0	0
	0	$\varepsilon_{22}\,\varepsilon_{12}$	0	$\varepsilon_{22}\,\varepsilon_{21}$	ε_{22}^2	$\varepsilon_{22}\,\varepsilon_{23}$	0	$\varepsilon_{22}\,\varepsilon_{32}$	0
	0	0	$\varepsilon_{23}\,\varepsilon_{13}$	$\varepsilon_{23}\,\varepsilon_{21}$	$\varepsilon_{23}\,\varepsilon_{22}$	ε_{23}^2	0	0	$\varepsilon_{23}\,\varepsilon_{33}$
Firm 3	$\varepsilon_{31}\,\varepsilon_{11}$	0	0	$\varepsilon_{31}\,\varepsilon_{21}$	0	0	ε_{31}^2	$\varepsilon_{31}\,\varepsilon_{32}$	$\varepsilon_{31}\,\varepsilon_{33}$
	0	$\varepsilon_{32}\,\varepsilon_{12}$	0	0	$\varepsilon_{32}\,\varepsilon_{22}$	0	$\varepsilon_{32}\,\varepsilon_{31}$	ε_{32}^2	$\varepsilon_{32}\,\varepsilon_{33}$
	0	0	$\varepsilon_{33}\,\varepsilon_{13}$	0	0	$\varepsilon_{33}\,\varepsilon_{23}$	$\varepsilon_{33}\,\varepsilon_{31}$	$\varepsilon_{33}\,\varepsilon_{32}$	ε_{33}^2

Figure 6
Residual cross product matrix: Firm and time effects
This figure shows a sample covariance matrix of the residuals. When standard errors are clustered by both firm and time, residuals of the same firm in different years, as well as residuals of the same year but on different firms, may be nonzero. Observations on different firms and different years are assumed to be zero and are reported as zero in the matrix.

long as there are a sufficient number of clusters, in this case, both enough firms and enough time periods (see Thompson, 2006). To illustrate the performance of standard errors clustered by firm, year, or both, data sets with a fixed firm and time effect are simulated:[20]

$$\begin{aligned} \varepsilon_{it} &= \gamma_i + \delta_t + \eta_{it} \\ X_{it} &= \mu_i + \zeta_t + \upsilon_{it}. \end{aligned} \tag{17}$$

One-third of the variability of the residual and the independent variable is due to the firm effect and one-third of the variability is due to the time effect [e.g., Var(γ) = Var(δ) = Var(η) and Var(μ) = Var(ζ) = Var(υ)]. Nine data sets are then simulated where the number of firms and time periods range from 10 to 1,000 so that the total number of observations is always 10,000 (e.g., 250 firms and 40 time periods, see Figure 7). Standard errors clustered by only one dimension are biased downward and produce confidence intervals that are too small. The magnitude of this bias varies widely depending upon the number of clusters. For example, the fraction of t-statistics clustered by time that are statistically significant at the 1% level (greater than 2.58) ranges from 73% when there are one thousand time periods (and ten firms) to 5% when there are only ten time periods (and one thousand firms).

[20] Although the firm and time effect are assumed to be fixed in the simulation, this is only for illustration. Clustered standard errors, whether we cluster on one or more dimensions, are robust to any form of within-cluster correlation. In Section 4, I will examine the performance of clustered standard errors when the firm effect is temporary (dies away as the time between observations grows). The reader can also refer to Cameron, Gelbach, and Miller (2006), and Thompson (2006) for results on standard errors clustered on more than one dimension when the firm and/or time effects are not fixed. Cameron, Gelbach, and Miller also generalize the procedure to allow for clustering on more than two dimensions.

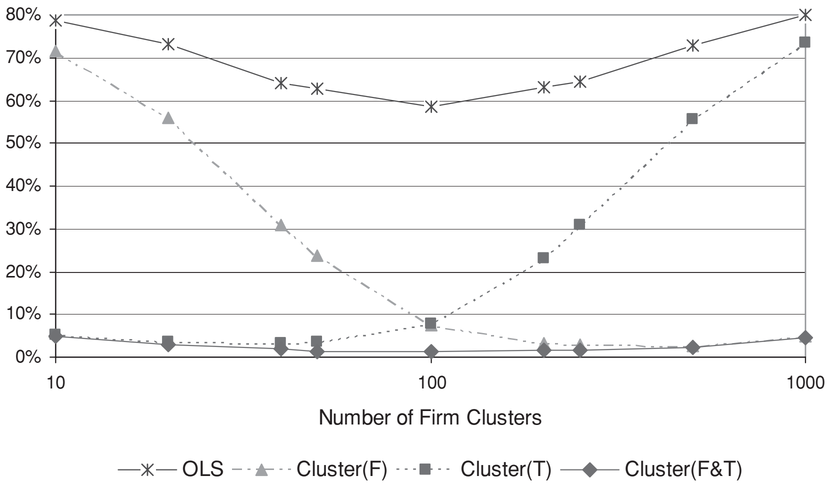

Figure 7
Rejection rates in the presence of a firm and a time effect for *t*-statistics clustered by firm, by time, or both
The figure graphs the fraction of *t*-statistics that are statistically significant at the 1% level (greater than 2.58 in absolute value). The number of firms and time periods range from 10 to 1,000 so that the total number of observations is always 10,000. The number of firms increases from 10 to 1,000 as we move left to right across the figure, while the number of time periods decreases from 1,000 to 10 as we move left to right across the figure. Thus the number of firms and the number of years in each of the nine simulations are (10, 1000), (20, 500), (40, 250), (50, 200), (100, 100), (200, 50), (250, 40), (500, 20), (1000, 10). In each simulation, OLS standard errors (stars), standard errors clustered by firm (triangles), standard errors clustered by year (squares), and standard errors clustered by firm and year (diamonds) are reported.

Clustering by two dimensions produces less biased standard errors. However, clustering by firm and time does not always yield unbiased estimates. When there are one hundred firms and one hundred years, 1% of the *t*-statistics are greater than 2.58. As the number of clusters—firms or years—declines, the standard errors clustered by firm and time are biased, although the magnitude of the bias is not large. In the simulations, the number of *t*-statistics that are greater than 2.58 rises to 5% when the number of firms or time periods falls to 10 (see Cameron, Gelbach, and Miller, 2006; and Thompson, 2006 for additional results). When there are only a few clusters in one dimension, clustering by the more frequent cluster yields results that are almost identical to clustering by both firm and time. For example, in the simulation with one thousand firms and ten years of data, the percentage of *t*-statistics that are greater than 2.58 is 5% whether the standard errors are clustered by firm or by firm and time (Figure 7).

4. Estimating Standard Errors in the Presence of a Temporary Firm Effect

The analysis thus far has assumed that the firm effect is fixed. Although this is common in the literature, it may not always be true in the data. The dependence between residuals may decay as the time between them increases [e.g., $\rho(\varepsilon_t, \varepsilon_{t-k})$ may decline with k]. In a panel with a short time series, distinguishing between a permanent and a temporary firm effect may be impossible. However, as the number of years in the panel increases, it may be feasible to empirically identify the permanence of the firm effect. In addition, if the performance of the different standard error estimates depends on the permanence of the firm effect, researchers need to know this.

4.1 Temporary firm effects: specifying the data structure

To explore the performance of the different standard error estimates in a more general context, a data structure is simulated that includes both a permanent component (a fixed firm effect) and a temporary component (nonfixed firm effect) that is assumed to be a first-order autoregressive process. This allows the firm effect to die away at a rate between a first-order autoregressive decay and zero. To construct the data, the nonfirm effect portion of the residual (η_{it} from Equation (4)) is specified as

$$\begin{aligned}\eta_{it} &= \zeta_{it} \quad \text{if } t = 1\\ &= \phi\eta_{it-1} + \sqrt{1-\phi^2}\zeta_{it \text{ if } t>1}.\end{aligned} \tag{18}$$

ϕ is the first-order auto correlation between η_{it} and η_{it-1}, and the correlation between η_{it} and η_{it-k} is ϕ^{k}.[21] Combining this term with the fixed firm effect (γ_i in Equation (4)) means that the serial correlation of the residuals dies off over time, but more slowly than implied by a first-order autoregressive process and asymptotes to ρ_ε (from Equation (6)). By choosing the relative magnitude of the fixed firm effect (ρ_ε) and the first-order auto correlation (ϕ), the pattern of auto correlations in the residual can be altered. The correlation of lag length k is

$$\begin{aligned}Corr(\varepsilon_{i,t}, \varepsilon_{i,t-k}) &= \frac{Cov(\gamma_i + \eta_{i,t}, \gamma_i + \eta_{i,t-k})}{\sqrt{Var(\gamma_i + \eta_{i,t})Var(\gamma_i + \eta_{i,t-k})}}\\ &= \frac{\sigma_\gamma^2 + \phi^k\sigma_\eta^2}{\sigma_\gamma^2 + \sigma_\eta^2}\\ &= \rho_\varepsilon + (1-\rho_\varepsilon)\phi^k.\end{aligned} \tag{19}$$

[21] I multiply the ς term by $\sqrt{(1-\phi^2)}$ to make the residuals homoscedastic. From Equation (18),

$$\begin{aligned}Var(\eta_{it}) &= \sigma_\zeta^2 \text{ if } t = 1\\ &= \phi^2\sigma_\zeta^2 + (1-\phi^2)\sigma_\zeta^2 = \sigma_\zeta^2 \text{ if } t > 1,\end{aligned}$$

where the last step is by recursion (if it is true for $t = m$, it is true for $t = m + 1$). Assuming homoscedastic residuals is not necessary since the Fama-MacBeth and clustered standard errors are robust to heteroscedasticity (Liang and Zeger, 1986; Moulton, 1986, 1990; Arellano, 1987; Andrews, 1991; Rogers, 1993; Jagannathan and Wang, 1998, and Williams, 2000). However, assuming homoscedasticity makes the interpretation of the results simpler. Any difference in the standard errors I find is due to the dependence of observations within a cluster, not heteroscedasticity.

Table 5
Estimated standard errors with a nonfixed firm effect

Panel A: OLS and clustered standard errors

Avg(β_{OLS}) Std(β_{OLS}) Avg(SE_{OLS}) % Sig(T_{OLS}) Avg(SE_C) % Sig(T_C)	I	II	III	IV
ρ_X/ρ_ε	0.50/0.50	0.00/0.00	0.25/0.25	0.60/0.35
ϕ_X/ϕ_ε	0.00/0.00	0.90/0.90	0.75/0.75	0.99/0.81
OLS	0.9994	1.0001	1.0009	0.9991
	0.0513	0.0659	0.0566	0.0677
	0.0283	0.0283	0.0283	0.0253
	[0.1578]	[0.2746]	[0.1996]	[0.3302]
	0.0508	0.0668	0.0569	0.0670
	[0.0114]	[0.0086]	[0.0104]	[0.0098]
OLS with firm dummies	1.0007	1.0003	1.0013	1.0046
	0.0299	0.0517	0.0442	0.1881
	0.0298	0.0298	0.0298	0.1101
	[0.0096]	[0.1382]	[0.0802]	[0.1288]
	0.0298	0.0516	0.0441	0.1886
	[0.0100]	[0.0098]	[0.0092]	[0.0108]

Panel B: GLS estimates with and without clustered standard errors

Avg(β_{GLS}) Std(β_{GLS}) Avg(SE_{GLS}) % Sig(T_{GLS}) Avg(SE_{GLS-C}) % Sig(T_{GLS-C})	I	II	III	IV
ρ_X/ρ_ε	0.50/0.50	0.00/0.00	0.25/0.25	0.60/0.35
ϕ_X/ϕ_ε	0.00/0.00	0.90/0.90	0.75/0.75	0.99/0.81
GLS	1.0005	1.0003	1.0012	1.0006
	0.0284	0.0475	0.0408	0.0731
	0.0283	0.0283	0.0283	0.0580
	[0.0090]	[0.1240]	[0.0730]	[0.0388]
	0.0282	0.0474	0.0408	0.0721
	[0.0090]	[0.0100]	[0.0090]	[0.0112]

The same data structure is specified for the independent variable in the first three columns of Table 5.[22] The correlation structures range from a fixed firm effect ($\rho = 0.50$ and $\phi = 0.00$) to a standard AR1 process ($\rho = 0.00$ and $\phi = 0.90$).

[22] In column IV, I use different data generating processes for the independent variable and the residual. The parameters were chosen to match the first- and tenth-order autocorrelation of the residuals and the independent variables from the capital structure regression that I examine in Section 5 (Table 7). These autocorrelations are graphed in Figure 10A.

Table 5
(Continued)

Panel C: Fama-MacBeth standard errors

$Avg(\beta_{FM})$ $Std(\beta_{FM})$ $Avg(SE_{FM})$ % $Sig(T_{FM})$ $Avg(SE_{FM-AR1})$ % $Sig(T_{FM-AR1})$	I	II	III	IV
ρ_X/ρ_ε	0.50/0.50	0.00/0.00	0.25/0.25	0.60/0.35
ϕ_X/ϕ_ε	0.00/0.00	0.90/0.90	0.75/0.75	0.99/0.81
Fama-MacBeth	0.9995	1.0001	1.0008	0.9991
	0.0514	0.0660	0.0567	0.0667
	0.0239	0.0187	0.0221	0.0138
	[0.2510]	[0.4696]	[0.3350]	[0.6094]
	0.0224	0.0389	0.0376	0.0289
	[0.3222]	[0.2876]	[0.2098]	[0.3900]
Avg(first-order autocorrelation)	−0.1157	0.4395	0.3250	0.4389

The table contains estimates of the coefficient and standard errors based on five thousand simulated panel data sets, each of which contains five hundred firms and ten years per firm. The true slope coefficient is 1, the standard deviation of the independent variable is 1, and the standard deviation of the error term is 2. Across the columns the magnitude of the fixed firm effect (ρ) and the first-order autocorrelation (ϕ) varies. ρ_X ($\rho\varepsilon$) is the fraction of the independent variable's (residual's) variance that is due to the fixed firm effect. ϕ_x ($\phi\varepsilon$) is the first-order autocorrelation of the nonfixed portion of the firm effect of the independent variable (residual). Combining Equation (4) with Equation (18), the residual is specified as

$$\begin{aligned}\varepsilon_{it} &= \gamma_i + \eta_{it} = \gamma_i + \zeta_{it} \quad \text{if } t = 1 \\ &= \gamma_i + \phi\,\eta_{it-1} + \sqrt{1-\phi^2}\zeta_{it} \quad \text{if } t > 1.\end{aligned}$$

The independent variable is specified in a similar manner.

Panel A contains coefficients estimated by OLS. In the first row, only the independent variable (*X*) is included; in the second row, firm dummies are also included in the regression. The first two entries in each cell contain the average slope estimated by OLS and the standard deviation of the coefficient (i.e., the true standard error). The third entry is the average standard error estimated by OLS. The percentage of OLS *t*-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets. The fifth entry is the average standard error clustered by firm. The percentage of clustered *t*-statistics that are significant at the 1% level is reported in square brackets.

Panel B contains coefficients and standard errors estimates of the random effects model using feasible generalized least-squares (FGLS). The first two entries in each cell contain the average slope estimated by GLS and the standard deviation of the coefficient (i.e., the true standard error). The third entry is the regular standard error estimated by GLS (i.e., not clustered). The percentage of regular GLS *t*-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets. The fifth entry is the average standard error clustered by firm. The percentage of clustered *t*-statistics that are significant at the 1% level is reported in square brackets.

Panel C contains coefficients and standard errors estimated by Fama-MacBeth. The first two entries in each cell contain the average slope estimated by Fama-MacBeth and the standard deviation of the coefficient (i.e., the true standard error). The third entry is the average standard error estimated by the Fama-MacBeth procedure (see Equation (9)). The percentage of *t*-statistics that are significant at the 1% level (e.g., $|t|>2.58$) is reported in square brackets. The fifth and sixth entries are the Fama-MacBeth standard error corrected for the first-order autocorrelation and the percentage of *t*-statistics that are significant at the 1% level. I adjust the standard error by multiplying the traditional Fama-Macbeth standard error (Equation (9)) by the square root of $(1+\theta)/(1-\theta)$), where θ is the first-order autocorrelation of β_t and β_{t-1}. The average first-order autocorrelation is reported in the last row of panel C.

4.2 Fixed effects—firm dummies

A significant minority of the papers in the survey include firm dummies in their regressions. Using the simulations, the relative performance of OLS and clustered standard errors can be compared both with and without firm dummies. The results are reported in Table 5, panel A. The fixed effect estimates are more efficient in this case (e.g., 0.0299 versus 0.0513), although this is not always true (for an example, see Table 5, panel A, column IV). Once the firm effects are included, the OLS standard errors are unbiased (Table 5, panel A, column I).[23] The clustered standard errors are unbiased with and without the firm dummies (see Kezdi, 2004 for examples where the clustered standard errors are too large in a fixed effect model). This conclusion, however, depends on the firm effect being fixed. If the firm effect decays over time, the firm dummies no longer fully capture the within-cluster dependence and OLS standard errors are still biased (Table 5, panel A, columns II–IV). In these simulations, the firm effect decays over time (in column II, 61% of the firm effect dissipates after nine years). Once the firm effect is temporary, the OLS standard errors again underestimate the true standard errors even when firm dummies are included in the regression (Baker, Stein, and Wurgler, 2003; and Wooldridge, 2003).[24] The magnitude of the underestimation depends on the magnitude of the temporary component of the firm effect (i.e., ϕ), ranging from 33% in column III to 42% in column II.

4.3 Generalized least-squares estimates of the random effects model

When the residuals of a panel regression are correlated, not only OLS standard error estimates are biased but also the coefficient estimates are inefficient (the estimates do not exploit all of the information in the data). Researchers can improve efficiency by estimating a random effects model using a GLS approach (i.e., a panel data set with an unobserved firm effect, see Wooldridge, 2007). This is rarely done in the finance literature. Less than 3% of the papers in the survey used this method (for examples, see Maksimovic and Phillips, 2002; Gentry, Kemsley, and Mayer, 2003; and Almazan et al., 2004). To test the performance of this model, the random effects model was estimated using feasible generalized least-squares (FGLS) in each of the four simulations reported in Table 5, panel B.

[23] I have assumed the model is correctly specified [i.e., $\mathrm{Corr}(X_{it}, \varepsilon_{it}) = 0$]. In this case, the only purpose of including firm dummies is to correct the standard errors. In practice, the model may not be correctly specified [i.e., $\mathrm{Corr}(X_{it}, \varepsilon_{it}) \neq 0$], and so including fixed effects would also be necessary to test the model's specification (see Hausman, 1978). Instead of including firm dummies, we could have first differenced the data within firm. However, it would still be necessary to use clustered as opposed to OLS standard errors, since the residuals would be correlated.

[24] Similar intuition applies to the standard errors of a difference-in-difference estimation (DID). In a DID estimation, the researcher has a panel data set with observations on different agents (e.g., companies) before and after a treatment such as a law change (see Bertrand, Duflo, and Mullainathan, 2004). By first-differencing the data (e.g., $\varepsilon_{it} - \varepsilon_{it-1}$), the objective is to remove the firm effect. By then differencing these changes across different companies in the same year (e.g., $[\varepsilon_{it} - \varepsilon_{it-1}]$–$[\varepsilon_{kt} - \varepsilon_{kt-1}]$, the objective is to remove the time effect. This method completely removes the firm and time effect only when they are permanent (i.e., Equations (4), (5), and (15)). When the firm and/or time effect are not permanent, than traditional OLS standard errors will be biased (for an example with only a firm effect, see Table 5, panel A). In this case, it is still necessary to estimate clustered standard errors when estimating a difference-in-difference model.

The GLS estimates are more efficient than the OLS estimates (both with or without firm dummies) when the residuals are correlated (compare Table 5, panels A and B, columns I–III). The standard errors produced by GLS, however, are unbiased only when the firm effect is permanent (e.g., column I). When the residuals are correlated, but the correlation dies away, the GLS coefficient estimates are still more efficient than the OLS estimates, but the standard errors are no longer unbiased (Table 5, panel B, columns II–IV). As with the OLS standard errors, the GLS standard errors are too small, although the magnitude of the bias is smaller. Thus, it is necessary to estimate standard errors clustered by firm even when using GLS (except when the firm effect is fixed).[25]

4.4 Adjusted Fama-MacBeth standard errors

As shown in Section 1, the presence of a firm effect causes the Fama-MacBeth standard error to be biased downward. Many authors have acknowledged the bias and have suggested adjusting the standard errors for the estimated first-order autocorrelation of the estimated slope coefficients (Christopherson, Ferson, and Glassman, 1998; Graham, Lemmon, and Schallheim, 1998; Chen, Hong, and Stein, 2001; Cochrane, 2001; Lakonishok and Lee, 2001; Fama and French, 2002; Kemsley and Nissim, 2002; Bakshi, Kapadia, and Madan, 2003; Pastor and Veronesi, 2003; Chakravarty, Gulen, and Mayhew, 2004; Nagel, 2005; and Schultz and Loughran, 2005). The proposed adjustment is to estimate the correlation between the yearly coefficient estimates (i.e., Corr$[\beta_{t,}\ \beta_{t-1}] = \theta$), and then multiply the estimated variance by $(1 + \theta)/(1 - \theta)$ to account for the serial correlation of the βs (see Chakravarty, Gulen, Mayhew, 2004; and Fama and French, 2002, footnote 1).[26] This would seem to make intuitive sense since the presence of a firm effect causes the yearly coefficient estimates to be serially correlated.

The same four simulated data structures as above are used to test the merits of this idea (Table 5). In each data set, the ten cross-sectional slope coefficients and the first-order autocorrelation of the slope coefficients are estimated, then the original and an adjusted Fama-MacBeth standard error are calculated. In the

[25] To understand why the regular GLS standard errors are biased, we need to examine how the GLS estimator of the random effects model is constructed. The GLS estimates are a matrix weighted average of the between and within estimates (Greene, 2002). The between estimates are obtained by running the regression on firm means [e.g., regress the mean value of Y for each firm on the mean value of X for each firm, $Y_{i*} = \beta_{\text{Between}} X_{i*} + \varepsilon_{i*}$]. Since the between regression contains only one observation per firm, there is no within-cluster correlation. The within estimates are obtained by running a regression of the deviations from firm means [e.g., $Y_{it} - Y_{i*} = \beta_{\text{Within}} (X_{it} - X_{i*}) + (\varepsilon_{it} - \varepsilon_{i*})$]. This is identical to including firm dummies. If the firm effect is entirely fixed, then there is no within-cluster correlation in the residuals of this regression either. However, if the firm effect is temporary, then the residuals are still correlated within the cluster and this is the source of the bias in the standard errors. This is why the regular GLS standard errors are correct only when the firm effect is fixed. The intuition is exactly the same as with the firm dummy estimates (compare panel A to panel B in Table 5).

[26] The literature has used two alternatives that are conceptually similar. Instead of using the infinite period adjustment (e.g., $[(1 + \theta)/(1 - \theta)]$), some papers have used a finite period adjustment. Given T years per firm, the correction is

$$Variance\ correction = \left(1 + 2\sum_{k=1}^{T-1}(T - k)\theta^k\right).$$

first column of Table 5, the fixed firm effect accounts for 50% of the variance. The autocorrelation is estimated very imprecisely as noted by Fama and French (2002). The ninetieth percentile confidence interval ranges from −0.60 to 0.41 and the mean autocorrelation is −0.12 (Table 5, panel C, column I). Since the average first-order autocorrelation is negative, the adjusted Fama-MacBeth standard errors are even more biased than the unadjusted standard errors. The fraction of *t*-statistics that is greater than 2.58 in absolute value rises from 25% to 32%.

The problem is that the correlation being estimated (the within sample autocorrelation of βs) is not the same as the one that is causing the bias in the standard errors (the population autocorrelation of βs). The covariance that biases the standard errors and that can be estimated across the five thousand simulations is

$$Cov(\beta_t, \beta_{t-1}) = E[(\beta_t - \beta_{True})(\beta_{t-1} - \beta_{True})]. \tag{20}$$

To see how the presence of a fixed firm effect influences this covariance, consider the case where the realization for firm i is a positive value of $\mu_i\gamma_i$ (i.e., the realized firm effect in the independent variable and the residual). This positive realization will result in an above average estimate of the slope coefficient in year t, and because the firm effect is fixed, it will also result in the same above average estimate of the slope coefficient in year $t-1$ (Equations (4), (5), and (8)). The realized value of the firm effect (μ_i and γ_i) in a given simulation does not change the average β across samples. The average β across data samples is the true β (one in the simulations). Thus, when the true correlation between β_t and β_{t-1} is estimated, the firm effect causes this correlation to be positive and the Fama-MacBeth standard errors to be biased downward.[27]

Researchers, however, are given only one data set. They must calculate the serial correlation of the βs within the sample they are given. This covariance is calculated as

$$Cov(\beta_t, \beta_{t-1}) = E[(\beta_t - \bar{\beta}_{\text{Within sample}})(\beta_{t-1} - \bar{\beta}_{\text{Within sample}})]. \tag{21}$$

The within-sample serial correlation measures the tendency of β_t to be above the within-sample mean when β_{t-1} is above the within-sample mean. To see how the presence of a fixed firm effect influences this covariance, consider the same case as above. A positive realization of $\mu_i\gamma_i$ will raise the estimate of β_1 through β_T, as well as the average of the βs by the same amount. Thus, a fixed firm effect will not influence the deviation of any β_t from the sample average β. The estimated serial correlation is asymptotically zero when the

[27] In the simulation, the covariance between β_t and β_s ranged from 0.0018 to 0.0023 and did not decline as the difference between t and s increased ($\rho_X = \rho_\varepsilon = 0.50$). The theoretical value of the covariance between β_t and β_s should be 0.0020 (according to Equation (11)) and would imply a true standard error of the Fama-MacBeth estimate of 0.0510 (according to Equation (12)). This matches the number reported in Table 2.

firm effect is constant and adjusting the standard errors based on this estimated serial correlation will still lead to biased standard error estimates (see Pastor and Veronesi, 2003 for an example).[28]

The adjusted Fama-MacBeth standard errors do better when there is an autoregressive component in the residuals (i.e., $\phi > 0$). In the three remaining simulations (Table 5, panel C), the estimated within-sample autocorrelation is positive, but the adjusted Fama-MacBeth standard errors are still biased downward. For example, in column II of Table 5, adjusting the standard errors reduces the bias from 72% to 41% (e.g., from $1 - 0.0187/0.0660$ to $1 - 0.0389/0.0660$); 29% of the *t*-statistics are still significant at the 1% level (Table 5, panel C, column II). The adjusted Fama-MacBeth standard errors do best when the firm effect dies off fast enough and the researcher has a sufficient number of time periods per firm. To show this, the simulations for the data structures in Table 5 are rerun with the number of time periods per firm varying from 5 to 50. When the firm effect is fixed, the adjusted Fama-MacBeth standard errors do no better than the unadjusted standard errors, even as the number of time periods grows (Figure 8). When the firm effect is a first-order autoregressive process (e.g., Table 5, column II), the adjusted Fama-MacBeth standard errors do very well. The percentage of *t*-statistics that is significant at the 1% level drops to 7% when there are fifty time periods per firm. When the firm effect has both a temporary and a fixed component (as found in the real world examples in the next section), the adjusted Fama-MacBeth standard errors are still biased, but significantly less biased than the unadjusted Fama-MacBeth standard errors. With fifty time periods per firm, 28% of the *t*-statistics are significant at the 1% level (Figure 8).

5. Empirical Applications

Simulated data are used in the previous sections. Thus, the data structure is known, which made choosing the method for estimating standard errors much easier. In real world applications, researchers may have priors about the structure of data (are firm effects or time effects more important, are they permanent or temporary), but the data structure is not known with certainty. Thus in this section, several techniques for estimating standard errors are applied to two real data sets. This allows demonstration of how the different methods for estimating standard errors compare, confirms that the methods used by some published papers could have produced standard errors estimates whose bias is large, and shows what can be learned from the different standard errors estimates.

[28] The average within-sample serial correlation which is estimated in Table 5, column I is actually less than zero, but this is due to a small sample bias. With only ten years of data per firm, there are only nine observations to estimate the serial correlation. To verify that this is correct, the simulation was rerun using five to fifty years of data. As the number of years increases, the adjusted Fama-MacBeth standard error converges to the unadjusted Fama-MacBeth standard error (Figure 8).

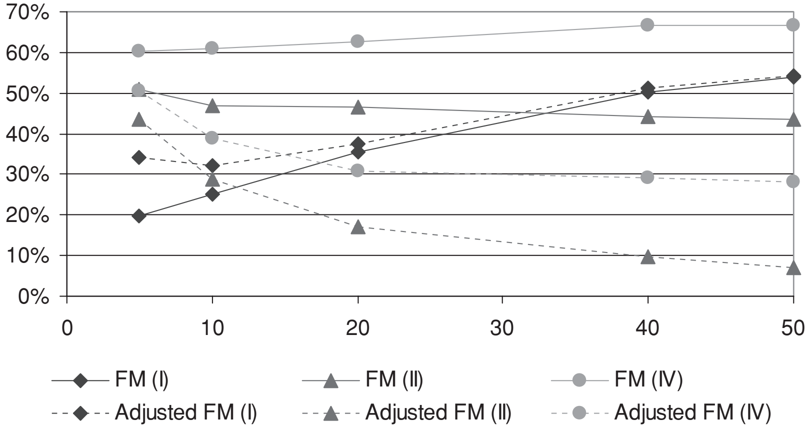

Figure 8
Rejection rates based on unadjusted and adjusted Fama-MacBeth standard errors
The figure graphs the fraction of t-statistics that are statistically significant at the 1% level (greater than 2.58 in absolute value) as a function of number of time periods per firm. The rejection rates are graphed for both standard (unadjusted) Fama-MacBeth t-statistics and Fama-MacBeth t-statistics adjusted for the first-order serial correlation. The rejection rates are based on three data structures from Table 5. The rejection rates for data with only a fixed firm effect ($\rho = 0.50$ and $\phi = 0.00$, column I in Table 5) are graphed with diamonds. Rejection rates for data with only a first-order autocorrelation ($\rho = 0.00$ and $\phi = 0.90$, column II in Table 5) are graphed with triangles. A data structure with both a fixed firm effect and an autoregressive firm effect ($\rho_X = 0.60$, $\phi_X = 0.99$, $\rho_\varepsilon = 0.35$, $\phi_\varepsilon = 0.81$, column IV in Table 5) is graphed with circles. The symbols with solid lines are the rejection rates based on unadjusted Fama-MacBeth t-statistics, while the symbols with dotted lines are the rejection rates based on adjusted Fama-MacBeth t-statistics.

For both data sets, the regression is first estimated using OLS, and White standard errors as well as standard errors clustered by firm, by time, and by both are reported (Tables 6 and 7, columns I–IV). By using White standard errors as the baseline, differences across columns are due only to within-cluster correlations, not to heteroscedasticity. If the standard errors clustered by firm are dramatically different than the White standard errors, then there is a significant firm effect in the data [e.g., Corr($X_{it}\, \varepsilon_{it}$, $X_{it-k}\, \varepsilon_{it-k}$) $\neq 0$]. The slope coefficients and the standard errors are then estimated using Fama-MacBeth (Tables 6 and 7, columns V). Each of the OLS regressions includes time dummies. This makes the efficiency of the OLS and Fama-MacBeth coefficients similar.[29]

5.1 Asset pricing application

The equity return regressions from Daniel and Titman (2006, "Market Reactions to Tangible and Intangible Information") are used for the asset pricing example. To demonstrate the effect of equity issues on future equity returns, they regress monthly returns on annual values of the lagged book-to-market ratio, historic

[29] The reported R^2s do not include the explanatory power attributable to the time dummies. This is done to make the R^2 comparable between the OLS and the Fama-MacBeth results. Although the Fama-MacBeth procedure estimates a separate intercept for each year, the constant is calculated as the average of the yearly intercepts. Thus, the Fama-MacBeth R^2 does not include the explanatory power of time dummies. Procedurally, I subtracted the yearly means off of each variable before running the OLS regressions.

Table 6
Asset pricing application: Equity returns and asset tangibility

	I	II	III	IV	V
Log$(B/M)_{t-5}$	0.2456**	0.2456**	0.2456**	0.2456**	0.2064**
	(0.0238)	(0.0247)	(0.0924)	(0.0926)	(0.0795)
Log(Book Return) (last five years)	0.2482**	0.2482**	0.2482**	0.2482**	0.2145**
	(0.0385)	(0.0396)	(0.0859)	(0.0864)	(0.0788)
Market return (last five years)	−0.3445*	−0.3445*	−0.3445*	−0.3445*	−0.3310**
	(0.0257)	(0.0261)	(0.1000)	(0.1001)	(0.0893)
Share issuance	−0.5245**	−0.5245**	−0.5245**	−0.5245**	−0.5143**
	(0.0426)	(0.0427)	(0.1440)	(0.1441)	(0.1235)
R^2	0.0008	0.0008	0.0008	0.0008	0.0008
Coefficient estimates	OLS	OLS	OLS	OLS	FM
Standard errors	White	$CL-F$	$CL-T$	$CL-F\&T$	FM

The table contains coefficient and standard error estimates of the equity return regressions from Daniel and Titman (2006). The data are briefly described in the Appendix and in detail in their paper. The sample runs from July 1968 to December 2003 and contains 844,155 firm-month observations. The estimates in columns I–IV are OLS coefficients and the regressions contain time (month) dummies. Standard errors are reported in parentheses. White standard errors are reported in column I, standard errors clustered by firm in column II, by month in column III, and by firm and month in column IV. Column V contains coefficients and standard errors estimated by Fama-MacBeth. Statistical significance at the 1% and 5% levels are denoted by ** and *, respectively.

Table 7
Corporate finance application: Capital structure regressions (1965–2003)

	I	II	III	IV	V
Ln(MV assets)	0.0460**	0.0460*	0.0460**	0.0460*	0.0394**
	(0.0055)	(0.0184)	(0.0074)	(0.0191)	(0.0076)
Ln(1 + Firm age)	−0.0432**	−0.0432	−0.0432**	−0.0432	−0.0479**
	(0.0084)	(0.0297)	(0.0067)	(0.0293)	(0.0077)
Profits/sales	−0.0330**	−0.0330	−0.0330**	−0.0330	−0.0299**
	(0.0107)	(0.0359)	(0.0098)	(0.0357)	(0.0097)
Tangible assets	0.1043**	0.1043**	0.1043**	0.1043**	0.1158**
	(0.0057)	(0.0197)	(0.0083)	(0.0206)	(0.0096)
Market-to-book (assets)	−0.0251**	−0.0251**	−0.0251**	−0.0251**	−0.0272**
	(0.0006)	(0.0020)	(0.0013)	(0.0023)	(0.0016)
Advertising/sales	−0.3245**	−0.3245	−0.3245**	−0.3245	−0.3965*
	(0.0841)	(0.2617)	(0.0814)	(0.2609)	(0.1712)
R&D/sales	−0.3513**	−0.3513*	−0.3513**	−0.3513*	−0.3359**
	(0.0469)	(0.1544)	(0.0504)	(0.1555)	(0.0501)
R&D > 0 (= 1 if yes)	0.0177**	0.0177*	0.0177**	0.0177*	0.0126**
	(0.0024)	(0.0076)	(0.0025)	(0.0077)	(0.0034)
R-squared	0.1360	0.1360	0.1360	0.1360	0.1300
Coefficient estimates	OLS	OLS	OLS	OLS	FM
Standard errors	White	$CL-F$	$CL-T$	$CL-F\&T$	FM

The table contains coefficient and standard error estimates of a capital structure regression. The dependent variable is the market debt ratio (book value of debt divided by the sum of the book value of assets minus the book value of equity plus the market value of equity). The data are annual observations between 1965 and 2003. The sample contains NYSE firms that pay a dividend in the previous year. There are 24,286 firm-years in the sample. The independent variables are lagged one year and are defined in the Appendix. The estimates in columns I–IV are OLS coefficients and the regressions contain time (year) dummies. Standard errors are reported in parentheses. White standard errors are reported in column I, standard errors clustered by firm in column II, by month in column III, and by firm and month in column IV. Column V contains coefficients and standard errors estimated by Fama-MacBeth. Statistical significance at the 1% and 5% levels are denoted by ** and *, respectively.

changes in book and market values, and a measure of the firm's equity issuance. The data are briefly described in the Appendix and in detail in their paper. Each observation of the dependent variable is a monthly equity return. However, the independent variables are annual values (based on the prior year). Thus, for the twelve observations in a year, the dependent variable (equity returns) changes each month, but the independent variable (e.g., past book value) does not, and is therefore highly persistent.

The White standard errors are essentially the same as standard errors clustered by firm (the difference ranges from 0% to 4%). This occurs because the autocorrelation in the residuals is effectively zero (Figure 9, panel A). The autocorrelation in the independent variable is large and persistent, starting at 0.99 the first month and declining to between 0.49 and 0.75 by the twenty-fourth month, depending on the variable examined. However, since the adjustment in the standard error, and the bias in White standard errors, is a function of the monthly autocorrelation in the *X*s (a large number) times the autocorrelation in the residuals (zero), the standard errors clustered by firm are equal to the White standard errors.

The story is very different when the data are clustered by time (month). The standard errors clustered by month are two to four times larger than the White standard errors. For example, the *t*-statistic on the lagged book-to-market ratio is 10.3 if the White standard error is used and 2.7 if the standard error clustered by month is used. This means there is a significant time effect in the data (Figure 9, panel B) even after including time dummies. Any constant time effect (i.e., one that raises the monthly return for every firm in a given month by the same amount) has already been removed from the data and will not affect the standard errors. The remaining correlation in the time dimension must vary across observations (e.g., Corr[$\varepsilon_{it}\,\varepsilon_{jt}$] varies across *i* and *j*).

Understanding a temporary firm effect is straightforward. The firm effect is temporary (dies off over time) if the 1980 residual for firm A is more highly correlated to the 1981 residual for firm A than to the 1990 residual. This is how the data in Section 4 are simulated. Understanding a nonconstant time effect is more difficult. For the time effect to be nonconstant, it must be that a shock in 1980 has a large effect on firm A and B, but has a significantly smaller effect on firm Z. If the time effect influenced each firm in a given month by the same amount, the time dummies would absorb the effect and clustering by time would not change the reported standard errors. The fact that clustering by time does change the standard errors, means there must be a nonconstant time effect.

If the data are known, economic intuition can be used to determine how the data should be organized and the source of the dependence within a cluster can be predicted. For example, since this data contain monthly equity returns, we might consider how a systematic (macro) shock would affect a firm's returns differently. If the economy booms in a given month, firms in the durable goods industry may have more positive returns than firms in the nondurable

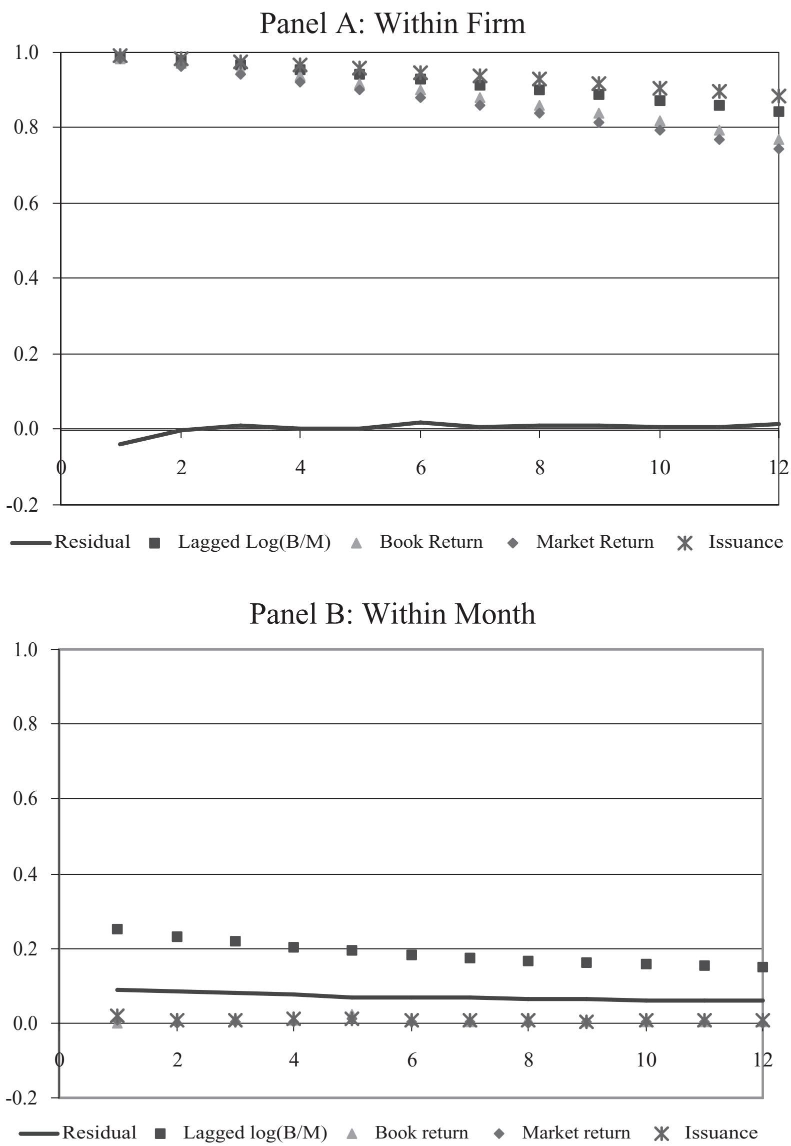

Figure 9
Residuals and independent variables autocorrelation: asset pricing example
The autocorrelations of the residual and the four independent variables are graphed for one to twelve lags. In panel A, the correlations are within firm and are only calculated for observations of the same firm [i.e., Corr (ε_{it} ε_{it-k}) for k equal 1 to 12]. In panel B, the correlations are within month and are only calculated for observations in the same month [i.e., Corr (ε_{it} ε_{i-kt}) for k equal 1 to 12]. The data were sorted by month and then industry (four-digit SIC code) in panel B. The independent variables are described in the Appendix.

goods industry. This can create a situation where the residuals of firms in the same industry are correlated (within the month) with each other but less correlated with firms in another industry. When the data are sorted by month and then industry (four-digit SIC code), there is evidence of this in the autocorrelation for the residuals and the independent variables within each month (Figure 9, panel B). The autocorrelation of the residuals is much larger than when the sort is by firm then month (compare Figure 9, panels A and B) and they die away as firms in more distant industries are considered.[30]

When calculating the standard errors clustered by time, an assumption about how to sort the data does not need to be made to obtain unbiased standard errors. However, if researchers are going to understand what the standard errors are telling them about the structure of the data, they need to consider the source of the dependence in the residuals. By examining how standard errors change when clustered by firm or time (i.e., compare columns I to II and I to III), the nature of the dependence that remains in the residuals can be determined, which can provide researchers guidance on how to improve their models.

Standard errors clustered by both firm and month are also estimated (e.g., Equation (16)). The standard errors clustered by firm and month are essentially identical to the standard errors clustered by month alone (compare column IV and III in Table 6). These two standard errors will be close when there are few firms per month (see, for example, Figure 7) or when there is no firm effect. Given that the data set has many firms per month (at least 1000), the results imply that the data do not contain a significant firm effect. The fact that the serial correlation in the residuals is effectively zero (Figure 9) and that the standard errors clustered by firm are the same as the White standard errors are both consistent with this interpretation.

According to the results reported in Sections 1 and 2, the Fama-MacBeth standard errors perform better in the presence of a time effect than a firm effect, and so given the above results, they should do well here. The Fama-MacBeth coefficients and standard errors are reported in column V of Table 6 (they are a replica of those reported by Daniel and Titman, 2006; Table 3, row 8). The coefficient estimates are similar to the OLS coefficients, and the standard errors are much larger than the White standard errors (2.0–3.4 times), as would be expected in the presence of a time effect. The Fama-MacBeth standard errors are close to the standard errors clustered by time, as both methods are designed to account for dependence in the time dimension.

[30] A nonconstant time effect can be generated by a random coefficient model (Greene, 2002). For example, if the firm's return depends on the firm's β times the market return, but only the market return or time dummies are included in the regression, then the residual will contain the term $\{[\beta_i - \text{Average}(\beta_i)]\ \textit{Market return}_t\}$. Firms that have similar βs will have highly correlated residuals within a month, and firms that have very different βs will have residuals whose correlation is smaller. This logic suggests that the data should be sorted by month and then β. When the data are sorted this way, the autocorrelations are smaller but die away more slowly (declining from 0.033 to 0.031 at a lag of 24) than when the data are sorted by month and then industry (declining from 0.091 to 0.044 at a lag of 24).

5.2 Corporate finance application

A capital structure regression is used for the corporate finance illustration since this is data that contains a significant firm effect, as the data will show, and for which researchers have used a wide variety of techniques to estimate the standard errors. In constructing the data set, independent variables are used that are common in the literature (firm size, firm age, asset tangibility, and firm profitability). This sample contains NYSE firms that pay a dividend in the previous year for the years 1965–2003. The independent variables are lagged one year relative to the dependent variable. The results are reported in Table 7.

The relative importance of the firm effect and the time effect can be seen by comparing the standard errors across the first four columns. The standard errors clustered by firm are dramatically larger than the White standard errors (3.1–3.5 times larger). For example, the *t*-statistic on the profit margin variable is −3.1 based on the White standard error and −0.9 based on the standard error clustered by firm. This is not surprising, since the autocorrelation of the profit margin is extremely high as is the autocorrelation in the residuals (see Figure 10A). Even after ten years it remains above 40%.

The importance of the time effect (after including time dummies) is small in this data set. One can see this in two ways. First, the standard errors clustered by year are only slightly larger than the White standard errors (except for the market-to-book ratio). Second, the standard errors clustered by firm and year are almost identical to the standard errors clustered by just firm (the standard error on market to book is still larger but now by only 16%). Clustering by time has little effect on the standard errors since the correlation of the residuals within a year is small. When the data are sorted by year then industry, the first-order autocorrelation of the residuals is less than 12% (Figure 10B).

In the two examples, clustering standard errors by both firm and time appears unnecessary. In the asset pricing example, these standard errors are identical to the standard errors clustered by time, since there is no firm effect (Table 6). In the corporate finance example, they are identical to the standard errors clustered by firm, since the time effect is small (Table 7). This pattern may not generalize. Thus the standard errors clustered by firm and time can be a useful robustness check (see Cameron, Gelbach, and Miller, 2006 for an example).

The Fama-MacBeth standard errors (Table 7, column V) are close to the standard errors clustered by year and the White standard errors. For example, the Fama-MacBeth *t*-statistic on the profit margin is −3.1, the same as the White *t*-statistic. The results are similar for firm size, firm age, asset tangibility (the ratio of property, plant, and equipment to assets), and R&D expenditure. The White and Fama-MacBeth *t*-statistics are significantly larger than the *t*-statistics clustered by firm. This is the conclusion of Section 1. In the presence of a firm effect, as in this capital structure regression, White and Fama-MacBeth standard errors are significantly biased.

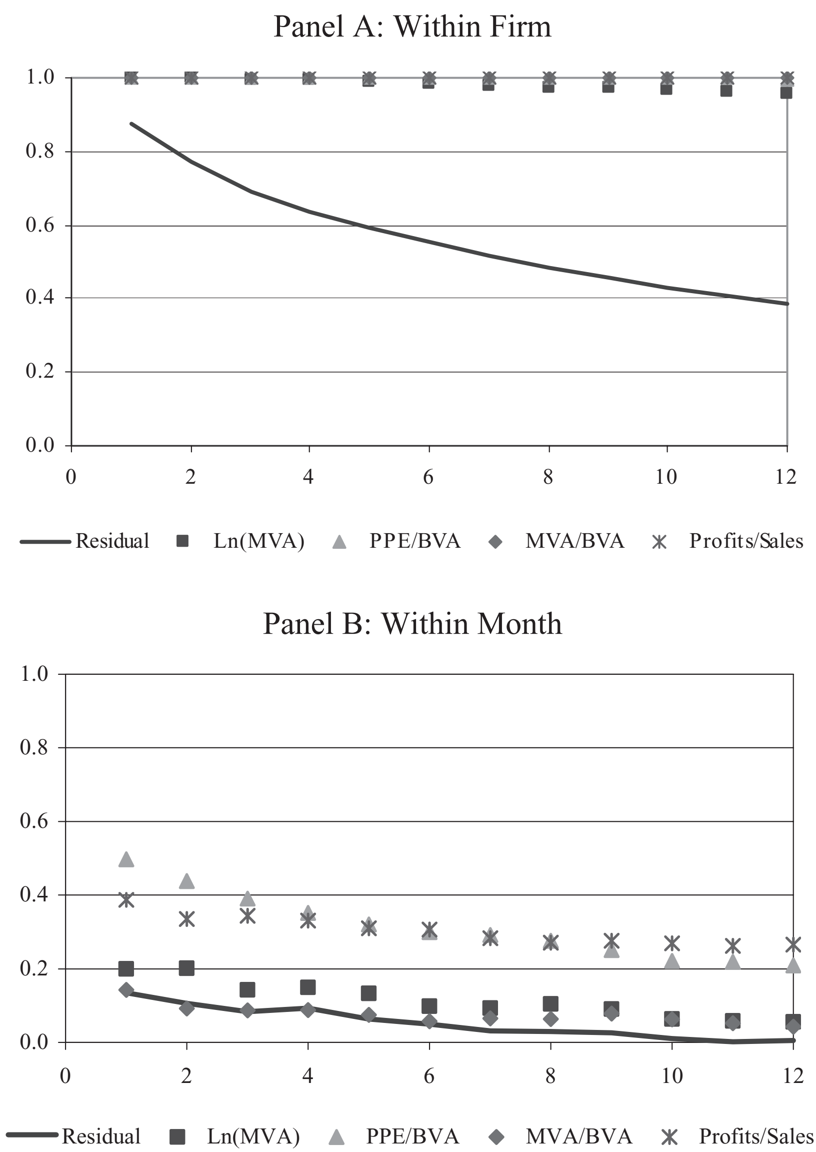

Figure 10
Residuals and independent variables autocorrelation: Corporate finance example
The autocorrelations of the residual and four of the eight independent variables are graphed for one to twelve lags. In panel A, the correlations are within firm and are only calculated for observations of the same firm [i.e., Corr (ε_{it} ε_{it-k}) for k equal 1 to 12]. In panel B, the correlations are within year and are only calculated for observations of the same year [i.e., Corr ($\varepsilon_{it,}$ ε_{i-kt}) for k equal 1 to 12]. The data were sorted by month and then industry (four-digit SIC code) in panel B. The independent variables are described in the Appendix. The graphs for the remaining four variables are similar and are available from the author.

6. Conclusions

It is well known that OLS and White standard errors are biased when the residuals are not independent. What has been less clear is how researchers should estimate standard errors when using panel data sets. The empirical finance literature has proposed and used a variety of methods for estimating standard errors when the residuals are correlated across firms or years. This paper shows how the performance of the different methods varies considerably and their relative accuracy depends on the structure of the data. Simply put, estimates that are robust to the form of dependence in the data produce unbiased standard errors and correct confidence intervals; estimates that are not robust to the form of dependence in the data produce biased standard errors and confidence intervals that are often too small. The two illustrations in Section 5 demonstrate that the magnitude of the biases can be very large.

Although it may seem obvious that choosing the correct method is important, the absence of good advice in the literature means that the correct decision has not always been made, as the literature survey demonstrates. The purpose of this paper is to provide such guidance. In the presence of a firm effect [e.g., $\mathrm{Cov}(X_{it}\ \varepsilon_{it}, X_{it-k}\ \varepsilon_{it-k}) \neq 0$], standard errors are biased when estimated by OLS, White, Newey-West (modified for panel data sets), Fama-MacBeth, or Fama-MacBeth corrected for first-order autocorrelation.[31] Despite this, these methods are often used in the literature when the regressions being estimated contain a firm effect. The standard errors clustered by firm are unbiased and produce correctly sized confidence intervals whether the firm effect is permanent or temporary. The fixed effect and random effects model also produces unbiased standard errors but only when the firm effect is permanent.

In the presence of a time effect [e.g., $\mathrm{Cov}(X_{it}\ \varepsilon_{it}, X_{kt}\ \varepsilon_{kt}) \neq 0$], Fama-MacBeth produces unbiased standard errors and correctly sized confidence intervals. This is not surprising since it was designed for just such a setting. Standard errors clustered by time also produce unbiased standard errors and correctly sized confidence intervals, but only when there are a sufficient number of clusters. When there are too few clusters, clustered standard errors are biased even when clustered on the correct dimension (Figures 5 and 7). When both a firm and a time effect are present in the data, researchers can address one parametrically (e.g., by including time dummies) and then estimate standard errors clustered on the other dimension. Alternatively, researchers can cluster on multiple dimensions. When there are a sufficient number of clusters in each dimension, standard errors clustered on multiple dimensions are unbiased

[31] Skoulakis (2006) proposes applying the logic of Fama-MacBeth to each firm, instead of each year. He demonstrates that running *N* time series regressions and using the standard deviation of the *N* coefficients produces an estimate that is correct in the presence of a firm effect. Pesaran and Smith (1995) make a similar suggestion but their focus is on coefficient estimation and they do not cite Fama-MacBeth. I found only two papers in the literature that have used the Fama-MacBeth approach in the way suggested by Skoulakis (Coval and Shumway, 2005; and Bloomfield, O'Hara, and Saar, 2006). A generalized form of the Fama-MacBeth approach can be found in Ibragimov and Muller (2007).

and produce correctly sized confidence intervals whether the firm effect is permanent or temporary.

Knowing that the OLS standard errors are biased means that there is information in the residual that the researcher is not using (i.e., the residuals are correlated). This suggests that researchers can improve the efficiency of their estimates (using a technique such as fixed effects, GLS or GMM) and may also use these techniques to test whether their model is correctly specified. Thus in addition to providing a guide to the correct estimation of standard errors, the techniques in this paper can be used to help researchers diagnose potential problems with their models. By comparing the different standard errors, one can quickly observe the presence and magnitude of a firm and/or a time effect. As seen in Section 5, when the standard errors clustered by firm are much larger than the White standard errors (three to four times larger), this indicates the presence of a firm effect in the data (Table 7). When the standard errors clustered by time are much larger than the White standard errors (two to four times larger), this indicates the presence of a time effect in the data (Table 6). When the standard errors clustered by firm and time are much larger than the standard errors clustered by only one dimension, this can indicate the presence of both a firm and a time effect in the data. Which dependencies are most important will vary across data sets and thus researchers must consult their data. This information can provide researchers with guidance in selecting the standard error estimation method that is most appropriate, intuition as to the deficiency of their models, and guidance for improving their models.

Appendix: Data Set Constructions

A. Asset pricing application

The data for the regressions in Table 6 are taken from Daniel and Titman's paper "Market Reactions to Tangible and Intangible Information" (2006). A more detailed description of the data can be found in their paper. The dependent variable is monthly returns on individual stocks from July 1968 to December 2001. The independent variables are:

- Log(Lagged book-to-market). The log of the total book value of the equity at the end of the firm's fiscal year ending anywhere in year $t-6$ divided by the total market equity on the last trading day of calendar year $t-6$.
- Log(Book return). This variable measures changes in the book value of the firm's equity over the previous five years. It is calculated as the log of one plus the percentage change in the book value over the past five years. Thus, if you purchased 1% of the book value five years ago, and neither invested additional cash nor took any cash out of the investment, the book return is the current percentage ownership divided by the initial 1%.
- Log(Market return). This variable measures changes in the market value of the firm over the previous five years. It is calculated as the log of one plus the market return from the last day of year $t-6$ to the last day of year $t-1$.
- The share issuance is a measure of the net equity issuance. It is calculated as minus the log of the percentage ownership at the end of five years, assuming the investor started with 1% of the firm. Thus, if an investor purchases 1% of the firm and five years later they own 0.5% of the firm, then the share issuance is equal to $-\log(0.5/1.0) = 0.693$. Investors are assumed to

neither take money out of their investment nor add additional money to their investment. Thus, any cash flow that investors receive (e.g., dividends) would be reinvested. For transactions such as equity issues and repurchases, the investor is assumed not to participate and thus these will lower or raise the investor's fractional ownership.

To make sure that the accounting information is available to implement a trading strategy, the independent variables are lagged at least six months. Thus, the independent variables for a fiscal year ending anytime during calendar year $t - 1$ are used to predict future monthly returns from July of year t through June of year $t + 1$. The independent variables are annual measures and are thus constant for each of the twelve monthly observations during the following year (July through June).

B. Corporate finance application

The regressions in Table 7 are constructed using Compustat data from 1965 to 2003. The dependent variable, the market debt ratio, is defined as the book value of debt (data9 + data34) divided by the sum of the book value of assets (data6) minus the book value of equity (data60) plus the market value of equity (data25 * data199). The independent variables are lagged one year, and I only include observations where the firm paid a dividend (data21 > 0) in the previous year. To reduce the influence of outliers, I capped ratio variables (e.g., profits to sales, tangible assets, advertising to sales, and R&D to sales) at the first and ninety-ninth percentile. The independent variables are:

- Ln(Market Value of Assets). The log of the sum of the book value of assets (data6) minus the book value of equity (data60) plus the market value of equity (data25 * data199).
- Ln(1 + Firm Age). Firm age is calculated as the number of years the firm's stock has been listed. Firm age is calculated as the current year (fyenddt) minus the year the stock began trading (linkdt).
- Profits/Sales. Operating profits before depreciation (data13) divided by sales revenue (data12).
- Tangible Assets. Property, plant, and equipment (data8) divided by the book value of total assets (data6).
- Advertising/Sales. Advertising expense (data45) divided by sales (data12).
- R&D/Sales. R&D expenditure (data46) divided by sales (data12). If R&D is missing, it is coded as zero.
- R&D > 0. A dummy variable equal to one if R&D expenditure is positive, and zero otherwise.

References

Almazan, A., K. Brown, M. Carlson, and D. A. Chapman. 2004. Why Constrain Your Mutual Fund Manager? *Journal of Financial Economics* 73:289–321.

Anderson, R., and D. Reeb. 2004. Founding-Family Ownership and Firm Performance: Evidence from the S&P 500. *Journal of Finance* 58:1301–28.

Andrews, D. W. K. 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* 59:817–58.

Arellano, M. 1987. Computing Robust Standard Errors for Within-Groups Estimators. *Oxford Bulletin of Economics and Statistic* 49:431–3.

Baker, M., J. Stein, and J. Wurgler. 2003. When Does the Market Matter? Stock Prices and the Investment of Equity-Dependent Firms. *Quarterly Journal of Economics* 118:969–1005.

Baker, M., and J. Wurgler. 2002. Market Timing and Capital Structure. *Journal of Finance* 57:1–32.

Bakshi, G., N. Kapadia, and D. Madan. 2003. Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies* 16:101–43.

Bertrand, M., E. Duflo, and Sendhil Mullainathan. 2004. How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics* 119:249–75.

Bloomfield, R., M. O'Hara, and G. Saar. 2006. The Limits of Noise Trading: An Experimental Analysis. Working Paper, Cornel University.

Brockman, P., and D. Chung. 2001. Managerial Timing and Corporate Liquidity: Evidence from Actual Share Repurchases. *Journal of Financial Economics* 61:417–48.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2006. Robust Inference with Multi-Way Clustering. Working paper, University of California-Davis.

Chakravarty, S., H. Gulen, and S. Mayhew. 2004. Informed Trading in Stock and Option Markets. *Journal of Finance* 59:1235–57.

Chen, J., H. Hong, and J. Stein. 2001. Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices. *Journal of Financial Economics* 61:345–81.

Cheng, S., V. Nagar, and M. Rajan. 2005. Identifying Control Motives in Managerial Ownership: Evidence from Antitakeover Legislation. *Review of Financial Studies* 18:637–72.

Choe, H., B. Kho, and R. M. Stulz. 2005. Do Domestic Investors Have an Edge? The Trading Experience of Foreign Investors in Korea. *Review of Financial Studies* 18:795–829.

Christopherson, J., W. Ferson, and D. Glassman. 1998. Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance. *Review of Financial Studies* 11:111–42.

Cochrane, J. 2001. *Asset Pricing*. Princeton, NJ: Princeton University Press.

Coval, J., and Tyler Shumway. 2005. Do Behavioral Biases Affect Prices. *Journal of Finance* 60:1–34.

Daniel, K., and S. Titman. 2006. Market Reactions to Tangible and Intangible Information. *Journal of Finance* 61:1605–43.

Denis, D., D. Denis, and K. Yost. 2002. Global Diversification, Industrial Diversification, and Firm Value. *Journal of Finance* 57:1951–79.

Doidge, C. 2004. U.S. Cross-Listings and the Private Benefits of Control: Evidence from Dual-Class Firms. *Journal of Financial Economics* 72:519–53.

Donald, S., and K. Lang. 2007. Inference with Difference in Differences and Other Panel Data. *Review of Economics and Statistics* 89:221–33.

Efron, B., and R. J. Tibshirani. 1986. Bootstrap Measures For Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science* 1:54–77.

Fama, E., and J. MacBeth. 1973. Risk, Return and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81:607–36.

Fama, E., and K. French. 2002. Testing Tradeoff and Pecking Order Predictions about Dividends and Debt. *Review of Financial Studies* 15:1–33.

Fama, E., and K. French. 2001. Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay? *Journal of Financial Economics* 60:3–43.

Faulkender, M., and M. Petersen. 2006. Does the Source of Capital Affect Capital Structure? *Review of Financial Studies* 19:45–79.

Gentry, W., D. Kemsley, and Christopher Mayer. 2003. Dividend Taxes and Share Prices: Evidence from Real Estate Investment Trusts. *Journal of Finance* 58:261–82.

Graham, J., M. Lemmon, and J. Schallheim. 1998. Debt, Leases, Taxes and the Endogeneity of Corporate Tax Status. *Journal of Finance* 53:131–62.

Greene, W. 2002. *Econometric Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Gross, D., and N. Souleles. 2004. An Empirical Analysis of Personal Bankruptcy and Delinquency. *Review of Financial Studies* 15:319–47.

Hansen, C. 2007. Asymptotic Properties of a Robust Variance Matrix Estimator for Panel Data When *T* Is Large. *Journal of Econometrics* 141:597–620.

Hausman, J. 1978. Specification Tests in Econometrics. *Econometrica* 46:1251–71.

Ibragimov, R., and U. Muller. 2007. *T*-statistic Based Correlation and Heterogeneity Robust Inference. Working paper, Harvard University.

Jagannathan, R., and Z. Wang. 1998. An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression. *Journal of Finance* 53:1285–309.

Johnson, S. 2003. Debt Maturity and the Effects of Growth Opportunities and Liquidity Risk on Leverage. *Review of Financial Studies* 16:209–36.

Kayhan, A., and S. Titman. 2007. Firms' Histories and Their Capital Structures. *Journal of Financial Economics*. 83:1–32.

Kemsley, D., and D. Nissim. 2002. Valuation of the Debt Tax Shield. *Journal of Finance*. 57:2045–73.

Kezdi, G. 2004. Robust Standard Error Estimation in Fixed-Effects Panel Models. *Hungarian Statistical Review* 9:95–116.

Lakonishok, J., and I. Lee. 2001. Are Insider Trades Informative? *Review of Financial Studies* 14:79–111.

Lamont, O., and C. Polk. 2001. Does Diversification Destroy Value? Evidence from the Industry Shocks. *Journal of Financial Economics* 63:51–77.

Liang, K., and Scott Zeger. 1986 Longitudinal Analysis Using Generalized Linear Models. *Biometrika*. 73:13–22.

MacKay, P. 2003. Real Flexibility and Financial Structure: An Empirical Analysis. *Review of Financial Studies*. 16:1131–65.

Maksimovic, V., and Gordon Phillips. 2002. Do Conglomerate Firms Allocate Resources Inefficiently Across Industries? Theory and Evidence. *Journal of Finance* 57:721–67.

Moulton, B. 1986. Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics* 32:385–97.

Moulton, B. 1990. An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *Review of Economics and Statistics* 72:334–8.

Nagel, S. 2005. Short Sales, Institutional Investors and the Cross-Section of Stock Returns. *Journal of Financial Economics*. 78:277–309.

Newey, W., and K. West. 1987. A Simple, Positive Semi-Definite, Heteroscedastic and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55:703–8.

Pastor, L., and P. Veronesi. 2003. Stock Valuation and Learning about Profitability. *Journal of Finance* 58:1749–89.

Pesaran, M. H., and R. Smith. 1995. Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics* 68:79–113.

Rogers, W. 1993. Regression Standard Errors in Clustered Samples. *Stata Technical Bulletin* 13:19–23.

Sapienza, P. 2004. The Effects of Government Ownership on Bank Lending. *Journal of Financial Economics* 72:357–84.

Schultz, P., and T. Loughran. 2005. Liquidity: Urban versus Rural Firms. *Journal of Financial Economics* 78:341–74.

Skoulakis, G. 2006. Panel Data Inference in Finance: Least-Squares vs. Fama-MacBeth. Working paper, University of Maryland.

Thompson, S. 2006. Simple Formulas for Standard Errors That Cluster by Both Firm and Time. Working paper, Harvard University.

White, H. 1984. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity. *Econometrica* 48:817–38.

Williams, R. 2000. A Note on Robust Variance Estimation for Cluster-Correlated Data. *Biometrics* 56:645–6.

Wooldridge, J. 2007. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wooldridge, J. 2003. Cluster-Sample Methods in Applied Econometrics. *American Economic Review* 93:133–8.

Wu, Y. 2004. The Choice of Equity-Selling Mechanisms. *Journal of Financial Economics* 74:93–119.