Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches

Mitchell A. Petersen
Kellogg School of Management, Northwestern University
and NBER

Abstract

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, the two literatures have used different solutions to this problem. Corporate finance has relied on clustered standard errors, while asset pricing has used the Fama-MacBeth procedure. 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.

I)   Introduction

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, forty-two percent 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 within cluster correlation varied among the remaining papers. Thirty-four percent of the remaining papers estimated both the coefficients and the standard errors using the Fama-MacBeth procedure (Fama-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 correlation within a cluster (e.g. within a firm or industry). Seven

---

[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 which 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 or industry) as a cluster. I included both linear regressions as well as non-linear techniques such as logits and tobits in my survey. I included only papers which report at least five observations in each dimension (e.g. firms and years). 207 papers met the selection criteria. Papers which did not report the method for estimating the standard errors, or reported correcting the standard errors only for heteroscedasticity (i.e. White standard errors which 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 regresses on earnings surprises). In these cases, the bias in the standard errors can be quite large. See Section VI for two illustrations.

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 percent of the papers reported clustered standard errors (Williams, 2000, Rogers, 1993, Andrews, 1991, Moulton, 1990, Arellano, 1987, Moulton, 1986, Liang, and Zeger, 1986) which are White standard errors adjusted to account for possible correlation within a cluster. These are also called Rogers standard errors in the finance literature.

Although the literature has used a diversity 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 which 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 (time series dependence) for a given firm. I will call this an unobserved firm effect (see Wooldridge, 2002). 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, I can estimate the coefficients and standard errors using each of the methods and compare their relative performance.

---

[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 which 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 which I have highlighted.

Section II examines the sensitivity of standard error estimates to the presence of a firm fixed effect, a feature common among many variables including financial leverage, dividends, and investment. My 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 III, 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 IV, were 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.

I initially specified the firm effect as a constant (e.g. it does not decay over time). In practice, the firm effect may decay and so the correlation between residuals declines as the time between them grows. In Section V, I simulate data with a more general correlation structure. This allows me to compare OLS, clustered, and Fama-MacBeth standard errors in a more general setting. Simulating the temporary firm effect also allows me to examine 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,

3

as suggested by some authors (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 VI, 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 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.

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

A)   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 we have 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 strong assumption that the model is correctly specified. The zero mean is without loss of generality and allows us to calculate variances as sums of the squares of the variable. The estimated coefficient is:

---

[3] The simulations in the paper are based on linear regressions. However, my results generalize to non-linear 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 non-linear models. The magnitude of the biases are similar to what I report for linear models (results available from the author).

$$\hat{\beta}_{OLS} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it} Y_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\left(X_{it}\beta + \varepsilon_{it}\right)}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}$$

$$= \beta + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} \tag{2}$$

and the asymptotic variance of the estimated coefficient is:

$$AVar\left[\hat{\beta}_{OLS} - \beta\right] = \plim_{\substack{N\to\infty \\ T\ fixed}}\left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\varepsilon_{it}\right)^2\left(\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N}\right)^{-2}\right]$$

$$= \plim_{\substack{N\to\infty \\ T\ fixed}}\left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2\right)\left(\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N}\right)^{-2}\right] \tag{3}$$

$$= \frac{1}{N}\left(T\sigma_X^2\sigma_\varepsilon^2\right)\left(T\sigma_X^2\right)^{-2}$$

$$= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT}$$

This is the standard OLS formula and is based on the assumption that the errors are independent and identically distributed (Greene, 1990, page 148). The independence assumption is used to move from the first to the second line 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 line.[4] The independence assumption is often violated in panel data and is the focus of the paper.

---

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

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

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

Assume that the independent variable X also has a firm specific component.

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

Each of the components of X (μ and υ) 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.

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

Given this data structure [equations (1), (4), and (5)], I can calculate the true standard error

---

[5] This language follows Wooldridge, 2002 (see Chapter 10). When I use the word fixed it means the unobserved firm effect does not die away over time. Wooldridge calls this a time-constant unobserved effect. This means the correlation between ϵ_{it} and ϵ_{it-k} is a constant with respect to k. In Section V, I will examine cases where this correlation does die 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.

of the OLS coefficient. Since the residuals are no longer independent within 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:

$$
\begin{aligned}
AVar\left[\hat{\beta}_{OLS} - \beta\right] &= \plim_{\substack{N \to \infty \\ T\ fixed}} \left[\frac{1}{N^2} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}\, \varepsilon_{it} \right)^2 \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\
&= \plim_{\substack{N \to \infty \\ T\ fixed}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}\, \varepsilon_{it} \right)^2 \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\
&= \plim_{\substack{N \to \infty \\ T\ fixed}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}^2\, \varepsilon_{it}^2 + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} X_{it}\, X_{is}\, \varepsilon_{it}\, \varepsilon_{is} \right) \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\
&= \frac{1}{N} \left(T\, \sigma_X^2\, \sigma_\varepsilon^2 + T(T-1)\, \rho_X\, \sigma_X^2\, \rho_e\, \sigma_e^2 \right) \left(T\, \sigma_X^2 \right)^{-2} \\
&= \frac{\sigma_\varepsilon^2}{\sigma_X^2\, NT} \left(1 + (T-1)\, \rho_X\, \rho_\varepsilon \right)
\end{aligned}
\qquad (7)
$$

I use the assumption that residuals are independent across firms in deriving the second line. 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 non-zero.[7] The magnitude of the error is also increasing in the number of years. To understand

---

[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 ϵ_{t-k} times the correlation between X_t and X_{t-k}, for all k<T (Wooldridge, 2002). It is equal to:

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

Since, the autocorrelations can be positive or negative, it is possible for the OLS standard error to under or over-estimate 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 is even larger than specified by equation (7) (see Moulton, 1986).

this intuition, consider the extreme case where the independent variables and residuals are perfectly correlated across time (i.e. $\rho_X = 1$ and $\rho_\epsilon = 1$). 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 each additional year provides N additional observations and the estimated standard error will shrink accordingly and incorrectly.

The correlation of the residuals within cluster is the problem the clustered standard errors are designed to correct.[8] By squaring the sum of $X_{it}\epsilon_{it}$ within each cluster, the covariance between residuals within cluster is estimated (see Figure 1). This correlation can be of any form; no parametric structure is assumed. However, the squared sum of $X_{it}\epsilon_{it}$ 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, 2001; and Wooldridge, 2002). I will return to this issue in Section III.

B)   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 we can observe the true standard error as well as the average estimated standard errors.[9] In the first version of the simulation, I include an unobserved firm effect that is

---

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

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

[9] Each simulated data set contains 5,000 observations (500 firms and 10 years per firm). The components of the independent variable (μ v) and the residual (γ η) are independent of each other 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 5,000 simulations.

fixed, but no time effect in the independent variable and the residual. Thus the data are simulated as described in equations (4) and (5). Across simulations I assumed that the standard deviation of the independent variable and the residual are both constant at one and two respectively. This will produce an R^2 of 20 percent. Across different simulations, I altered the fraction of the variance in the independent variable which is due to the firm effect. This fraction ranges from zero to seventy-five percent in twenty-five percent increments (see Table 1). I did the same for the residual. This allows me to demonstrate 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 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 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 (see 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 (see 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 fifty percent of the variability in both the residual and the independent variable is due to the firm effect ($\rho_X = \rho_\epsilon = 0.50$), 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 (i.e. in the independent variable) or across the rows (i.e. in the residual). The true standard error does rise.

When I estimate the standard error of the coefficient using clustered standard errors, the estimates (the fifth entry in each cell) are very close to the true standard error. 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, 2005). The fraction of OLS t-statistics which are statistically significant at the one percent 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 (see Figure 2-A and Table 1). As you move down the diagonal in Table 1, the percent of t-statistics which are statistically significant at the 1 percent level rises. For example 15.3% percent of the OLS t-statistics are statistically significant at the 1 percent level when $\rho_X = \rho_\epsilon = 0.50$. The clustered standard errors are unbiased (see Table 1) and the empirical distribution of the t-

---

[10] All of the regressions contain a constant whose true value is zero. The intuition carries over to the intercept estimation. The estimated intercept averages -0.0003 with a standard deviation of 0.0669, when $\rho_X = \rho_\epsilon = 0.50$. 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 which 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 $\rho_X = \rho_\epsilon = 0.50$.

statistics are also correct (see Figure 2-B). 0.9 percent of the clustered t-statistics are significant at the one percent level. The reason the t-statistics give us the same intuition as the standard errors is the standard errors are estimated very precisely. For example, the mean OLS standard error 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 $\rho_X = \rho_\epsilon = 0.50$).[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 two. The bias rises from about 30 percent when there are five years of data per firm to 73 percent when there are 50 years (when $\rho_X = \rho_\epsilon = 0.50$, see Figure 3). The robust standard errors are consistently close to the true standard errors independent of the number of time periods (see Figure 3).

C)   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:

---

[11] I do not report the mean squared error (MSE) of the standard error estimates, 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 is 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 V-D. Tables of MSEs are available from the author.

$$\hat{\beta}_{FM} = \sum_{t=1}^{T} \frac{\hat{\beta}_t}{T}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \left(\frac{\sum_{i=1}^{N} X_{it} Y_{it}}{\sum_{i=1}^{N} X_{it}^2} \right) = \beta + \frac{1}{T} \sum_{t=1}^{T} \left(\frac{\sum_{i=1}^{N} X_{it} \varepsilon_{it}}{\sum_{i=1}^{N} X_{it}^2} \right) \quad \textbf{(8)}$$

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

$$S^2\left(\hat{\beta}_{FM}\right) = \frac{1}{T} \sum_{t=1}^{T} \frac{\left(\hat{\beta}_t - \hat{\beta}_{FM}\right)^2}{T-1} \quad \textbf{(9)}$$

The variance formula assumes that the yearly estimates of the coefficient (β_t) are independent of

each other. This is only correct if $X_{it} \epsilon_{it}$ is independent of $X_{is} \epsilon_{is}$ for $t \neq s$. As discussed above, this

is not true when there is a firm effect in the data (i.e. $\rho_X \rho_\epsilon \neq 0$). 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:

$$AVar\left(\hat{\beta}_{FM}\right) = \frac{1}{T^2} AVar\left(\sum_{t=1}^{T} \hat{\beta}_t\right)$$

$$= \frac{AVar\left(\hat{\beta}_t\right)}{T} + \frac{2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} ACov\left(\hat{\beta}_t, \hat{\beta}_s\right)}{T^2} \quad \textbf{(10)}$$

$$= \frac{AVar\left(\hat{\beta}_t\right)}{T} + \frac{T(T-1)}{T^2} ACov\left(\hat{\beta}_t, \hat{\beta}_s\right)$$

Given our specification of the data (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 calculated as follows for $t \neq s$:

$$ACov\left(\hat{\beta}_t,\hat{\beta}_s\right)=\plim_{N\to\infty}\left[\left(\frac{\sum_{i=1}^{N}X_{it}^2}{N}\right)^{-1}\left(\frac{\sum_{i=1}^{N}X_{it}\,\varepsilon_{it}}{N}\right)\left(\frac{\sum_{i=1}^{N}X_{is}\,\varepsilon_{is}}{N}\right)\left(\frac{\sum_{i=1}^{N}X_{is}^2}{N}\right)^{-1}\right]$$

$$=\left(\sigma_X^2\right)^{-2}\plim_{N\to\infty}\left[\left(\frac{\sum_{i=1}^{N}X_{it}\,\varepsilon_{it}}{N}\right)\left(\frac{\sum_{i=1}^{N}X_{is}\,\varepsilon_{is}}{N}\right)\right]$$

$$=\left(\sigma_X^2\right)^{-2}\plim_{N\to\infty}\left[\frac{\sum_{i=1}^{N}X_{it}\,X_{is}\,\varepsilon_{it}\,\varepsilon_{is}}{N^2}\right]$$

$$=\left(\sigma_X^2\right)^{-2}\frac{N\,\rho_X\,\sigma_X^2\,\rho_\varepsilon\,\sigma_\varepsilon^2}{N^2}$$

$$=\frac{\rho_X\,\rho_\varepsilon\,\sigma_\varepsilon}{N\,\sigma_X^2}$$

(11)

Combining equations (10) and (11) gives us the expression for the asymptotic variance of the Fama-MacBeth coefficient estimates.

$$AVar\left(\hat{\beta}_{FM}\right)=\frac{AVar\left(\hat{\beta}_t\right)}{T}+\frac{T(T-1)}{T^2}ACov\left(\hat{\beta}_t,\hat{\beta}_s\right)$$

$$=\frac{1}{T}\left(\frac{\sigma_\varepsilon^2}{N\,\sigma_X^2}\right)+\frac{T(T-1)}{T^2}\left(\frac{\rho_X\,\rho_\varepsilon\,\sigma_\varepsilon^2}{N\,\sigma_X^2}\right)$$

$$=\frac{\sigma_\varepsilon^2}{NT\,\sigma_X^2}\left(1+(T-1)\rho_X\,\rho_\varepsilon\right)$$

(12)

This result is the same as our expression for the variance of the OLS coefficient (see equation 7). The Fama-MacBeth standard error 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).

D)   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 5,000 simulated data sets which 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 Table 1 and 2). These results demonstrate that both OLS and Fama-MacBeth standard errors are biased downward (see 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 percent, the OLS standard error has a bias of 60% ($0.595 = 1 - 0.0283/0.0698$, see Table I) and the Fama-MacBeth standard error has a bias of 74 percent ($0.738 = 1 - 0.0183/0.0699$, see Table II). 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 ($\rho_X \rho_\epsilon$ increases), the OLS bias grows, and the Fama-MacBeth bias grows even faster.[12] The incremental bias of the Fama-MacBeth standard errors is due to the way in which the estimated variance is calculated. To see this we need to expand the expression of the estimated variance (equation 9).

---

[12] The distribution of empirical t-statistics is even wider for the Fama-MacBeth than for OLS (compare Figures 2-A and 2-C). 25 percent of the Fama-MacBeth t-statistics are statistically significant at the 1 percent level compared to 15 percent of the OLS t-statistics when $\rho_X = \rho_\epsilon = 0.50$ (see Table 1 and 2).

$$S^2\left(\hat{\beta}_{FM}\right) = \frac{1}{T(T-1)}\sum_{t=1}^{T}\left[\frac{\sum_{i=1}^{N}X_{it}\,\varepsilon_{it}}{\sum_{i=1}^{N}X_{it}^2} - \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N}X_{it}\,\varepsilon_{it}}{\sum_{i=1}^{N}X_{it}^2}\right)\right]^2$$

$$= \frac{1}{T(T-1)}\sum_{t=1}^{T}\left[\frac{\sum_{i=1}^{N}\left(\mu_i+\upsilon_{it}\right)\left(\gamma_i+\eta_{it}\right)}{\sum_{i=1}^{N}\left(\mu_i+\upsilon_{it}\right)^2} - \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N}\left(\mu_i+\upsilon_{it}\right)\left(\gamma_i+\eta_{it}\right)}{\sum_{i=1}^{N}\left(\mu_i+\upsilon_{it}\right)^2}\right)\right]^2 \quad \textbf{(13)}$$

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 $\rho_X\rho_\epsilon$) 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 $\rho_X\rho_\epsilon$ is equal to one; the true standard error is $(\sigma_\epsilon/N\sigma_X)^{\frac{1}{2}}$ 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 (see Figure 3).

E)   Incorrect Standard Error Estimates in Published Papers.

Although I have just demonstrated 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 contains a firm effect. As part of my literature survey, I looked for papers which 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 will be biased. Since I am not able to replicate each of the studies, I will discuss a few

examples where Fama-MacBeth standard errors have been used with such data.

The first example is a logit estimate of whether a firm pays a dividend (a highly persistent

variable) on firm characteristics such as the firm's market to book ratio, the earnings to assets ratio,

and relative firm size (Fama and French, 2001). A second example are the papers which 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 (Pastor and

Veronesi, 2003, and Kemsley and Nissim, 2002).[13] A third example are papers which run capital

structure regressions. In these papers, the 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 (see for example Baker and Wurgler, 2002, Fama

and French, 2002, and Johnson, 2003).[14] As I will show in Section VI, the serial correlation among

these variables is quite large (usually greater than 0.95 after ten years). Since both the left and right

---

[13] Both of these papers correct the Fama-MacBeth standard errors for the first order autocorrelation of the estimated slopes. Pastor and Veronesi (2003) report that this does not change their answer. I will show in Section V-D that this correction still produces biased standard errors and this may explain Pastor and Veronesi's finding that the adjustment has little effect on their estimated standard errors.

[14] 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 II. 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 (see 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 V-D and show it can generate biased standard errors as well.

hand side variables in these three regressions are highly persistent, this is the kind of data where Fama-MacBeth standard errors are biased. In Section VI-B, I will estimate a capital structure regression and show that the magnitude of the bias is indeed large. Despite the potential for biased standard errors and thus incorrect inferences, Fama-MacBeth standard errors are still used in the literature.

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) standard errors in the kinds of regressions I have listed, they believe (incorrectly) that this approach is correct. The problem is actually worse. The published finance literature has not only used incorrect methods but has gone on to provide (incorrect) advice which 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 $\rho_\epsilon \neq 0$). Wu (2004) 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) 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, et. al. (2005) explain that "The Fama-McBeth 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

17

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 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 isn't with the Fama-MacBeth method, only with its use. It was developed to account for correlation between observations on different firms in the same year, not to account for correlation between observations on the same firm in different years. It is now being used and recommended in cases where it produces 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 III.

F)   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 was initially designed to account for serial correlation of unknown form in the residuals of a single time series. It has been modified for use in a panel data set by estimating only correlations between lagged residuals in the same cluster (see Bertrand, Duflo, and Mullainathan, 2004, Doidge, 2004, MacKay, 2003, Brockman and Chung, 2001). 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.[15] To examine the relative performance of the Newey-West, I simulated 5,000 data sets where the firm effect is fixed and assumed to account for twenty-five percent of the variability of both the independent variable

---

[15] In the standard application of Newey-West, a lag length of M implies that the correlation between ϵ_t and ϵ_{t-k} are included for k running from -M to M. When Newey-West has been applied to panel data sets, correlations between lagged and leaded values are only included when they are drawn from the same cluster. Thus a cluster which 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 tth observation where t runs from 1 to T.

and the residual.

The standard error estimated by the Newey-West is an increasing function of the lag length in the simulation. When the lag length is set to zero, the estimated standard error is numerically identical to the White standard error, which is only robust to heteroscedasticity (White, 1984). This is the same as the OLS standard error in my simulation. Not surprisingly, this estimate significantly underestimates the true standard error (see Figure 4). As the lag length is increased 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 (see Figure 4). 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. However, even with the maximum lag length of 9, the Newey-West estimates still eliminate only 60 percent of the bias that OLS/White standard errors produce [(0.0328-0.0283)/(0.0358-0.0283)].

Unlike clustered standard errors, the Newey-West approach to estimating standard errors in a panel data set produces biased estimates. The source of the bias is the weighting function. When estimating standard errors, Newey-West multiplies the covariance of lag j (e.g. $\epsilon_t \epsilon_{t-j}$) by the weight [1-j/(M+1)], where M is the specified maximum lag. If I set the maximum lag equal to T-1, then the central matrix in the variance equation of the Newey-West standard error is:

$$\sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}\,\varepsilon_{it}\right)^{2} = \sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}^{2}\,\varepsilon_{it}^{2} + 2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} w(t-s)\,X_{it}\,X_{is}\,\varepsilon_{it}\,\varepsilon_{is}\right)$$

$$= \sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}^{2}\,\varepsilon_{it}^{2} + 2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t} w(j)\,X_{it}\,X_{it-j}\,\varepsilon_{it}\,\varepsilon_{it-j}\right) \qquad \textbf{(14)}$$

$$= \sum_{i=1}^{N}\left(\sum_{t=1}^{T} X_{it}^{2}\,\varepsilon_{it}^{2} + 2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t}\left(1-\frac{j}{T}\right)X_{it}\,X_{it-j}\,\varepsilon_{it}\,\varepsilon_{it-j}\right)$$

This is identical to the clustered standard error formula except for the weighting function [w(j)]. The clustered standard errors use a weighting function of one for all covariances. 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 semi-definite. For fixed j, the weight w(j) approaches one asymptotically. Newey and West show that if M is allowed to grow at the correct rate with the sample size (T), then their estimate is consistent. However, in the panel data setting, the number of time periods may be small. The consistency of the clustered standard error is based on the number of clusters (N) being large, as opposed to the number of time periods (T). Thus the Newey-West weighting function is unnecessary and leads to standard error estimates which have a small bias in a panel data setting.[16]

III)   Estimating Standard Errors in the Presence of a Time Effect.

---

[16] Although the bootstrap method of estimating standard errors was rarely used in the articles which I surveyed, it is another alternative for estimating standard errors in a panel data set (see for example Kayhan and Titman, 2004 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 5,000 firm-years), the estimated standard errors are the same as the OLS standard errors reported in Table I (e.g. 0.0282 for the bootstrap versus 0.0283 for OLS when $\rho_X = \rho_\epsilon = 0.50$). When I drew observations as a cluster (e.g. I drew 500 firms with replacement and took all 10 years for any firm which was drawn), the estimated standard errors are 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). They 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.

To demonstrate how the techniques work in the presence of a time effect, I generated data sets which contain only a time effect (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-MacBeth, 1973). If I assume that the panel data structure contains only a time effect, the equations I derived above are essentially unchanged. The expressions for the standard errors in the presence of only a time effect are correct once I exchange N and T.

A)   Clustered Standard Error Estimates.

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

$$\varepsilon_{it} = \delta_t + \eta_{it}$$
$$X_{it} = \zeta_t + \upsilon_{it}$$

(15)

As before, I simulated 5,000 data sets of 5,000 observations each. I allowed the fraction of variability in both the residual and the independent variable which is due to the time effect to range from zero to seventy-five percent in twenty-five percent 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 (Var(ζ)=0) or the residual (Var(δ)=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 percent of t-statistics are significant at the 1 percent level.

21

The clustered standard errors are much more accurate, but unlike our results with the firm effect, they underestimate the true standard error. The magnitude of the underestimate is small, ranging from 13 percent [1-0.1297/0.1490] when the time effect accounts for 25 percent of the variability to 19 percent [1-0.3986/0.4927] when the time effect accounts for 75 percent of the variability. The problem arises due to the limited number of clusters (e.g. years). When I estimated the standard errors in the presence of the firm effects, I had 500 firms (clusters). When I estimated the standard errors in the presence of a time effect, I have only 10 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. Based on these results, 10 clusters is too small and 500 is sufficient (see Kezdi, 2004 and Hansen, 2005).

To explore this issue, I simulated data sets of 5,000 observations with the number of years (or clusters) ranging from 5 to 100. In all of the simulations, 25 percent of the variability in both the independent variable and the residual is due to the time effect [i.e. $\rho_X = \rho_\epsilon = 0.25$]. The bias in the clustered standard error estimates declines with the number of clusters, dropping from 27 percent when there are 5 years (or clusters) to 3 percent when there are 40 years to 1 percent when there are 100 years (see 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.

B)   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 are unbiased (see equation 12).

22

This is what I find in the simulation (see 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 three percent across the simulations (using a 1 percent critical value).

IV)   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 has 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 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 – for example 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, 2005). If the time effect is fixed (e.g. equation), the time dummies completely remove the correlation between observations in the same time period. In this case, we have only a firm effect left in the data. As we saw in Section II, 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 and even standard errors clustered by firm can be biased. I will return to this issue in more detail in Sections V and VI. 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 (e.g. firm and time). Thompson (2006) and Cameron, Gelbach, and Miller (2006) proposed the following estimate of the variance-covariance matrix

$$V_{Firm \& Time} = V_{Firm} + V_{Time} - V_{White} \qquad \textbf{(16)}$$

which combines the standard errors clustered by firm with the standard errors clustered by time. The standard errors clustered by firm (the first term) captures 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) captures 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.[17]

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

---

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

firm and time effect.[18]

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

(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. $V(\gamma)=V(\delta)=Var(\eta)$ and $Var(\mu)= V(\zeta)=V(\nu)$]. I then simulated nine data sets 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, if we examine the t-statistics clustered by time, the fraction of t-statistics that are statistically significant at the one percent level (greater than 2.58) ranges from 73% when there are 1,000 time periods (and 10 firms) to 5% when there are only 10 time period (and 1,000 firms).

Clustering by two dimensions produces less biased standard errors. However, clustering by firm and time does not always yield unbiased estimates. When there are 100 firms and 100 years, one percent of the t-statistics are greater then 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 my simulations, the number of t-statistics which are greater than 2.58 rises to 5%

---

[18] 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 V, 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 Thompson (2006) and Cameron, Gelbach, and Miller (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 (2006) also generalize the procedure to allow for clustering on more than two dimensions.

when the number of firms or time periods falls to 10 (see Thompson (2006) and Colin, Gelbach, and Miller (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 1,000 firms and 10 years of data, the percent of t-statistics that are greater than 2.58 is 5% whether I cluster only by firm or when I cluster by firm and time (see Figure 7).

V)   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. $\rho(\epsilon_t, \epsilon_{t-k})$ 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.

A)   Temporary Firm Effects: Specifying the Data Structure.

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

$$\eta_{it} = \zeta_{it} \qquad\qquad\qquad if\ t=1$$
$$= \phi\,\eta_{it-1} + \sqrt{1-\phi^2}\ \zeta_{it} \qquad if\ t>1 \tag{18}$$

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.[19] 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 auto-regressive 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), I can alter the pattern of auto correlations in the residual. The correlation of lag length k is:

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

The same data structure is specified for the independent variable in the first three columns.[20] The correlation structures range from a fixed firm effect ($\rho=0.50$ and $\phi=0.00$) to a standard AR1 process

---

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

$$Var\left(\eta_{it}\right) = \sigma_\zeta^2\ if\ t=1$$
$$= \phi^2\,\sigma_\zeta^2 + \left(1-\phi^2\right)\sigma_\zeta^2 = \sigma_\zeta^2\ if\ t>1$$

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 (Williams, 2000, Jagannathan and Wang, 1998, Rogers, 1993, Andrews, 1991, Moulton, 1990, Arellano, 1987, Moulton, 1986, Liang, and Zeger, 1986). 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.

[20] In column IV, I use a 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 which I examine in Section VI (see Table 7). These autocorrelations are graphed in Figure 10.

(ρ=0.00 and φ=0.90).

B)   Fixed Effects – Firm Dummies.

A significant minority of the papers in the survey include firm dummies in their regressions. Using the simulations, I can compare the relative performance of OLS and clustered standard errors 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. Once we include the firm effects, the OLS standard error are unbiased (see Table 5 - Panel A, column I).[21] The clustered standard errors are unbiased with and without the fixed effects (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 (see Table 5 - Panel A, columns II-IV). In these simulations, the firm effect decays over time (in column II, 61 percent of the firm effect dissipates after 9 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 (Wooldridge, 2003, Baker, Stein, and Wurgler, 2003). The magnitude of the underestimation depends on the magnitude of the temporary component of the firm effect (i.e. φ), ranging from 33 percent in column III to 42 percent in column II.

C)   Generalized Least Squares Estimates of the Random Effects Model.

When the residuals of a panel regression are correlated, not only are OLS standard error

---

[21] I have assumed the model is correctly specified [i.e. Corr(X_{it}, ϵ_{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. Corr(X_{it}, ϵ_{it}) ≠ 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 opposed to OLS standard errors, since the residuals would be correlated.

estimates biased, but 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 generalized least squares approach (i.e. a panel data set with an unobserved firm effect, see Wooldridge, 2002). This is rarely done in the finance literature. Less than 3 percent of the papers in my survey used this method (for examples see Maksimovic and Phillips, 2002, Gentry, Kemsley, and Mayer, 2003, and Almazan, Brown, Carlson, and Chapman, 2004). To test the performance of this model, I estimated the random effects model estimated with feasible generalized least squares (FGLS) in each of the four simulations reported in Table 5 (see Panel B).

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). 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 (see 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).[22]

D)   Adjusted Fama-MacBeth Standard Errors.

---

[22] 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, 1990). 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, $Y_{i*} = \beta_{Within} X_{i*} + \epsilon_{i*}$]. Since the between regression contains only one observation per firm there is no within cluster correlation. The within estimates are obtained by running the regression on the deviations from firm means [e.g. $Y_{it} - Y_{i*} = \beta_{Between} (X_{it} - X_i^*) + (\epsilon_{it} - \epsilon_{i*})$]. This is identical to including firm dummies. If the firm effect is entirely fixed, then there is no within cluster correlation in this regression either. However, if the firm effect is temporary, then the residuals are still correlated within 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 Table 5 – Panel A to Panel B).

As I showed in Section II, 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 (Graham, Lemmon, and Schallheim, 1998; Christopherson, Ferson, and Glassman, 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; Schultz and Loughran, 2005). The proposed adjustment is to estimate the correlation between the yearly coefficient estimates (i.e. Corr[β_t, β_{t-1}] = θ), and then multiply the estimated variance by $(1+\theta)/(1-\theta)$ to account for serial correlation of the βs (see Chakravarty, Gulen, Mayhew, 2004 and Fama and French, 2002, footnote 1).[23] This would seem to make intuitive sense since the presence of a firm effect causes the yearly coefficient estimates to be serially correlated.

To test the merits of this idea, I use the same four simulated data structures as above (see Table 5). In each data set, I estimated ten slope coefficients and the first-order autocorrelation of the slope coefficients. I then calculate the original and an adjusted Fama-MacBeth standard error. In the first column, the fixed firm effect accounts for 50 percent of the variance. The autocorrelation is estimated very imprecisely as noted by Fama and French (2002). The 90th percentile confidence interval ranges from -0.91 to 0.91 and the mean autocorrelation is -0.12 (see Table 5 - Panel C –

---

[23] The literature has used two alternatives which 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:

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

column I). Since the average first-order auto-correlation is negative, the adjusted Fama-MacBeth standard errors are even more biased than the unadjusted standard errors. The fraction of t-statistics which are greater than 2.58 in absolute value rises from 25 to 32 percent.

The problem is the correlation being estimated (the within sample autocorrelation of βs) is not the same as the one which is causing the bias in the standard errors (the population autocorrelation of βs). The covariance that biases the standard errors and that I can estimate across the 5,000 simulations is:

$$Cov\left(\beta_t, \beta_{t-1}\right) = E\left[\left(\beta_t - \beta_{True}\right)\left(\beta_{t-1} - \beta_{True}\right)\right] \qquad \textbf{(20)}$$

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 an above average estimate of the slope coefficient in year t-1 (see 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 I estimate the true correlation between β_t and β_{t-1}, the firm effect causes this correlation to be positive and the Fama-MacBeth standard errors to be biased downward.[24]

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:

---

[24] 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 ($\rho_X = \rho_\epsilon = 0.50$). 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 I report in Table II.

$$Cov\left(\beta_t, \beta_{t-1}\right) = E\left[\left(\beta_t - \overline{\beta}_{Within\,sample}\right)\left(\beta_{t-1} - \overline{\beta}_{Within\,sample}\right)\right] \qquad \textbf{(21)}$$

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 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).[25]

The adjusted Fama-MacBeth standard errors do better when there is an auto-regressive component in the residuals (i.e. $\phi > 0$). In the three remaining simulations (Table 5 – Panel C), the estimated within sample auto correlation is positive, but the adjusted Fama-MacBeth standard errors are still biased downward. For example, in column II, 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 percent of the t-statistics are still significant at the 1 percent 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, I re-ran the simulations for the data structures in Table 5, but varied the number of time periods per firm 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 grow (see Figure 8). When the firm effect is a first-order auto-

---

[25] The average within sample serial correlation I estimate is actually less than zero, but this is due to a small sample bias. With only ten years of data per firm, I have only nine observations to estimate the serial correlation. To verify that this is correct, I re-ran the simulation using 5 to 50 years of data. As the number of years increases, the adjusted Fama-MacBeth standard error converges to the standard Fama-MacBeth standard error (see Figure 8).

regressive process (e.g. Table 5, column II), then the adjusted Fama-MacBeth standard errors do very well. The percent of t-statistics that are significant at the one percent level drops to 7 percent when there are fifty time periods per firm. When the firm effect has both a temporary and a fixed component (as we will find in the real world examples I discuss 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 percent of the t-statistics are significant at the one percent level (see Figure 8).

VI)   Empirical Applications.

I used simulated data in the previous sections. Thus, I had the advantage of knowing the data structure, which made choosing the method for estimating standard errors much easier. In real world applications, we may have priors about the data's structure (are firm effects or time effects more important, are they permanent or temporary), but we do not know the data structure for certain. Thus in this section, I will apply several of the techniques for estimating standard errors to two real data sets. This way I can demonstrate how the different methods for estimating standard errors compare, confirm that the methods used by some published papers have produced significantly biased results, and show what we can learn from the different standard errors estimates.

For both data sets, I first estimate the regression using OLS, and report White standard errors as well as standard errors clustered by firm, by time, and by both (Tables 6 and 7, columns I-IV). By using White standard errors as my 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 we know that there is a significant firm effect in the data [e.g. Corr($X_{it} \epsilon_{it}$, $X_{it-k} \epsilon_{it-k}$) $\neq 0$]. I then estimate the slope coefficients and the

standard errors using Fama-MacBeth (Tables 6 and 7, columns V). Each of the OLS regressions include time dummies. This makes the efficiency of the OLS and Fama-MacBeth coefficients similar.[26]

A)   Asset Pricing Application.

For the asset pricing example, I used the equity return regressions from Daniel and Titman (2006, "Market Reactions to Tangible and Intangible Information"). 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 changes in book and market values, and a measure of the firm's equity issuance. The data is briefly described in Appendix I 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 (ranging from three percent larger to one percent smaller). This occurs because the autocorrelation in the residuals is effectively zero (see Figure 9 - Panel A). The autocorrelation in the independent variable is large and persistent, starting at 0.98 the first month and declining to between 0.49 to 0.75 by the 24th 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

_____

[26] 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.

the Xs (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 I 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 7.2 if we use the White standard error and 1.9 if we cluster by month. This means there is a significant time effect in the data (see Figure 9 - Panel B), even after including time dummies. Any constant time effect (i.e. one which 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[$\epsilon_{it} \epsilon_{jt}$] 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 I simulated the data in Section V. Understanding a non-constant time effect is more difficult. For the time effect to be non-constant, it must be 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 non-constant time effect.

If we know the data, we can use our economic intuition to determine how the data should be organized and predict the source of the dependence within a cluster. For example, since this data contains monthly equity returns we might consider how a systematic (macro) shock would effect firm's returns differently. If the economy booms in a given month, firms in the durable goods

35

industry may have more positive returns than firms in the non-durable 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 I sort the data by month and then industry (4 digit), I see evidence of this in the autocorrelation for the residuals and the independent variables within each month (see Figure 9 - Panel B). The autocorrelations of the residual is much larger than when I sorted by firm then month (compare Figure 9 - Panels A and B) and they die away as we consider firms in more distant industries.[27]

When calculating the standard errors clustered by time, we don't need to make an assumption about how to sort the data 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 we cluster by firm or time (i.e. compare columns I to II and I to III), we can determine the nature of the dependence which remains in the residuals and this can guide us on how to improve our models.

I also estimated standard errors clustered by both firm and month (see Table 6, column IV). The standard errors clustered by firm and month are essentially identical to the standard errors clustered by month alone. 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 the data set has many firms

---

[27] A non-constant time effect can be generated by a random coefficient model (Greene, 1990). 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 { $[\beta_i$ -Average$(\beta_i)]$ Market return$_t$ }. Firms which have similar βs will have highly correlated residuals within a month, and firms which have very different βs will have residuals whose correlation is smaller. This logic suggests that I should instead sort by month and then β. When I sort this way, the autocorrelations are smaller but die away more slowly (declining from 0.030 to 0.028 at a lag of 24) than when I sorted by month and industry (declining from 0.096 to 0.042 at a lag of 24).

per month (at least 1000), we can conclude the data does not contain a significant firm effect. The fact that the serial correlation in the residuals is effectively zero (see 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 results in Sections II and III, 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 (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 to 3.4 times) as we would expect in the presence of a time effect. The Fama-MacBeth standard errors are close to the standard errors when we cluster by time, as both methods are designed to account for dependence in the time dimension. The Fama-MacBeth standard errors are consistently smaller than the clustered standard errors, but the magnitude of the difference is small (twelve to eighteen percent, compare columns III and V of Table 6).

B)   Corporate Finance Application.

For the corporate finance illustration, I used a capital structure regression since this is an example of where OLS, White, and Fama-MacBeth standard errors have been used incorrectly (see Section II-E). The independent variables are those which are common from the literature (firm size, firm age, asset tangibility, and firm profitability). The sample contains NYSE firms that pay a dividend in the previous year for the years 1965-2003. I lagged the independent variables 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 to 3.5 times larger). For example, the t-statistic on the profit margin variable is -3.1 when I use the White standard errors and -0.9 when I use the standard errors 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 10). Even after 10 years it remains above 40 percent.

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 percent). Clustering by time has little effect on the standard errors since the correlation of the residuals within a year is small. When I sort by year then industry, the first-order autocorrelation of the residuals is less than 12 percent (see Figure 10-B).

In my two examples, clustering standard errors by both firm and time appears unnecessary. In the asset pricing example, these standard errors were identical to the standard errors clustered by time, since there was no firm effect (Table 6). In the corporate finance example, they were 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 are a useful check on a researcher's results, as future research will likely demonstrate.

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 was the conclusion of Section II. In the presence of a firm effect, as in a capital structure regression, White and Fama-MacBeth standard errors are significantly biased.

VII)   Conclusions.

It is well known from first-year econometrics classes 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. In this paper, I show that 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 VI, 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 the correct decision has not always been made, as my literature survey demonstrates. The purpose of this paper is to provide such guidance. In the presence of a firm effect [e.g. $\text{Cov}(X_{it} \epsilon_{it} , X_{it-k} \epsilon_{it-k}) \neq 0$], 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.[28] 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. $\text{Cov}(X_{it} \, \epsilon_{it} \, , \, X_{kt} \, \epsilon_{kt}) \neq 0$], 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 (see 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 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.

---

[28] Skoulakis (2005) 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 one paper in the literature that has used the Fama-MacBeth approach in the way suggested by Skoulakis (see Coval and Shumway, 2005).

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 we saw in Section VI, 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 I: 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) is 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) 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 book value over the past five years. Thus if you purchased one percent of book value five years ago, and neither invested additional cash or nor took any cash out of the investment, book return is the current percentage ownership divided by the initial one percent.

Log(Market return) 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.

Share issuance is a measure of 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 percent of the firm. Thus if investor purchases 1 percent of the firm and five years later they own 0.5 percent of the firm, then 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 which 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 data for the regressions in Table 7 are constructed from Compustat and include 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 (see Fama and French, 2002). 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 one and 99th percentile. The independent variables are:

   Ln(Market Value of Assets) is 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 is defined as operating profits before depreciation (data13) divided by sales revenue (data12).

   Tangible assets is defined as property, plant, and equipment (data8) divide by the book value of total assets (data6).

   Advertising / Sales is defined as advertising expense (data45) divided by sales (data12).

   R&D / Sales is defined as R&D expenditure (data46) divided by sales (data12). If R&D is missing, it is coded as zero.

   R&D > 0 is a dummy variable equal to one if R&D expenditure is positive, and zero otherwise.

References

Almazan, Andres, Keith Brown, Murray Carlson, and David A. Chapman, 2004, "Why Constrain your Mutual Fund Manager?" *Journal of Financial Economics* 73, 289-321.

Anderson, Ronald, and David Reeb, 2004, "Founding-Family Ownership and Firm Performance: Evidence from the S&P 500" *Journal of Finance* 58, 1301-1328.

Andrews, D. W. K, 1991, "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica* 59, 817-858.

Arellano, Manuel, 1987, "Computing Robust Standard Errors for Within-Groups Estimators,"*Oxford Bulletin of Economics and Statistic* 49, 431-433.

Baker, Malcolm, Jeremy Stein, and Jeffrey Wurgler, 2003, "When Does the Market Matter? Stock Prices and the Investment of Equity-Dependent Firms" *Quarterly Journal of Economics* 118, 969-1005.

Baker, Malcolm and Jeffrey Wurgler, 2002, "Market Timing and Capital Structure" *Journal of Finance* 57, 1-32.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options" *Review of Financial Studies* 16, 101-143.

Bertrand, Marianne; Duflo, Esther and Mullainathan, Sendhil. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 2004, *119*(1), pp. 249-75.

Brockman, Paul, and Dennis Chung, 2001, "Managerial timing and corporate liquidity: evidence from actual share repurchases" *Journal of Financial Economics* 61, 417-448.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, 2006, "Robust Interference wiht Multi-way Clustering," University of California-Davis working paper.

Chakravarty, Sugato, Huseyin Gulen, and Stewart Mayhew, 2004, "Informed Trading in Stock and Option Markets" *Journal of Finance* 59, 1235-1257.

Chen, Joseph, Harrison Hong, and Jeremy Stein, 2001, "Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices" *Journal of Financial Economics* 61, 345-381

Cheng, Shijun, Venky Nagar, and Madhav Rajan, 2005, "Identifying Control Motives in Managerial Ownership: Evidence from Antitakeover Legislation" *Review of Financial Studies* 18, 637-672.

Choe, Hyuk, Bong-Chan Kho, Rene 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, Jon, Wayne Ferson, and Debra Glassman, 1998, "Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance" *Review of Financial Studies* 11, 111-142.

Cochrane, John, 2001, *Asset Pricing*, Princeton University Press, Princeton, NJ.

Coval, Joshua, and Tyler Shumway, 2005, "Do Behavioral Biases Affect Prices" *Journal of Finance* 60, 1-34.

Daniel, Kent and Sheridan Titman, 2006, "Market Reactions to Tangible and Intangible Information" *Journal of Finance* 61, 1605-1643.

Denis, David, Diane Denis, and Keven Yost, 2002, "Global Diversification, Industrial Diversification, and Firm Value" *Journal of Finance* 57, 1951-1979.

Doidge, Craig, 2004, "U.S. cross-listings and the private benefits of control: evidence from dual-class firms" *Journal of Financial Economics* 72, 519-553.

Donald, Stephen and Kevin Lang, 2001, "Inference with Difference in Differences and Other Panel Data." *forthcoming Review of Economics and Statistics*.

Efron, B., Tibshirani, R. J., 1986. Bootstrap measures for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1, 54-77.

Fama, Eugene, and J. MacBeth, 1973, "Risk, return and equilibrium: Empirical tests," *Journal of Political Economy* 81, 607-636.

Fama, Eugene, and Ken French, 2002, "Testing Tradeoff and Pecking Order Predictions about Dividends and Debt" *Review of Financial Studies* 15, 1-33.

Fama, Eugene and Kenneth French, 2001, "Disappearing dividends: changing firm characteristics or lower propensity to pay?" *Journal of Financial Economics* 60, 3-43.

Faulkender, Mike, and Mitchell Petersen, 2005, "Does the Source of Capital Affect Capital Structure?" *forthcoming Review of Financial Studies*.

Gentry, William, Deen Kemsley, and Christopher Mayer, 2003, "Dividend Taxes and Share Prices: Evidence from Real Estate Investment Trusts," *Journal of Finance* 58, 261-282.

Graham, John, Michael Lemmon, and James Schallheim, 1998, "Debt, Leases, Taxes and the Endogeneity of Corporate Tax Status," *Journal of Finance* 53, 131-162.

Greene, William, 1990, *Econometric Analysis*, Macmillian Publishing Company, New York.

Gross, David and Nicholas Souleles, 2004, "An Empirical Analysis of Personal Bankruptcy and Delinquency" *Review of Financial Studies* 15, 319-347.

Hansen, Christian, 2005, "Asymptotic Properties of a Robust Variance Matrix Estimator for Panel Data When T Is Large" *forthcoming Journal of Econometrics*.

Hausman, Jerry, 1978, "Specification Tests in Econometrics, *Econometrica* 46, 1251-1271.

Jagannathan, Ravi, and Zhenyu Wang, 1998, "An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression" *Journal of Finance* 53, 1285-1309.

Johnson, Shane, 2003, "Debt Maturity and the Effects of Growth Opportunities and Liquidity Risk on Leverage" *Review of Financial Studies* 16, 209-236.

Kayhan, Ayla and Sheridan Titman, 2004, "Firms' Histories and Their Capital Structures" *forthcoming Journal of Financial Economics*.

Kemsley, Deen and Doron Nissim, 2002, "Valuation of the Debt Tax Shield" *Journal of Finance* 57, 2045-2073.

Kezdi, Gabor, 2004, "Robust Standard Error Estimation in Fixed-Effects Panel Models", *forthcoming. Hungarian Statistical Review*.

Lakonishok, Josef, and Inmoo Lee, 2001, "Are insider trades informative?" *Review of Financial Studies* 14, 79-111.

Lamont, Owen and Chris Polk, 2001, "Does diversification destroy value? Evidence from the industry shocks" *Journal of Financial Economics* 63, 51-77.

Liang, Kung-Yee and Scott Zeger, 1986 "Longitudinal analysis using generalized linear models" *Biometrika* 73, 13-22.

MacKay, Peter, 2003, "Real Flexibility and Financial Structure: An Empirical Analysis" *Review of Financial Studies* 16, 1131-1165.

Maksimovic, Vojslav, and Gordon Phillips, 2002, "Do Conglomerate Firms Allocate Resources Inefficiently Across Industries? Theory and Evidence" *Journal of Finance* 57, 721-767.

Moulton, Brent, 1986, "Random Group Effects and the Precision of Regression Estimates" *Journal of Econometrics* 32, 385-397.

Moulton, Brent, 1990, "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units" *Review of Economics and Statistics* 72, 334-338.

Nagel, Stefan, 2005, "Short sales, institutional investors and the cross-section of stock returns" *Journal of Financial Economics* 78, 277-309.

Newey, Whitney, and Kenneth West, 1987, "A Simple, Positive Semi-Definite, Heteroscedastic and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, 703-708.

Pastor, Lubos and Pietro Veronesi, 2003, "Stock Valuation and Learning about Profitability" *Journal of Finance* 58, 1749-1789.

Pesaran, M Hashem and Ron Smith, 1995, "Estimating Long-run Relationships from Dynamic Heterogeneous Panels" *Journal of Econometrics* 68, 79-113.

Rogers, William, 1993, "Regression Standard Errors in Clustered Samples," *Stata Technical Bulletin* 13, 19-23.

Sapienza, Paola, 2004, "The effects of government ownership on bank lending" *Journal of Financial Economics* 72, 357-384.

Schultz, Paul and Tim Loughran, 2005, "Liquidity: Urban versus Rural Firms" *Journal of Financial Economics* 78, 341-374.

Skoulakis, Georgios, 2005, "Panel Data inference in Finance: Least-Squares vs Fama-MacBeth" University of Maryland working paper.

Thompson, Samuel, 2006, "Simple Formulas for Standard Errors that Cluster by Both Firm and Time," Harvard working paper.

White, Hal, 1984, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity," *Econometrica* 48, 817-838.

Williams, Rick, 2000, "A Note on Robust Variance Estimation for Cluster-Correlated Data," *Biometric*s 56, 645-646.

Wooldridge, Jeffrey, 2002, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.

Wooldridge, Jeffrey, 2003, "Cluster-sample Methods in Applied Econometrics" *American Economic Review* 93, 133-138.

Wu, YiLin, 2004, "The choice of equity-selling mechanisms" *Journal of Financial Economics* 74, 93-119.

Table 1: Estimating Standard Errors with a Firm Effect
OLS and Clustered Standard Errors

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

Source of Residual Volatility

Notes:

The table contains estimates of the coefficient and standard errors based on 5,000 simulated panel data set each of which contains 500 firms and 10 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:

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

(22)

where the fraction of X's variance which is due to a firm specific component $[Var(\mu)/Var(X)]$ varies across the columns of the table from 0 percent (no firm effect) to 75 percent and the fraction of the residual variance which is due to a firm specific component $[Var(\gamma)/Var(\epsilon)]$ varies across the rows of the table from 0 percent (no firm effect) to 75 percent. 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 percent of OLS t-statistics which are significant at the one percent 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 percent of clustered t-statistics which are significant at the one percent level is reported in square brackets. For example, when fifty percent of the variability in both the residual and the independent variable is due to the fixed firm effect ($\rho_X = \rho_\epsilon = 0.50$), 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 percent of the OLS t-statistics are greater than 2.58 in absolute value (only one percent should be), while 0.9 percent of the clustered t-statistics are greater than 2.58 in absolute value.

Table 2: Estimating Standard Errors with a Firm Effect
Fama-MacBeth Standard Errors

Avg(β_{FM}) Std(β_{FM}) Avg(SE_{FM}) % 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]

Notes:

The table contains estimates of the coefficient and standard errors based on the same 5,000 simulated panel data sets which are used in Table 1. Each data set contains 500 firms and 10 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:

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

(23)

where 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 percent (no firm effect) to 75 percent 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 percent (no firm effect) to 75 percent. The first entry is the average slope coefficient based on a Fama-MacBeth estimation (e.g. a regression is run for each of the 10 years and the estimate is the average of the 10 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 percent of Fama-MacBeth t-statistics which are significant at the one percent level (e.g.|t|>2.58) is reported in square brackets. For example, when fifty percent of the variability in both the residual and the independent variable is due to the firm effect ($\rho_X = \rho_\epsilon = 0.50$), the true standard error of the Fama-MacBeth coefficient is 0.0509. The Fama-MacBeth standard error estimate is 0.0238, and 24.6 percent of t-statistics are greater than 2.58 in absolute value.

51

Table 3: Estimating Standard Errors with a Time Effect
OLS and Clustered Standard Errors

Avg(β_{OLS}) Std(β_{OLS}) Avg(SE_{OLS}) % Sig(T_{OLS}) Avg(SE_{C}) % Sig(T_{C})		Source of Independent Variable Volatility			
		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]

Notes:

The table contains estimates of the coefficient and standard errors based on 5,000 simulated panel data set each of which contains 500 firms and 10 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:

$$X_{it} = \zeta_t + \upsilon_{it}$$
$$\varepsilon_{it} = \delta_t + \eta_{it}$$

(24)

where the fraction of X's variance which is due to a time specific component $[Var(\zeta)/Var(X)]$ varies across the columns of the table from 0 percent (no time effect) to 75 percent and the fraction of the residual variance which is due to a time specific component $[Var(\delta)/Var(\epsilon)]$ varies across the rows of the table from 0 percent (no time effect) to 75 percent. 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 percent of OLS t-statistics which are significant at the one percent 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 percent of clustered t-statistics which are significant at the one percent level (e.g. $|t|>2.58$) is reported in square brackets.

Table 4: Estimating Standard Errors with a Time 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%
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]

Source of Residual Volatility

Notes

The table contains estimates of the coefficient and standard errors based on the same 5,000 simulated panel data sets which are used in Table 3. Each data set contains 500 firms and 10 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:

$$X_{it} = \zeta_t + \upsilon_{it}$$
$$\varepsilon_{it} = \delta_t + \eta_{it}$$

(25)

where the fraction of X's variance which is due to a time specific component $[Var(\zeta)/Var(X)]$ varies across the columns of the table from 0 percent (no time effect) to 75 percent and the fraction of the residual variance which is due to a time specific component $[Var(\delta)/Var(\epsilon)]$ varies across the rows of the table from 0 percent (no time effect) to 75 percent. The first entry is the average slope coefficient based on a Fama-MacBeth estimation (e.g. the regression is run for each of the 10 years and the estimate is the average of the 10 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 percent of Fama-MacBeth t-statistics which are significant at the one percent level (e.g.|t|>2.58) is reported in square brackets.

Table 5: Estimated Standard Errors with a Non-Fixed Firm Effect
Panel A: OLS and Clustered Standard Errors

$\text{Avg}(\beta_{\text{OLS}})$ $\text{Std}(\beta_{\text{OLS}})$ $\text{Avg}(\text{SE}_{\text{OLS}})$ $\% \text{ Sig}(T_{\text{OLS}})$ $\text{Avg}(\text{SE}_C)$ $\% \text{ Sig}(T_C)$	I	II	III	IV
ρ_X / ρ_ϵ	0.50 / 0.50	0.00 / 0.00	0.25 / 0.25	0.60 / 0.35
$\varphi_X / \varphi_\epsilon$	0.00 / 0.00	0.90 / 0.90	0.75 / 0.75	0.99 / 0.81
OLS	0.9994 0.0513 0.0283 [0.1578] 0.0508 [0.0114]	1.0001 0.0659 0.0283 [0.2746] 0.0668 [0.0086]	1.0009 0.0566 0.0283 [0.1996] 0.0569 [0.0104]	0.9991 0.0677 0.0253 [0.3302] 0.0670 [0.0098]
OLS with firm dummies	1.0007 0.0299 0.0298 [0.0096] 0.0298 [0.0100]	1.0003 0.0517 0.0298 [0.1382] 0.0516 [0.0098]	1.0013 0.0442 0.0298 [0.0802] 0.0441 [0.0092]	1.0046 0.1881 0.1101 [0.1288] 0.1886 [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 0.0284 0.0283 [0.0090] 0.0282 [0.0090]	1.0003 0.0475 0.0283 [0.1240] 0.0474 [0.0100]	1.0012 0.0408 0.0283 [0.0730] 0.0408 [0.0090]	1.0006 0.0731 0.0580 [0.0388] 0.0721 [0.0112]

Panel C: Fama-MacBeth Standard Errors

Avg(β_{FM}) Std(β_{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 0.0514 0.0239 [0.2510] 0.0224 [0.3222]	1.0001 0.0660 0.0187 [0.4696] 0.0389 [0.2876]	1.0008 0.0567 0.0221 [0.3350] 0.0376 [0.2098]	0.9991 0.0667 0.0138 [0.6094] 0.0289 [0.3900]
Avg(1st order auto-correlation)	-0.1157	0.4395	0.3250	0.4389

Notes:

The table contains estimates of the coefficient and standard errors based on 5,000 simulated panel data sets each of which contains 500 firms and 10 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 auto-correlation (ϕ) varies. ρ_X (ρ_ϵ) is the fraction of the independent variable's (residual's) variance which is due to the fixed firm effect. ϕ_x (ϕ_ϵ) is the first order auto-correlation of the non-fixed portion of the firm effect of the independent variable (residual). Combining equations (4) with equation (18), the residual is specified as:

$$
\begin{aligned}
\varepsilon_{it} = \gamma_i + \eta_{it} = \gamma_i + \zeta_{it} & \qquad\qquad if\ t = 1 \\
= \gamma_i + \phi_\varepsilon\, \eta_{it-1} + \sqrt{1-\phi^2}\ \zeta_{it} & \qquad if\ t > 1
\end{aligned}
\tag{26}
$$

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 percent of OLS t-statistics which are significant at the one percent level (e.g. |t|>2.58) is reported in square brackets. The fifth entry is the average standard error clustered by firm. The percent of clustered t-statistics which are significant at the one percent 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 percent of regular GLS t-statistics which are significant at the one percent level (e.g. |t|>2.58) is reported in square brackets. The fifth entry is the average standard error clustered by firm. The percent of clustered t-statistics which are significant at the one percent 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 percent of t-statistics which are significant at the one percent level (e.g. |t|>2.58) is reported in square brackets. The fifth and sixth entry are the Fama-MacBeth standard error corrected for first order auto-correlation and the percent of t-statistics which are significant at the one percent 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 auto-correlation of β_t and β_{t-1}. The average first-order auto-correlation is reported in the last row of Panel C.

Table 6: Asset Pricing Application
Equity Returns and Asset Tangibility

	I	II	III	IV	V
Log(B/M)$_{t-5}$	0.1883[1]	0.1883[1]	0.1883[10]	0.1883[10]	0.1728[5]
	(0.0261)	(0.0270)	(0.1007)	(0.1009)	(0.0824)
Log(Book Return)	0.1946[1]	0.1946[1]	0.1946[5]	0.1946[5]	0.1691[5]
(last 5 years)	(0.0421)	(0.0433)	(0.0973)	(0.0979)	(0.0848)
Market Return	-0.3177[1]	-0.3177[1]	-0.3177[1]	-0.3177[1]	-0.3002[1]
(last 5 years)	(0.0283)	(0.0292)	(0.1092)	(0.1094)	(0.0957)
Share issuance	-0.5012[1]	-0.5012[1]	-0.5012[1]	-0.5012[1]	-0.5172[1]
	(0.0471)	(0.0466)	(0.1529)	(0.1527)	(0.1275)
R^2	0.0006	0.0006	0.0006	0.0006	0.0006
Coefficient Estimates	OLS	OLS	OLS	OLS	FM
Standard Errors	White	CL - F	CL - T	CL-F&T	FM

Notes:
The table contains coefficient and standard error estimates of the equity return regressions from Daniel and Titman (2005). The data is briefly described in Appendix I and in detail in their paper. The sample runs from July, 1968 to December, 2001 and contains 699,707 firm-month observations. The estimates in columns I-IV are OLS coefficients and the regressions contain time (month) dummies. Standard errors are reported in parenthesis. 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.

[10] significant at 10%; [5] significant at 5%; [1] significant at 1%

Table 7: Corporate Finance Application
Capital Structure Regressions (1965-2003)

	I	II	III	IV	V
Ln(MV Assets)	0.0460[1]	0.0460[5]	0.0460[1]	0.0460[5]	0.0394[1]
	(0.0055)	(0.0184)	(0.0074)	(0.0191)	(0.0076)
Ln(1+Firm Age)	-0.0432[1]	-0.0432	-0.0432[1]	-0.0432	-0.0479[1]
	(0.0084)	(0.0297)	(0.0067)	(0.0293)	(0.0077)
Profits / Sales	-0.0330[1]	-0.0330	-0.0330[1]	-0.0330	-0.0299[1]
	(0.0107)	(0.0359)	(0.0098)	(0.0357)	(0.0097)
Tangible assets	0.1043[1]	0.1043[1]	0.1043[1]	0.1043[1]	0.1158[1]
	(0.0057)	(0.0197)	(0.0083)	(0.0206)	(0.0096)
Market to book (Assets)	-0.0251[1]	-0.0251[1]	-0.0251[1]	-0.0251[1]	-0.0272[1]
	(0.0006)	(0.0020)	(0.0013)	(0.0023)	(0.0016)
Advertising / Sales	-0.3245[1]	-0.3245	-0.3245[1]	-0.3245	-0.3965[5]
	(0.0841)	(0.2617)	(0.0814)	(0.2609)	(0.1712)
R&D / Sales	-0.3513[1]	-0.3513[5]	-0.3513[1]	-0.3513[5]	-0.3359[1]
	(0.0469)	(0.1544)	(0.0504)	(0.1555)	(0.0501)
R&D > 0 (=1 if yes)	0.0177[1]	0.0177[5]	0.0177[1]	0.0177[5]	0.0126[1]
	(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

Notes:

The table contains coefficient and standard error estimates of a capital structure regressions. 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 is annual observations between 1965 and 2003. The sample contains NYSE firms which 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 Appendix I. The estimates in columns I-IV are OLS coefficients and the regressions contain time (year) dummies. Standard errors are reported in parenthesis. 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.

[10] significant at 10%; [5] significant at 5%; [1] significant at 1%

Figure 1: Residual Cross Product Matrix
Assumptions About Zero Covariances

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

Notes:

The figure shows a sample covariance matrix of the residuals. Assumptions about 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 non-zero. Standard errors clustered by firm assumes that the correlation of the residuals within the cluster may be non-zero (these elements are shaded). This cluster assumption assumes that residuals across clusters are uncorrelated. These are recorded as zero in the above matrix.

Figure 2: Distribution of Simulated T-Statistics
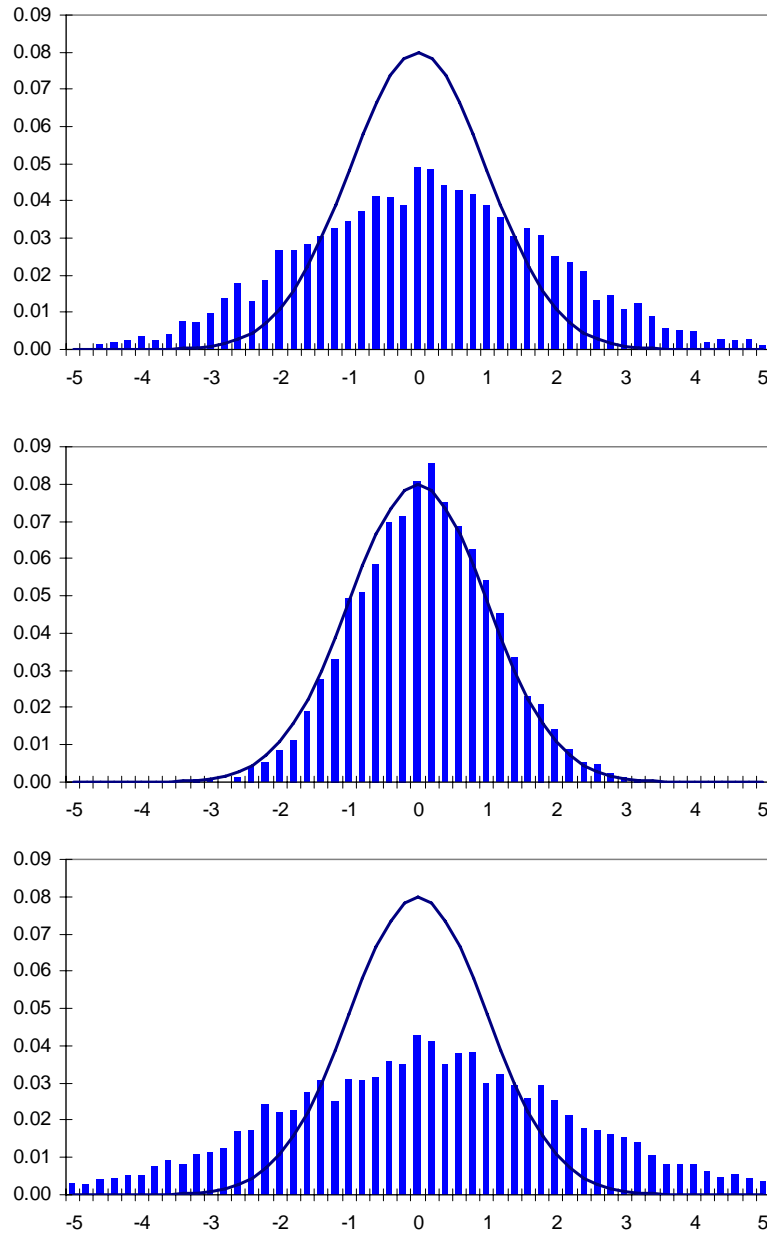
Notes:

The figures contain the theoretical t-distribution (the line), and the distribution of t-statistics produced by the simulation (the bars) when fifty percent 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 which are large in absolute value.

Figure 3: Bias in Estimated Standard Errors
as a function of years per cluster
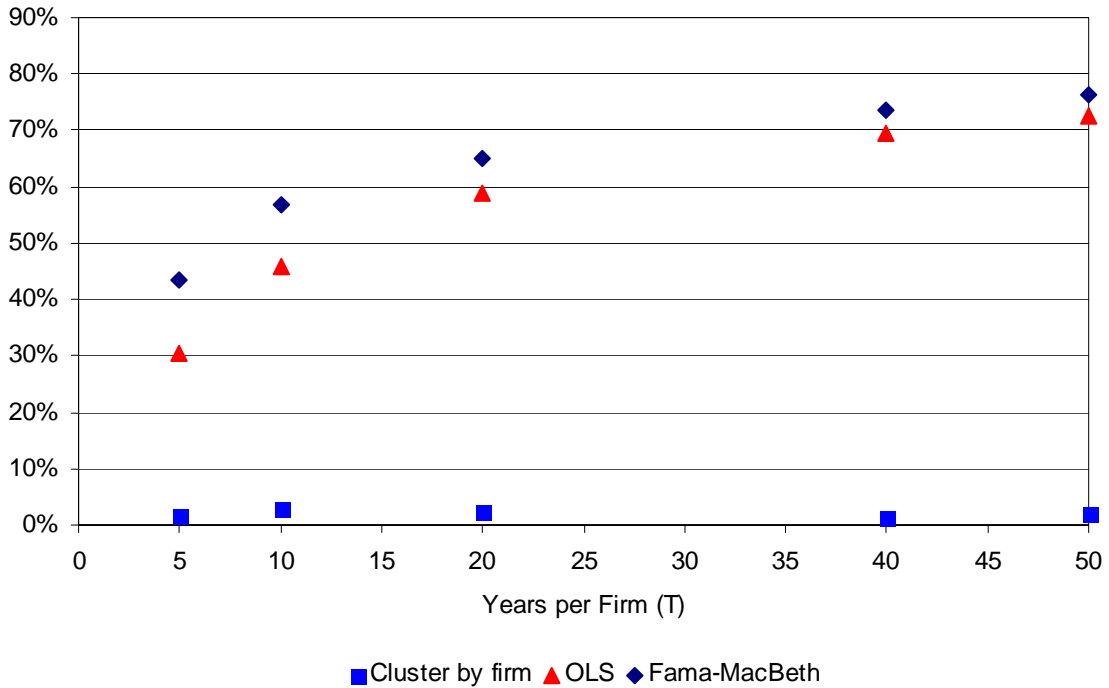
Notes:

The figure graphs the percentage by which OLS (triangles), clustered (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 5,000 simulations of a data set with 5,000 observations. The number of years per firm ranges from five to fifty. The firm effect is assumed to comprise fifty percent 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).

Figure 4: Relative Performance of OLS, Clustered, and Newey-West Standard Errors
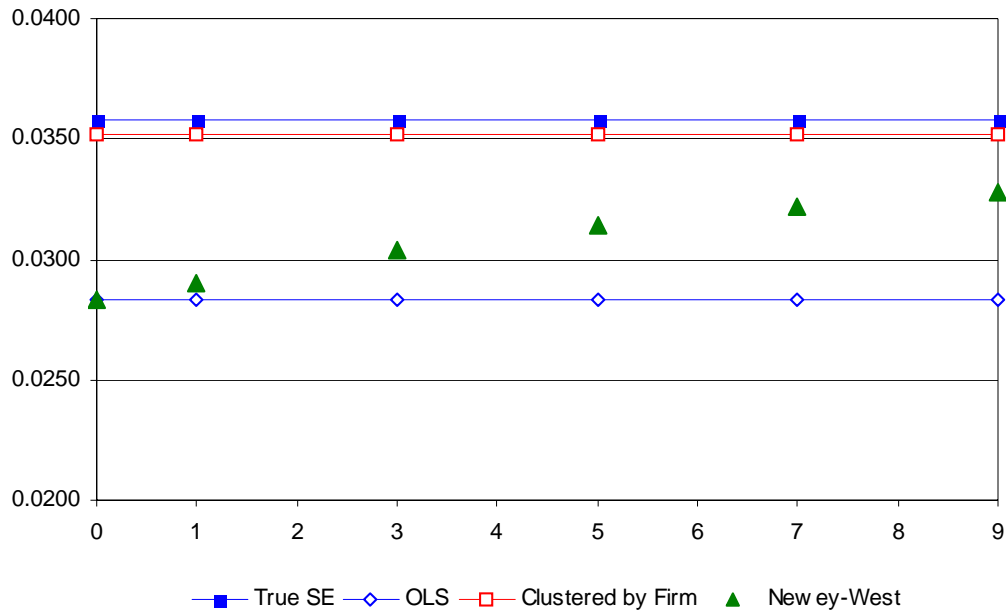
Notes:

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 5,000 simulated data sets. Each data set contains 5,000 observations (500 firms and 10 years per firm). In each simulation, twenty-five percent of the variability in both the independent variable and the residual is due to a firm effect [i.e. $\rho_X = \rho_\epsilon = 0.25$]. 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.

Figure 5: True Standard Errors and Clustered Standard Errors
as a function of cluster size (T)
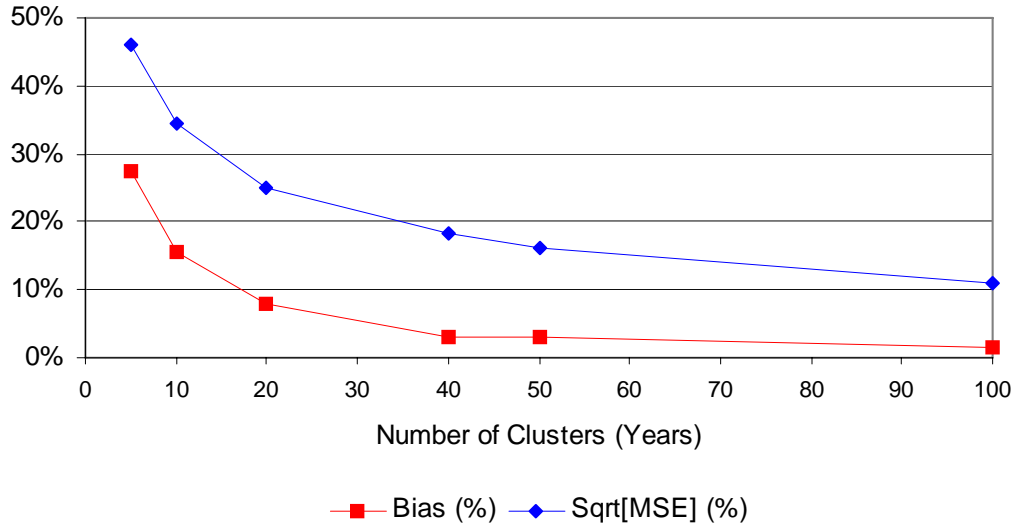
Notes:

The figure graphs the bias (squares) and mean squared error (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 mean squared error (MSE) is the average value of the squared difference between the estimated standard error and the true standard error.

$$
\begin{aligned}
MSE &= E\left[\left(\widehat{SE} - SE_{true}\right)^2\right] \\
&= E\left[\left(\widehat{SE} - \overline{SE}\right)^2 + \left(\overline{SE} - SE_{true}\right)^2\right] \\
&= Var\left(\widehat{SE}\right) + Bias\left(\widehat{SE}\right)^2
\end{aligned}
\tag{27}
$$

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 percent. Each simulated data set has 5,000 observations. In each simulation, twenty-five percent of the variability in both the independent variable and the residual is due to the time effect [i.e. $\rho_X = \rho_\epsilon = 0.25$]. The standard errors are averaged across 5,000 simulations. In these simulations, underestimation of the standard errors ranges from 27 percent when there were 5 years in the simulated data set to 1 percent when there were 100 years in the simulated data set.

Figure 6: Residual Cross Product Matrix
Firm and Time Effects

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

Notes:

    This figure shows a sample covariance matrix of the residuals. When standard errors clustered by both firm and time, residuals of the same firm in different year as well as residuals of the same year but on different firms may be non-zero. Observations on different firms and different years are assumed to be zero and are reported as zero in the above matrix.

Figure 7: Rejection Rates in the Presence of a Firm and a Time Effect
for T-Statistics Clustered by Firm, by Time, or Both
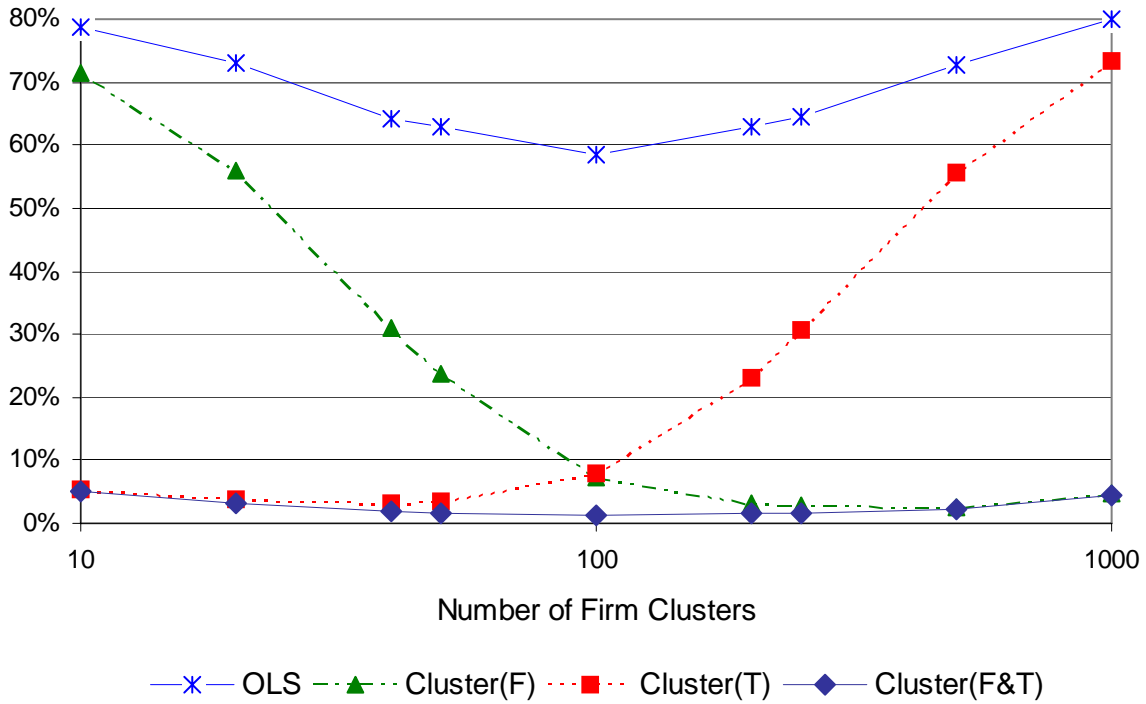
Notes:

The figure graphs the fraction of t-statistics which are statistically significant at the one percent 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 increase 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, I estimated 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).

Figure 8: Rejection Rates based on Standard and Adjusted Fama-MacBeth Standard Errors
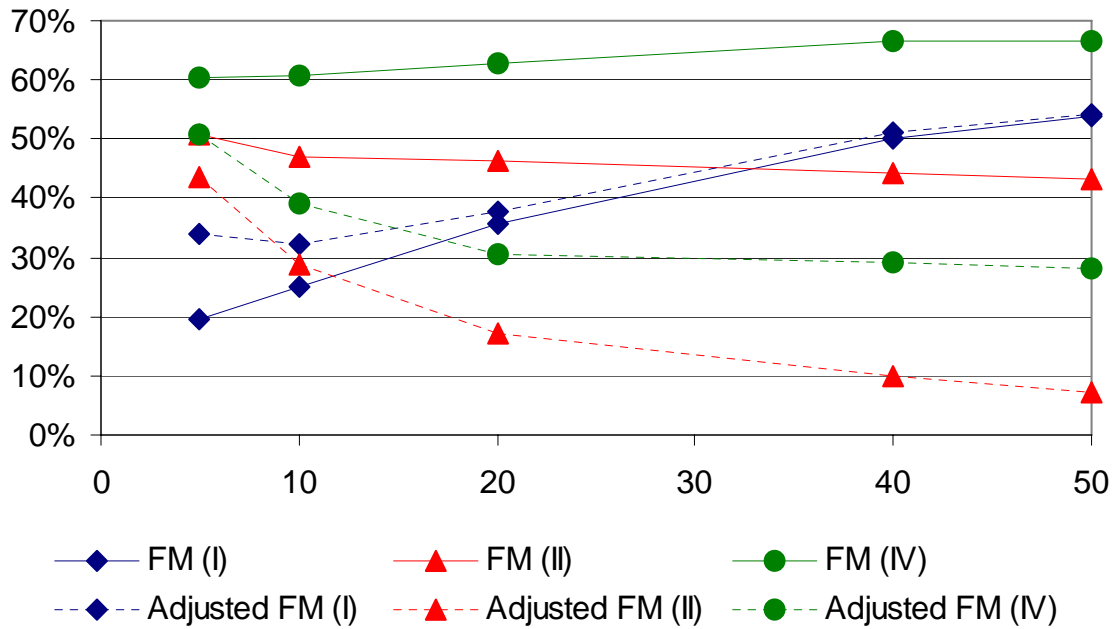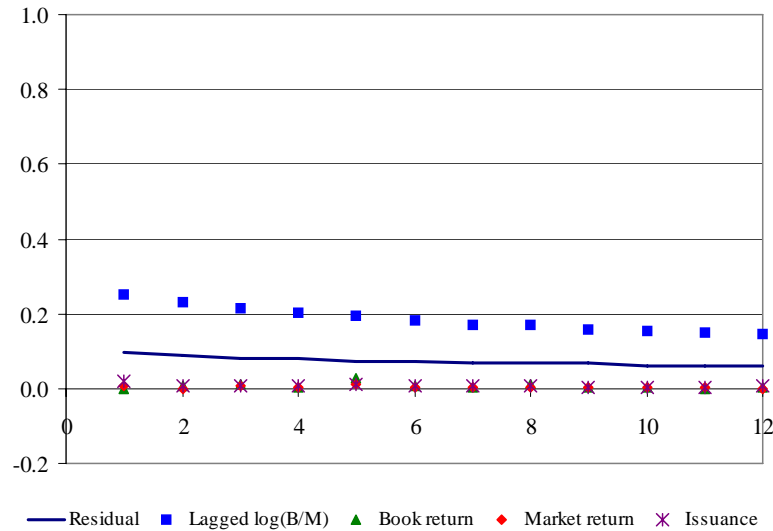

Notes:

The figure graphs the fraction of t-statistics which are statistically significant at the one percent 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 Fama-MacBeth t-statistics and Fama-MacBeth t-statistics adjusted for 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 auto-correlation ($\rho = 0.00$ and $\phi = 0.90$, column II in Table 5) are graphed with triangles. A data structure with both a fixed firm effect and an auto-regressive firm effect ($\rho_X = 0.60$, $\phi_X = 0.99$, $\rho_\epsilon = 0.35$, $\phi_\epsilon = 0.81$, column IV in Table 5) is graphed with circles. The symbols with solid lines are the rejection rates based on standard Fama-MacBeth t-statistics, while the symbols with dotted lines are the rejection rates based on adjusted Fama-MacBeth t-statistics.

Figure 9: Residuals and Independent Variables Auto-Correlation: Asset Pricing Example
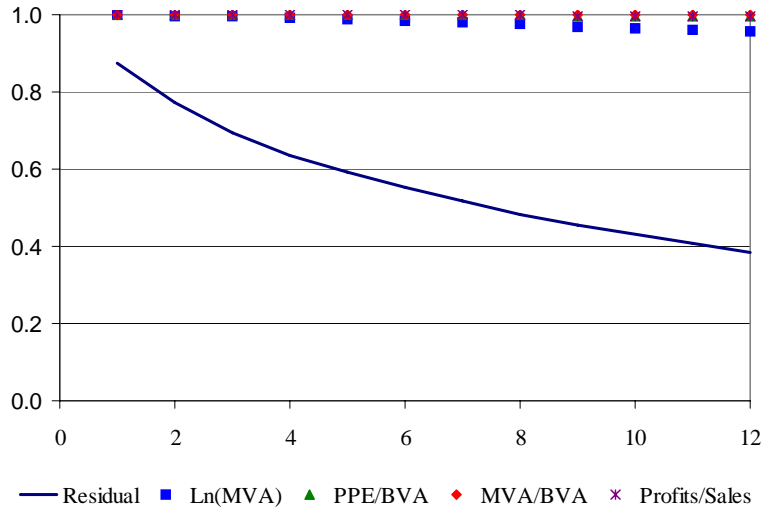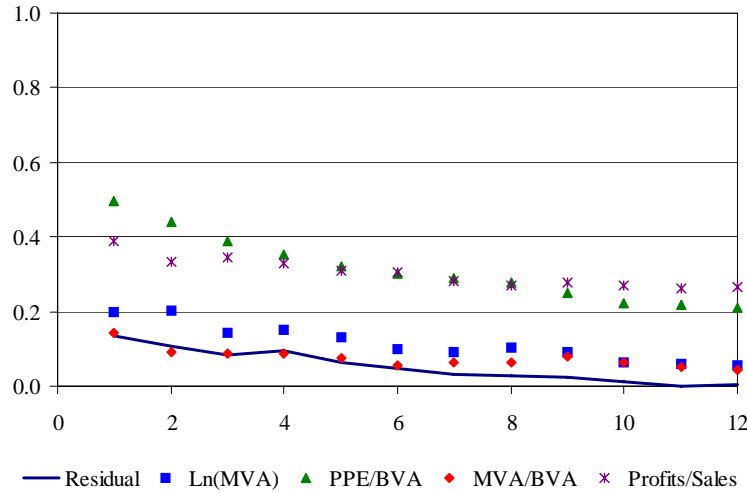Panel A: Within Firm
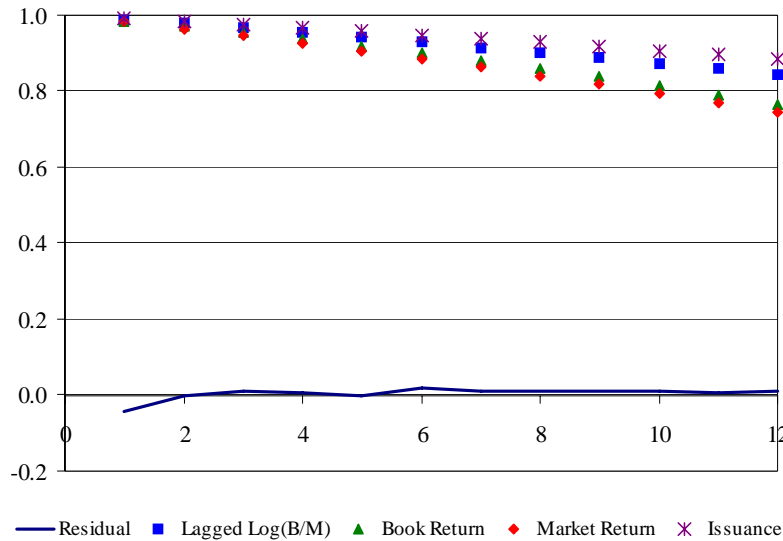
Panel B: Within Month

Notes:

The auto-correlations 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 one to twelve]. 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 one to twelve]. The data was sorted by month and then industry (4 digit) in Panel B.

Figure 10: Residuals and Independent Variables Auto-Correlation: Corporate Finance Example

Panel A: Within Firm

Panel B: Within Month

Notes:

The auto-correlations 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 ($\epsilon_{i\,t}\,\epsilon_{i\,t\text{-}k}$) for k equal one to twelve]. In Panel B, the correlations are within year and are only calculated for observations of the same year [i.e. Corr ($\epsilon_{i\,t}\,\epsilon_{i\text{-}k\,t}$) for k equal one to twelve]. The data was sorted by month and then industry (4 digit) in Panel B. The independent variables are described in Appendix I. The graph for the remaining four variables are similar and are available from the author.

71