

# Momentum, Reversals, and other Puzzles in Fama-MacBeth Cross-Sectional Regressions

Mark J. Kamstra\*

January 2018

**JEL Classification:** G12, G14

**Keywords:** momentum, reversals, autocorrelation, Fama-MacBeth

\* Schulich School of Business, York University, 4700 Keele St., Toronto ON M3J 1P3, Canada, email: mkamstra@yorku.ca. I am grateful for the helpful suggestions of Hank Bessembinder, Sreedhar Bharath, Joseph Engleberg, David Hirshleifer, Steve Heston, Christopher Hrdlicka, Christopher Jones, Alon Kfir, Lisa Kramer, Pete Kyle, Juhani Linnainmaa, Wei Lu, Michael Melvin, Seth Pruitt, Alberto Rossi, Shrihari Santosh, David Schreindorfer, Georgios Skoulakis, Zheng Sun, Allan Timmerman, Rossen Valkanov, Kirk Vandezande, Sunil Wahal, Junbo Wang, David Yang, workshop participants at Arizona State University, Brandeis University, Laval University, the University of Maryland Smith School of Business, Monash University, the UBC Sauder School of Business, the UC Irvine Merage School of Business, the UC San Diego Rady School of Business, the USC Marshall School of Business, UW Foster School of Business, and Wilfrid Laurier University, and conference participants at the Northern Finance Association 2017 meetings and the 2017 Australasian Finance & Banking Conference. This research was funded in part by the Social Sciences and Humanities Research Council of Canada and the Canadian Securities Institute Research Foundation.

# Momentum and Reversals in Fama-MacBeth Cross-Sectional Regressions

## Abstract

Evidence of momentum/reversal autocorrelation patterns in the cross-section of expected returns has challenged proponents of efficient markets for over 30 years. Stocks that have had high returns over the past year, or low returns 36 to 60 months ago, tend to have high returns going forward, without any good risk/return explanation. In this paper I show that the magnitude of time-series momentum patterns of individual firm returns has been exaggerated due to a bias in common implementations of the Fama-MacBeth procedure. This bias is so large that time-series momentum largely disappears once corrected. Interestingly, reversals become very prominent because the bias in the estimates of time-series autocorrelation coefficients is systematically positive. I outline the conditions under which this bias occurs, prove the existence of bias under these conditions, conduct Monte Carlo simulations showing the magnitude of the bias, provide and explore simple fixes for this bias, and demonstrate the impact of this bias with reference to published results in the literature. Additionally, based on insights from this observation, I introduce a Hausman test for misspecification of an asset pricing model.

# Momentum, Reversals, and other Puzzles in Fama-MacBeth Cross-Sectional Regressions

Two of the best-established anomalies in empirical asset pricing are time-series reversals and momentum in individual firm returns. These anomalies are particularly challenging to market efficiency because they appear to violate weak-form market efficiency (Fama 1970, 1991). Much of this evidence has centered on simple time-series autocorrelation from Fama-MacBeth regressions, and I will largely restrict myself to time-series issues.<sup>1</sup> I demonstrate that typical implementations of the Fama-MacBeth procedure produce upward-biased estimates of time-series autocorrelation in returns. This demonstration has three parts, first a proof of the bias under conditions that face researchers in finance, next a simulation experiment to estimate the magnitude of the bias (which is as large as 0.01-0.02 for an autoregressive coefficient estimate) and last, several empirical exercises replicating well-known studies.

A central takeaway from this study is that time-series momentum in individual firm returns largely disappears when the bias is corrected, but the evidence for reversals becomes very strong. An important qualifier is that the evidence of simple autocorrelation – time-series momentum – has exhibited substantial variation over time, and is not everywhere and always reduced to insignificance.

Consider the following proto-typical model, used by Lee and Swaminathan (2000), estimated with the Fama-MacBeth regression method:

$$r_{t+k,i} = \rho_{t,0} + \rho_{t,k}r_{t,i} + u_{t+k,i} \tag{1}$$

where  $t = 1, \dots, T$  is typically measured in months,  $i = 1, \dots, N$  indicates the cross-section of firms,  $r_{t,i}$  is the return for firm  $i$  at time  $t$ ,  $u_{t,i}$  is the firm  $i$  idiosyncratic residual, and  $\rho_{t,0}$  and  $\rho_{t,k}$  are parameters to be estimated ( $\rho_{t,0}$  is the intercept and  $\rho_{t,k}$  is the autocorrelation coefficient for lag  $k$ , both imposed to be identical for all firms in the cross-section at time period  $t$ ). Use of this model or a simple variation thereof is conventional when

---

<sup>1</sup>The two-step method of Black, Jensen, and Scholes (1972), or the weighted two-pass cross-sectional estimator of Gagliardini, Ossola, and Scaillet (2016) are sensible alternatives to the Fama-MacBeth methodology, but in spite of the many alternatives, the Fama-MacBeth cross-sectional method is still widely used, employed in papers published in virtually every issue of every top finance journal. Hence, although there are alternative methodologies, I am focusing on the Fama-MacBeth method and fixes for it.

time-series autocorrelation in firm level returns is explored with Fama-MacBeth regressions. Estimation is typically performed with OLS, with parameters estimated independently for each time period  $t$ . Then this sample over  $t = 1, \dots, T$  of cross-sectional estimates of  $\rho_{t,k}$  is used to produce the estimate of  $\rho_k$  as well as its standard error. I will refer to this as the “classic” model.<sup>2</sup> A few examples in the literature of work that employs this model specification include Jegadeesh (1990), Brennan Chordia, and Subrahmanyam (1998), Lee and Swaminathan (2000), Heston and Sadka (2008, 2010), Heston, Sadka and Thorson (2011), Novy-Marx (2012), Gong, Liu and Liu (2015), Bogousslavsky (2016), and D’Souza et al. (2016).

The intuition for the bias of the OLS estimate of time-series autocorrelation,  $\rho_k$ , produced by this method is easiest to see under the null of no autocorrelation. In this case, all firms in the cross-section are restricted to have the same mean, equal to  $\rho_{t,0}$ . If different firms have different (unconditional) expected values, the estimate of  $\rho_{t,k}$  is biased because lagged returns are a good proxy for the cross-sectional dispersion of mean returns. That is, lagged returns will (spuriously) correlate with time  $t$  returns as both have the same (unconditional) expected value.

The influence of this regression methodology is also widespread. For instance, Hong and Stein (1999) is motivated by a set of stylized facts including return continuation (momentum) in the short-to-medium run and reversals in the longer run. Their model of bounded rational agents explains these momentum and reversal effects. Bogousslavsky (2016) sets out to explain the Heston and Sadka (2008, 2010) results as an outcome of infrequent re-balancing and replicates the model and results of Heston and Sadka (2008).<sup>3</sup> I will explore a small number of these papers to see the impact of this specification error on their conclusions.

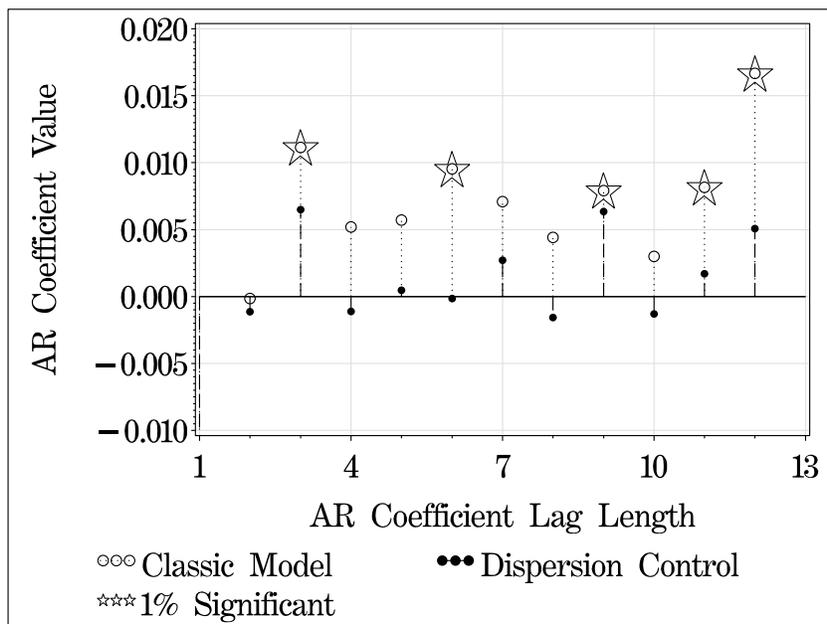
Interestingly, even if there is autocorrelation ( $\rho_k \neq 0$ ), this bias persists. A way to think about this is as an omitted variables bias. The traditional implementation of the Fama-MacBeth regression, with the CAPM beta included in the model, does not fall victim to this model misspecification problem, under the assumption that the CAPM beta measures expected returns. Indeed, any model that includes term(s) to capture the cross-sectional

---

<sup>2</sup>The classic model estimate of  $\rho_k$  can be thought of as a weighted long-short portfolio, as Heston and Sadka (2008) discuss. See Appendix A for a remark on Heston and Sadka (2008) and their interpretation of  $\rho_k$  as a feasible portfolio strategy.

<sup>3</sup>Bogousslavsky (2016) also estimates corrected models inspired by the work of Keloharju, Linnainmaa and Nyberg (2016a).

dispersion of firm expected returns will be immune to this bias.



**Figure 1**

Plots of  $\rho_k$  comparing the classic model with no control for cross-sectional dispersion of mean returns to a model with controls for cross-sectional dispersion of mean returns.

Figure 1 displays estimates of  $\rho_k$  for  $k = 1, \dots, 12$  from estimating Equation 1 with OLS (the line with circles, the “classic” model) and from a model controlling for the cross-sectional dispersion of mean returns (the line with solid dots). Note that the AR lag 1 value is truncated. The data period I use for Figure 1 covers 1962 to 2015. I describe both the data and estimation techniques in detail below. As the figure makes obvious, the failure of the classic model, Equation 1, to control for the cross-sectional dispersion of mean returns leads to a strong upward bias in the magnitude of the autocorrelation coefficient estimates, relative to the model controlling for cross-sectional dispersion.

This work contributes to the literature in empirical finance focusing on biased estimation, and methods to eliminate or reduce this bias, including Lyon, Barber, and Tsai (1999), Berk (2000), Ferson, Sarkissian, and Simin (2003, 2008), Petersen (2009), Hjalmarsson (2010), Lewellen, Nagel, and Shanken (2010), Asparouhova, Bessembinder, and Kalcheva (2010),

Kan, Robotti, and Shanken (2013), Gospodinov, Kan, and Robotti (2014), Pastor, Staambaugh, and Taylor (2015), Burt and Hrdlicka (2016), Harvey, Liu, and Zhu (2016), and Korteweg and Sorensen (2016). Distinct from this literature I explore biased estimation of the magnitude of serial correlation in individual stock returns with the use of Fama-MacBeth regressions.<sup>4</sup> Unlike Burt and Hrdlicka (2016), who document biased estimation of the predictability of firm returns in the context of information diffusion, the bias I document does not extend to analysis using long/short portfolios, but similar to Burt and Hrdlicka (2016), the bias I identify is removed by controlling for the firm-specific mean return, and this bias similarly grows more severe as cross-sectional dispersion in firm mean returns increases.

One central takeaway from the results I present is that the evidence of positive time-series momentum is much more tenuous than reported in the literature. Strong and statistically significant reversals, however, are supported by my analysis. The implications for more recent findings such as those of Heston and Sadka (2008, 2010), is that mod-12 lags *are* different from other lags, as Heston and Sadka establish in a variety of ways, and this *is* suggestive of seasonality in the cross-section of returns. But the time-series autocorrelation effects that substantiate this seasonality are from *reversals* at non-mod-12 lags, *not* the positive autocorrelation pulse at mod-12. These results also support, albeit indirectly, the notion that momentum portfolio profits, to the extent that they exist, are attributable to cross-sectional differences in expected returns, not time-series dependence in firm-specific returns.

Another takeaway is more prescriptive. The bias I identify is a function of not capturing the dispersion in the cross-section of firm-specific mean returns. Controlling for the cross-sectional dispersion of firm-specific (unconditional) expected returns can be accomplished in a variety of ways, and I find that a fixed-effects model, based on market equity ranking into deciles, or based on a five-by-five market equity crossed with book-to-market ranking, works well both in simulations I conduct and in practice. This solution has the advantage of being straightforward to apply to the Fama-MacBeth methodology. An obvious alternative, a panel/time-series estimation with firm fixed effects, also presents itself but I do not explore

---

<sup>4</sup>This bias can arise in similar situations, and has, at least occasionally, also been noted in the literature. A recent example is Korteweg and Sorensen (2016), where they identify this bias in the context of private equity performance and Fama-MacBeth-style cross-sectional regressions, producing evidence for long run performance differences.

it here.<sup>5</sup> My goal here, after documenting the bias, is to provide adjustments to the Fama-MacBeth method that correct for this bias.

I provide formal analysis to show the conditions under which OLS will produce unbiased estimates of autocorrelation in returns in Section I, and I report on Monte Carlo experiments which explore bias and test size and power in common return regression models. I outline some approaches to correct the bias in parameter estimates, describing advantages and disadvantages of each, in Section II. Sections III and IV contain demonstrations of the impact of this bias on published work and the effectiveness of corrections for this bias. I sketch out a class of asset pricing tests that exploit this OLS bias for a Hausman-type specification test in Section V, and Section VI concludes.

## I Biased Estimation

Assume the following model for time period  $t$  and cross-sections  $i = 1, \dots, N$ .

$$r_{t,i} = \rho_{0,i} + \rho_k r_{t-k,i} + \epsilon_{t,i}, \quad \rho_{0,i} = \rho_0 + \eta_i, \quad (2)$$

where

$$\eta_i, \epsilon_{t,i} \sim iid, \quad \forall t, i,$$

with both  $\eta_i$  and  $\epsilon_{t,i}$  having mean 0, variances  $\sigma_\eta^2$  and  $\sigma_\epsilon^2$ , and  $E[\epsilon_{t,i} r_{t-1,i}] = 0$ . This framework allows returns to have different means across firms ( $i$ ) and to exhibit autocorrelation.<sup>6</sup> It imposes that the parameters  $\rho_{0,i}$  and  $\rho_k$  are constant over time (though we are considering bias in the context of a single cross-sectional regression) and that variation in mean returns be independent across firms.<sup>7</sup> Estimation of  $\rho_{0,i}$  and  $\rho_k$ , say with the standard Fama-MacBeth methodology, is intractable with a single cross-section as there more parameters than observations.

Consider the estimation of  $\rho_0$  and  $\rho_k$  in the equation

---

<sup>5</sup>The large cross-section of stocks can make this procedure problematic to implement, though two-pass estimators are available. See, for instance, Gagliardini, Ossola, and Scaillet (2016).

<sup>6</sup>The structure of this model immediately invites comparison to a classic random effects model. However, as the regressor is a lagged dependent variable, independence of the regressor and  $\eta_i$  is not satisfied and a random effects model is inappropriate.

<sup>7</sup>Independence across firms is just a convenience.

$$r_{t,i} = \rho_0 + \rho_k r_{t-k,i} + \xi_{t,i}$$

where  $\xi_{t,i} = \eta_i + \epsilon_{t,i}$ . This specification is recognizable as the standard model used in the literature documenting time-series momentum and reversals, referred to as the classic model above.<sup>8</sup> Stacking elements over our parameters, writing  $\beta$  as a column vector  $\beta = \begin{pmatrix} \rho_0 \\ \rho_k \end{pmatrix}$  and stacking our regressors and dependent variable over the cross-section  $i$ , write  $X_t = (I_N \ r_{t-k})$  where  $I_N$  is an  $N \times 1$  column vector of ones and  $r_{t-k}$  (and equivalently,  $r_t$ ) is an  $N \times 1$  column vector of returns. I will suppress the subscript  $t$  on  $X$  and  $r$  for convenience in the derivations below. Write the OLS estimate of  $\beta$  (call this estimate  $\hat{\beta}$ ) using the  $N$  cross-sectional observations at time  $t$  as

$$\hat{\beta} = (X'X)^{-1} X'r$$

or

$$\hat{\beta} = (X'X)^{-1} X'(X\beta + \xi).$$

Hence

$$\hat{\beta} = \beta + (X'X)^{-1} X'(\eta + \epsilon)$$

where  $\xi$ ,  $\eta$  and  $\epsilon$  are conformably stacked vectors of  $\xi_{t,i}$ ,  $\eta_i$ , and  $\epsilon_{t,i}$ . This standard expansion makes explicit that the expectation of  $\hat{\beta}$  equals  $\beta$  if and only if both  $\epsilon$  and  $\eta$  are uncorrelated with  $X$ .

I am interested in the conditions under which the correlation of  $X$  with each of  $\epsilon$  and  $\eta$  can be assumed to equal zero. I will also need to account for the fact that  $X$  contains a lagged dependent variable, accommodated conventionally through the use of conditional expectations. Notice first that

$$E \left[ \hat{\beta} | X \right] = \beta + (X'X)^{-1} X'E[(\eta + \epsilon) | X].$$

Recall the assumption that  $E[\epsilon_{t,i} r_{t-1,i}] = 0$ , which, in our context of Fama-MacBeth regressions, is really just that return innovations (news) cannot be anticipated using past price

---

<sup>8</sup>Recall that we are estimating  $\rho_k$  with a single cross-section. Were we estimating  $\rho_k$  with a time series we would face the well-known downward bias of OLS estimation, but in the context of estimating  $\rho_k$  with a cross-section, as is the convention in this literature, we do not face this particular bias.

data. Hence  $E[\epsilon|X] = 0$  and

$$E\left[\hat{\beta} | X\right] = \beta + (X'X)^{-1} X'E[\eta|X].$$

It follows directly that the estimate of  $\rho_k$  is unbiased if  $(X'X)^{-1}$  is positive definite and  $E[\eta|X] = 0$  (or equivalently  $E[\eta_i|r_{t-k,i}] = 0$ ). Positive definiteness of  $(X'X)^{-1}$  is a standard, weak assumption to make of the data. But is it true that  $E[\eta_i|r_{t-k,i}] = 0$ ? Consider for simplicity  $k = 1$ . Notice that

$$E[r_{t-1,i}\eta_i] = E\left[E[r_{t-1,i}\eta_i|r_{t-1,i}]\right] = E\left[r_{t-1,i}E[\eta_i|r_{t-1,i}]\right]$$

by the Law of Iterated Expectations. Hence if  $E[\eta_i|r_{t-1,i}] = 0$  then  $E[r_{t-1,i}\eta_i] = 0$  and if  $E[r_{t-1,i}\eta_i] \neq 0$  then  $E[\eta_i|r_{t-1,i}] \neq 0$ . So we need only explore the value of  $E[r_{t-1,i}\eta_i]$  to determine the value of  $E[\eta_i|r_{t-1,i}]$ . Fortunately, it is straightforward to work out the value of  $E[r_{t-1,i}\eta_i]$ . Notice that

$$\begin{aligned} E[r_{t-1,i}\eta_i] &= E[(\rho_0 + \rho_k r_{t-2,i} + \eta_i + \epsilon_{t,i})\eta_i] \\ &= 0 + E[\rho_k r_{t-2,i}\eta_i] + E[\eta_i^2] + 0 \\ &= \sigma_\eta^2/(1 - \rho_k), \end{aligned}$$

as  $E[\rho_k r_{t-2,i}\eta_i] = E[\rho_k r_{t-1,i}\eta_i]$ . Hence the OLS estimate of  $\rho_k$  is biased (upwards) if expected returns vary in the cross-section of firms. Ironically, the estimate of  $\rho_k$  is biased because lagged returns proxy for the cross-sectional dispersion of mean returns. Momentum will be exaggerated in magnitude by this bias and reversals will be attenuated. Monte Carlo simulations verify that the bias can be large, in the order of .01 to .02, but more troublingly, this bias leads to strong over-rejection of the null of momentum when there is none.

## A Simulation Evidence

The bias of OLS estimates of  $\rho_k$  varies with the cross-sectional dispersion of firm mean returns, but also with the within-firm variance of returns.<sup>9</sup> This bias will also be more

---

<sup>9</sup>There is a well-known downward bias to the OLS estimate of  $\rho_k$ , which has a very small impact for autocorrelation values near 0. See Kendall (1954). As the Fama-MacBeth methodology estimates  $\rho_k$  with a single cross-section at a time and averages the coefficient estimates over many cross-sections, even this small bias does not impact the regression estimates here.

or less statistically significant, depending on the size of the cross-section and the number of cross-sections of Fama-MacBeth regressions. I will calibrate to US monthly return data to evaluate the size of the bias with a Monte Carlo simulation, looking at two cuts of the monthly CRSP files, going back to December 1925 or July 1962, and both extending up to December 2015.

I estimate with OLS an autoregressive coefficient ( $\rho_k$ ) of one lag ( $k=1$ ), setting the actual autocorrelation to values in a grid of -0.05 to 0.05, and I perform 25,000 replications, sufficient to reduce the simulation error to negligible amounts. The model I use to generate that data is

$$r_{t,i} = \rho_{0,i} + \rho_1 r_{t-1,i} + \epsilon_{t,i} \quad (3)$$

where  $\rho_{0,i} \sim N(\rho_0, \sigma_{\rho_0})$ , and  $\epsilon_{t,i} \sim N(0, \sigma_\epsilon)$ , for time periods  $t = 1, \dots, T$  and cross-sections  $i = 1, \dots, N$ , with independence imposed across time and the cross-section.<sup>10</sup>

I estimate  $\rho_1$  with the classic implementation, allowing for no cross-sectional dispersion of mean returns (Equation 4 below), and the model an econometrician could estimate with perfect knowledge of the cross-sectional dispersion of mean returns (Equation 5 below). Specifically, for each time period  $t$ , I estimate  $\rho_{0,t}$  and  $\rho_{1,t}$  for Equations 4 and 5

$$r_{t,i} = \rho_{0,t} + \rho_{1,t} r_{t-1,i} + \epsilon_{t,i} \quad (4)$$

$$r_{t,i} - \rho_{0,i} = \rho_{1,t} (r_{t-1,i} - \rho_{0,i}) + \epsilon_{t,i} \quad (5)$$

and form  $\hat{\rho}_1 = \sum_{t=1}^T \hat{\rho}_{1,t} / T$  for Equations 4 and 5, where  $\hat{\rho}_{1,t}$  is the least squares estimate of  $\rho_{1,t}$ .

Table 1 details simulation results ( $\hat{\rho}_1$ ) for values of  $\rho_1$  equal to -.05, -.01, 0, .01 and .05. The settings for simulations are indicated in the top 5 rows of the table, starting with the setting for cross-sectional variation in returns ( $\sigma_{\rho_0}$ ).

The calibration in columns 1 and 3 is to post-June 1962, NYSE-listed firms, common shares only, with a price greater than \$5. This is a sample of 4,150 firms with an average span

---

<sup>10</sup>The use of the normal distribution here is a convenient approximation to the empirical distribution of return innovations but does admit the possibility of a return less than -100%. However, because the standard deviations of  $\rho_{0,i}$  and  $\epsilon_{t,i}$  are small and the mean return far from -100%, a Monte Carlo simulation event with a return outcome less than -100% would be a 10 sigma or more event, and does not impact these simulations.

		(1)	(2)	(3)	
	$\sigma_{\rho_0} \rightarrow$	.000	.0146	.0141	
	$\rho_0 \rightarrow$	.0152	.0107	.0152	
	$\sigma_\epsilon \rightarrow$	0.109	0.172	0.109	
	T $\rightarrow$	198	196	198	
	N $\rightarrow$	1100	2400	1100	
Model	$\rho_1 \downarrow$	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)	
Classic Model	-0.05	-0.0500 (100.0)	-0.0425 (100.0)	-0.0328 (100.0)	
	-0.01	-0.0100 (99.58)	-0.0028 (47.34)	0.0067 (85.17)	
	0	0.0000 ( 4.86)	0.0071 (99.81)	0.0164 (100.0)	
	<i>No Control for Return Dispersion</i>	0.01	0.0100 (99.62)	0.0171 (100.0)	0.0263 (100.0)
	0.05	0.0500 (100.0)	0.0568 (100.0)	0.0656 (100.0)	
Correct Model	-0.05	-0.0500 (100.0)	-0.0500 (100.0)	-0.0500 (100.0)	
	-0.01	-0.0100 (99.62)	-0.0100 (100.0)	-0.0100 (99.65)	
	0	0.0000 ( 4.82)	-0.0000 ( 4.93)	-0.0000 ( 4.93)	
	<i>Controls for Dispersion with <math>\rho_{0,i}</math></i>	0.01	0.0100 (99.63)	0.0100 (100.0)	0.0100 (99.65)
	0.05	0.0500 (100.0)	0.0500 (100.0)	0.0500 (100.0)	

**Table 1 Simulation Results**

This table presents  $\rho_1$  estimates averaged over 25,000 replications, and the frequency of rejection of the null of no autocorrelation at the 5% level, two-sided. Under the null an unbiased 5% level test will reject 5% of the time, and in a simulation with 25,000 replications the standard deviation of the rejection rate is 0.1398%.

of 198 months in the data, an average monthly return of 1.52% with a standard deviation of 10.9% and a cross-sectional standard deviation of firm-specific mean returns of 1.41%. There are, on average, roughly 1,100 firms in the cross-section. The calibration in column 2

is to data from 1925 to 2015, comprising of 18,950 firms with an average span in the data of 196 months. The average monthly return is 1.07% with a standard deviation of 17.2% and the cross-sectional standard deviation of firm-specific mean returns is 1.46%. There are on average, roughly 2,400 firms in the cross-section. For both samples I restrict the data to firms with at least 5 years of monthly returns.<sup>11</sup> I set  $\rho_0$  to equal either 1.07% or 1.52%;  $\sigma_{\rho_0}$  to 0, 1.46%, or 1.41%; and  $\sigma_\epsilon$  to 0, 17.2%, or 10.9%. The values of  $T$  and  $N$  are set to  $T$  to 196 or 198 and  $N$  to either 2,400 or 1,100.

With no cross-sectional variation in returns ( $\sigma_{\rho_0} = 0$ , column 1 of the table), both estimation schemes produce unbiased estimates of  $\rho_1$  and tests of the null of  $\rho_1 = 0$  have correct size, with test rejections within 2 standard deviations of the nominal size<sup>12</sup> Consider now the case of cross-sectional variation in mean returns,  $\sigma_{\rho_0} = 1.46\%$  and  $1.41\%$ . The econometrician's model that correctly controls for cross-sectional dispersion in mean returns still has correct size and is unbiased. The classic estimation approach shows a substantial bias, between 0.71% and 1.64% from the true  $\rho_1$  value depending on the calibration, and these estimates are strongly statistically significant. The bias varies imperceptibly as the size of the cross-section changes from 1,100 to 2,400 firms, and the bias is also fairly insensitive to variation in the length of the time series and the magnitude of  $\rho_{0,i}$  (untabulated). However, the bias is very sensitive to  $\sigma_\epsilon$  (and  $\sigma_{\rho_0}$  of course). The variation in the magnitude of the bias of the  $\rho_1$  estimate in columns 2 and 3 is almost entirely a function of  $\sigma_\epsilon$ , with lower values of  $\sigma_\epsilon$  dramatically increasing this bias. This increase in bias as  $\sigma_\epsilon$  declines results from  $r_{t-1,i}$  becoming a better proxy for  $\eta_i$  as  $\sigma_\epsilon$  declines. The over-rejection of the null of no autocorrelation of the classic model is fairly steady at close to 100%. For every calibration I have investigated there is very strong over-rejection of the null. The bias is so large that statistically significant evidence of positive autocorrelation is virtually assured with cross-sections of several thousand firms even if the true return process is one of no autocorrelation.

---

<sup>11</sup>The restriction to firms with at least 5 years of data with which to calibrate these simulations reflects a trade-off between using a representative sample of firm mean returns and the precision of the estimates of the cross-sectional variability of the firm mean returns. If the sample period over which the individual firm average return is calculated is too short, such as when we include very short-lived firms, the dispersion in average returns may be dominated by the variance of the estimation error of the average return for the short-lived firms. Indeed, the cross-sectional standard deviation of mean return triples to over 4% when all firms are included.

<sup>12</sup>With a 5% nominal size and 25,000 replications, the standard deviation of the simulation error is just under 0.14%.

The positive bias of this estimate also means that small reversal effects, of 1 to 2 percent, are unlikely to be observed. In the simulations presented here, the power to discriminate  $\rho = -.01$  from 0, in the presence of cross-sectional dispersion in firm mean returns (columns 2 and 3 of Table 1) is about 47% for column 2 and 85% for column 3. The high rejection rate for column 3 of 85% comes from tests that suggest there is *positive* autocorrelation, so although we are rejecting the null, it is in favor of an incorrect alternative.

Empirical estimates of momentum effects from the literature, using the classic model presented in Equation 1, rarely find  $\hat{\rho}_1$  larger than 3 or 4%, which throws doubt on the reliability of this evidence of time-series momentum from the Fama-MacBeth cross-sectional regression methodology. Altogether, these simulation results suggest that a substantial bias may arise from the Fama-MacBeth methodology commonly used to estimate time-series momentum effects in the literature.

## II Several Simple Solutions

There are many possible solutions to this problem: we simply need a regressor that controls for the expected return in the cross-section. Asset pricing models provide candidates, such as the CAPM firm-specific beta. Another candidate is the firm-specific average of past returns, and a variation of this is used to great effect by Keloharju et al. (2016a). A very simple approach is to modify the Fama-MacBeth regression to have dummy variables for categories of firms, just as an industry fixed-effects model would use. These fixed effects would proxy for the expected return of each category and provided as we have sorted the firms into distinct expected return categories, this approach could also work, albeit with some misclassification almost guaranteed to arise, and with this misclassification, some small bias. Another solution is to abandon the Fama-MacBeth approach entirely and estimate a panel/time-series model with individual firm fixed effects. I will focus on approaches that are simple modifications to the Fama-MacBeth cross-sectional methodology, not on panel/time-series approaches, because my goal here is to suggest ways to use the Fama-MacBeth method with little or no bias.

## A Simulation Evidence: Perfect Identification of Cross-Sectional Expected Returns

The first solution I will consider is adding the average of past returns into the model of Equation 1, as in Equation 6 below. The second solution I will consider is the fixed-effects model, estimated by binning firms into quantiles with other similar expected-return firms with a separate dummy variable per quantile (Equation 7 below).

I will first assume that returns have identical distributions over time, so that the average of past returns, given enough observations, perfectly identifies the cross-sectional distribution of expected returns, and that the binning of firms into expected return quantiles is also perfect with no misclassifications. I will turn to the impact of imperfect identification of expected returns following this abstraction. The models I first investigate are:

$$r_{t,i} = \rho_0 \left( \sum_{l=1}^L r_{t-l,i} / L \right) + \rho_k r_{t-k,i} + \epsilon_{t,i} \quad (6)$$

$$r_{t,i} = \sum_{j=1}^{10} \rho_{0,j} D_{i,j} + \rho_k r_{t-k,i} + \epsilon_{t,i} \quad (7)$$

where  $L$  is the number of lags of returns used to form the average return for firm  $i$ ,  $D_{i,j}$  equals 1 when firm  $i$  is in expected return quantile  $j$ , 0 otherwise, and  $k = 1$ . In practice, the fixed-effects categories could be from an industry or size sorting. For the purposes of this simulation I will sort firms into quantiles by their true expected return value, and I will explore two cases, quintiles and deciles. The simulation will tell us how much bias is produced with this binned approximation to the true expected return. These solutions have the attractive features of ease of application and few (or no) additional parameters to estimate relative to the model of Equation 1. Although these equations could be estimated with panel/time-series methods, again, I will be considering only the Fama-MacBeth methodology, with parameters estimated individually by cross-section  $t$ , and then the estimate of the parameter  $\rho_k$  averaged across time periods.

Table 2 details simulation results for values of  $\rho_1$  ( $k=1$ ) equal to -.01, 0, and 0.01, for the models shown in Equations 6-7. The settings for simulations are indicated in the top 5 rows of the table, starting with the setting for cross-sectional variation in returns ( $\sigma_{\rho_0}$ ).

Again, for the case of no cross-sectional variation in returns ( $\sigma_{\rho_0} = 0$ , column 1 of the

		(1)	(2)	(3)
	$\sigma_{\rho_0} \rightarrow$	.000	.0146	.0141
	$\rho_0 \rightarrow$	.0152	.0107	.0152
	$\sigma_\epsilon \rightarrow$	0.109	0.172	0.109
	T $\rightarrow$	198	196	198
	N $\rightarrow$	1100	2400	1100
Model	$\rho_1 \downarrow$	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)
Average Return Model	-0.01	-0.0100 (97.14)	-0.0100 (99.96)	-0.0100 (96.99)
<i>Equation 6, L=60</i>	0	0.0000 ( 4.72)	-0.0000 ( 5.41)	0.0000 ( 4.82)
	0.01	0.0100 (96.70)	0.0100 (100.0)	0.0101 (97.24)
Fixed-Effects Model	-0.01	-0.0100 (99.58)	-0.0092 (100.0)	-0.0083 (96.75)
<i>Quintile Dummy Controls Equation 7</i>	0	-0.0000 ( 5.01)	0.0007 ( 7.78)	0.0017 (12.72)
	0.01	0.0100 (99.62)	0.0107 (100.0)	0.0117 (99.95)
Fixed-Effects Model	-0.01	-0.0100 (99.60)	-0.0097 (100.0)	-0.0093 (99.01)
<i>Decile Dummy Controls Equation 7</i>	0	0.0000 ( 4.92)	0.0003 ( 5.38)	0.0007 ( 6.10)
	0.01	0.0100 (99.61)	0.0103 (100.0)	0.0107 (99.84)

**Table 2 Simulation Results, Perfect Identification Case**

This table presents  $\rho_1$  estimates averaged over 25,000 replications, and the frequency of rejection of the null of no autocorrelation at the 5% level, two-sided. Under the null an unbiased 5% level test will reject 5% of the time, and in a simulation with 25,000 replications the standard deviation of the rejection rate is 0.1378%. The historical average return model, Equation 6, has  $L = 60$  equivalent to a 5 year period with which to form the average return.

table), each estimation scheme produces unbiased estimates of  $\rho_1$  and tests of the null of  $\rho_1 = 0$  have correct size, with test rejections within 2 standard deviations of the nominal

size<sup>13</sup>. The fixed effects models have better power than the historical mean model.

For the cases of cross-sectional variation in returns,  $\sigma_{\rho_0} = 1.46\%$  and  $1.41\%$ , we see substantial improvement over the classic model with very little or no bias in the estimate of  $\rho_1$  for each estimation scheme and only small rates of over-rejection of the null of no autocorrelation when there is, in fact, no autocorrelation. The historical average return method, Equation 6, has the best performance, with no appreciable bias and slight under-rejection of the null. The fixed-effects model with quintiles has the worse performance, displaying an over-rejection of the null of no autocorrelation of approximately 13% and an upwards bias in the estimate of  $\rho_1$  of almost 0.2% for the most difficult calibration (column 3 of the table). The fixed-effects model with deciles performs reasonably well with a small (though statistically significant) tendency to over-reject and with a very small upward bias of 0.0007 in the estimate of  $\rho_1$ , again for the calibration displayed in column 3 of the table. There is also a bias that the fixed-effects model is exposed to under the alternative that could compromise power. Since this model is not an individual-firm-effects model, similar firms are grouped and their average firm mean return is used to proxy for each individual firm mean return. This firm-effects intercept (average return over deciles or quintiles, for instance) could pick up the autocorrelation impact across these firms and under-estimate the  $\rho$  value when the alternative is true.<sup>14</sup> In any finite sample this will in fact occur, but this bias will decline linearly in magnitude with the square root of the cross-sectional number of observations. With the samples I consider in these simulations, of 1,100 and 2,400 firms, this bias in the estimate of  $\rho$  is negligible, but untabulated simulations show that this bias can be considerable if the cross-section of firms is less than a few hundred.

## **B Simulation Evidence: Imperfect Identification of Cross-Sectional Expected Returns**

The models in Equations 6 and 7 incorporate unrealistic features and should be thought of as limiting-case/best-performance examples of such approaches. I explore two types of errors of identification of the cross-section of expected returns, a simple misclassification error for the fixed-effects approach and time-varying persistent expected returns model which should

---

<sup>13</sup>With a 5% nominal size and 25,000 replications, the standard deviation of the simulation error is just under 0.14%.

<sup>14</sup>I thank David Hirshleifer for pointing out this bias.

impact the simple historical average approach.<sup>15</sup>

### B.1 Imperfect Identification of Cross-Sectional Expected Returns, Persistent and Variable Expected Return Model

A variation of Equation 1 is suggested by the work of Campbell (1991, 2001), Valkanov (2003) and Pastor and Stambaugh (2009). These authors explore the idea of time-varying persistent expected returns, and they document and exploit the existence of a negative correlation between innovations in this persistent expected return and subsequent innovations in ex post realized returns.<sup>16</sup> This negative correlation can make it so that the ex post return is white noise even when expected returns are variable and persistent. This feature of ex post return data can impact the historical average return method in practice, making it so that the ex ante individual firm historical average return is a very poor forecast of ex post firm returns and hence unlikely to capture cross-sectional variation in returns.

The model outlined in Campbell (1991) and subsequent work is well described by the following template:

$$r_{t,i} = \mu_{t-1,i} + u_{t,i}, \quad \mu_{t,i} = (1 - \phi)E_{r_i} + \phi\mu_{t-1,i} + w_{t,i} \quad (8)$$

where  $E_{r_i}$  is firm  $i$  unconditional expected return,  $cov(u_{t,i}, w_{t,i}) < 0$ , and  $-1 < \phi < 1$ .

Pastor and Stambaugh (2009) observe that estimated expected returns will depend importantly on past returns with variable and persistent expected returns. In the context of estimating expected returns using only past returns, they observe that a simple (equal-weighted) average return would not be appropriate. They find that recent returns should have negative weights, and more distant returns positive weights, when forming the expected return. Simulations of returns generated as described in Equation 8 are particularly helpful because the demonstration of bias in Section I does not apply to this more complicated conditional model. Regressing firm  $i$  returns on its lagged returns, with returns following the process described in Equation 8, produces an errors-in-variables bias as lagged returns are a noisy measure of  $\mu_{t-1,i}$ , with the noise correlated with the regression error. Complicating

---

<sup>15</sup>The notion of time-varying expected returns is now well established. See, for instance, Campbell and Cochrane (1999) who find that a slowly time-varying, countercyclical risk premium can generate substantial swings in prices at the frequency of business cycles, or Pastor and Stambaugh (2009).

<sup>16</sup>I thank Ross Valkanov for drawing my attention to this literature.

this errors-in-variables bias is the cross-sectional restriction that all firms have the same intercept (the central focus of this paper), and this leads to an analytically difficult problem. However, it is straightforward to determine that either positive or negative estimation bias of  $\rho$  in Equation 1 is possible if the return generating process is as described in Equation 8, depending on the values of parameters in Equation 8.<sup>17</sup>

The simulations presented in Table 3 set  $\phi$  to 0.9, as suggested in Pastor and Stambaugh (2009), with settings for unconditional return mean and variance unchanged relative to the simulations reported in Tables 1 and 2. These settings are detailed in the top 5 rows of Table 3. For these simulations, the volatility of the innovation to expected returns  $w$  was chosen to set the correlation of  $u_{t,i}$  and  $w_{t,i}$  to -10% or -32%. The range of correlations reported in Table 2 of Campbell (1991) is typically negative (with one exception) laying between -.106 and -.664. I restrict myself to the lower range of these estimates because the bias of the least squares estimate of  $\rho_1$  in the classic model, Equation 1, flips to negative from positive at correlations in the higher range. The bias in the estimate of  $\rho_1$  also varies with the value of  $\rho_1$ , so I report results for  $\rho_1$  equal to (-.05, -.01, 0, .01, .05), which spans the magnitude of typically reported values of  $\rho_1$  estimates.

This persistent variable return process poses a significant challenge to the historical average return model, as documented in in Table 3. Each calibration investigated, as well as unreported variations in model parameterization values, leads to significant bias for estimates of  $\rho_1$  when the true value of  $\rho_1 < 0$ , with upward biases as much as 0.0196 (and large downward biases for some calibrations). When the true value of  $\rho_1$  is zero, depending on the model parameterization, the bias in the estimate of  $\rho_1$  can be even greater than that simulated for the classic (no control) model, close to 0.018. The bias decreases as  $\rho_1$  increases, but is fairly substantial for values considered here. Over-rejection of the null of no autocorrelation when the null is true is particularly severe for the lower ranges of the correlation of  $u_{t,i}$  and  $w_{t,i}$ , but is very large for all cases explored here. Depending on  $corr(u_{t,i}, w_{t,i})$ , the historical average return correction can perform very poorly under the alternative as well.

These results for the historical average return model suggest its performance will depend critically on the return generating process, and given the work of Campbell (1991, 2001), Valkanov (2003) and Pastor and Stambaugh (2009), caution is advised. If expected returns are persistent and variable, the historical average return model will be severely impacted as

---

<sup>17</sup>I observe this feature in exploratory simulation analysis, untabulated.

		(1)	(2)	(3)
	$\sigma_{\rho_0} \rightarrow$	.000	.0146	.0141
	$\rho_0 \rightarrow$	.0152	.0107	.0152
	$\sigma_\epsilon \rightarrow$	0.109	0.172	0.109
	T $\rightarrow$	198	196	198
	N $\rightarrow$	1100	2400	1100
Model	$\rho_1 \downarrow$	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)
Average Return Model	-0.05	-0.0468 (100.0)	-0.0536 (100.0)	-0.0470 (100.0)
<i>Equation 6, L=60</i>	-0.01	-0.0066 (71.48)	-0.0136 (100.0)	-0.0065 (69.71)
<i>corr(u<sub>t,i</sub>, w<sub>t,i</sub>) = -0.32</i>	0	0.0034 (25.84)	-0.0035 (51.21)	0.0036 (28.54)
	0.01	0.0135 (99.90)	0.0065 (95.54)	0.0137 (99.92)
	0.05	0.0536 (100.0)	0.0465 (100.0)	0.0541 (100.0)
Average Return Model	-0.05	-0.0332 (100.0)	-0.0447 (100.0)	-0.0327 (100.0)
<i>Equation 6, L=60</i>	-0.01	0.0075 (81.11)	-0.0043 (67.45)	0.0082 (87.38)
<i>corr(u<sub>t,i</sub>, w<sub>t,i</sub>) = -0.10</i>	0	0.0177 (100.0)	0.0058 (90.34)	0.0185 (100.0)
	0.01	0.0278 (100.0)	0.0159 (100.0)	0.0287 (100.0)
	0.05	0.0685 (100.0)	0.0563 (100.0)	0.0696 (100.0)

**Table 3 Simulation Results, Campbell-Valkanov-Pastor-Stambaugh Persistent and Variable Expected Returns**

This table presents  $\rho_1$  estimates from the historical average return model averaged over 25,000 replications, and the frequency of rejection of the null of no autocorrelation at the 5% level, two-sided. Under the null an unbiased 5% level test will reject 5% of the time, and in a simulation with 25,000 replications the standard deviation of the rejection rate is 0.1378%. The historical average return model, Equation 6, has  $L = 60$  equivalent to a 5 year period with which to form the average return.

a candidate to control for the cross-sectional dispersion of firm mean returns.

## B.2 Imperfect Identification of Cross-Sectional Expected Returns, Misclassification of Fixed Effects

There are many ways in which firms could get misclassified while implementing the fixed-effects method; for instance, we may simply not have a firm-specific characteristic that ever reliably sorts firms into the correct expected return quantile. But misclassifications will likely be to “nearby” risk categories. I report on a simulation with a misclassification error independent of the expected return<sup>18</sup> and the two cases I look at are misclassification rates of a little over 20% and a misclassification rate of almost 50%. As the bias in the estimate of  $\rho_1$  is invariant to the value of  $\rho_1$ , I report only results for  $\rho_1$  equal to  $(-.01, 0, .01)$ . This set of values of  $\rho_1$  allows me to evaluate the bias and the rate of over-rejection (at  $\rho_1 = 0$ ) as well as the power of a test of  $\rho_1 \neq 0$ .

Table 4 details simulation results for the fixed-effects (decile) model of Equation 7. The settings for simulations are indicated in the top 5 rows of the table, starting with the setting for cross-sectional variation in returns ( $\sigma_{\rho_0}$ ). Even with a high misclassification rate of over 47%, the fixed-effects decile model performs very nearly identically to the case with no misclassification, with the bias still under 0.002, though t-tests of no autocorrelation are biased to over-reject, as much as a 10% rejection rate for a 5% test. When the alternative is true and the null should be rejected, the test has good power even for small deviations of 0.01 from the null of  $\rho = 0$ , correctly rejecting approximately 90% or more of the time when  $\rho = -.01$ , even with a misclassification rate of over 47%. I also investigated a counterfactual dummy variable (fixed-effects) model with quantile categories assigned randomly to firms. Results for this case, untabulated, indicate that a categorization uncorrelated with the dispersion in firm mean return performs almost identically to the classic implementation with no control for cross-sectional mean return dispersion for the size of cross-section looked at here (1,100 and 2,400 firms).

Altogether, these results point to a simple estimator with as few as 5 or 10 dummy variables for firm categories to pick up cross-sectional variation in expected returns. This estimator shows little bias and good power, even with a high misclassification error rate.

---

<sup>18</sup>Untabulated simulations, with a misclassification error that becomes larger as the firm-specific expected return deviates from the average (market) return and is also skewed to positive values for above median expected returns and skewed to negative values for below median expected returns, produce similar results.

		(1)	(2)	(3)
	$\sigma_{\rho_0} \rightarrow$	.000	.0146	.0141
	$\rho_0 \rightarrow$	.0152	.0107	.0152
	$\sigma_\epsilon \rightarrow$	0.109	0.172	0.109
	T $\rightarrow$	198	196	198
	N $\rightarrow$	1100	2400	1100
Model	$\rho_1 \downarrow$	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)	$\hat{\rho}_1$ (% Reject)
Fixed-Effects Model	-0.01	-.0100	-.0096	-.0091
<i>Decile Dummy Controls</i>		(97.24)	(99.94)	(94.07)
<i>Equation 7</i>	0	0.0000	0.0004	0.0009
<i>Misclassification Rate: 22.3%</i>		( 4.79)	( 5.69)	( 6.24)
	0.01	0.0100	0.0104	0.0109
		(96.82)	(100.0)	(98.82)
Fixed-Effects Model	-0.01	-.0100	-.0093	-.0083
<i>Decile Dummy Controls</i>		(97.20)	(99.91)	(89.51)
<i>Equation 7</i>	0	-.0000	0.0007	0.0016
<i>Misclassification Rate: 47.1%</i>		( 5.32)	( 6.44)	( 9.68)
	0.01	0.0100	0.0107	0.0117
		(96.64)	(100.0)	(99.48)

**Table 4 Simulation Results, Misclassification Case**

This table presents  $\rho_1$  estimates averaged over 25,000 replications, and the frequency of rejection of the null of no autocorrelation at the 5% level, two-sided. Under the null an unbiased 5% level test will reject 5% of the time, and in a simulation with 25,000 replications the standard deviation of the rejection rate is 0.1378%.

### III Impact of Bias on Empirical Estimates of Serial Correlation

In this section I consider Equations 6 (the average model) and 7 (the fixed-effects model) to evaluate how these approaches work in practice compared to the classic methodology, Equation 1. This exercise focuses on data from July 1962 onwards (though return data pre-1962 are used as conditioning information where needed), uses fixed-effects firm categories based on firm-level accounting data, and a maximum (minimum) of a twenty (five) year window to calculate the historical average firm-specific return. I also remark on results using the entire cross-section of firms available through CRSP back to 1925.

**Panel A**

1962-07-01 - 2015-12-01

Variable	Observations	Mean	Std	Min	Max	Skew	Kurt
Return	642,925	1.379	10.90	-94.01	416.94	1.263	18.91
Price	642,925	33.487	54.40	5.01	4736.0	38.120	2210.9
Market Value of Equity	642,925	4.24E6	1.67E7	1770.1	5.24E8	11.559	192.70
Book-to-Market	642,925	0.786	0.72	0.00	41.54	11.173	385.23

**Panel B**

1925-12-01 - 2015-12-01

Variable	Observations	Mean	Std	Min	Max	Skew	Kurt
Return	4,100,746	1.058	17.18	-100.0	2400.0	6.784	360.68
Price	4,112,515	26.777	852.46	0.01	226000	167.57	31906
Market Value of Equity	4,112,515	1.2E6	8.39E6	0.00	7.51E8	24.852	968.16

**Table 5 Summary Statistics**

This table presents summary statistics for the two return datasets used to produce the main empirical results documenting an apparent bias in estimates of autocorrelation coefficients.

Table 5 contains summary statistics for the monthly firm-level data. Panel A displays results for a data set restricted to common share equity (CRSP share classes 10 and 11) listed on the NYSE, with positive book-value and a minimum price of \$5, while Panel B displays results for the full set of data available. On the shorter, more recent sample, the mean returns are higher than the sample going back to 1925 at almost 1.4% per month and the volatility is much lower, at roughly 11%. All the data are positively skewed and exhibit strong kurtosis. Book-to-market values are not available in the extended sample back to 1925. For the implementation of Equation 7 on the 1925-2015 sample, the fixed-effects categories are restricted to firm market capitalization categories, while market capitalization and book-to-market (10 deciles of ME and 5x5 MExBM categories) will be used for the 1962-2015 sample. To construct the fixed-effects category breakpoints only NYSE firms were used. When book-to-market values are the basis for forming fixed-effects categories and breakpoints, the book-to-market values were lagged conventionally, by 6 to 18 months.

I provide plots of the Fama-MacBeth (1973) regression estimates of  $\rho_k$  for  $k = 1, \dots, 12$  in Figure 2. These estimates are the 1962-2015 averages of the autocorrelation estimates from the cross-sectional regressions defined in Equations 1 (the classic model), 6 (the average model), and 7 (the fixed-effects model), performed using monthly data. Panel A displays the case corresponding to Equation 1, with no control for cross-sectional dispersion of mean returns, and Panel B compares the no-control estimates (from the classic model) to those produced controlling for the cross-section of firm mean return dispersion with the historical average return,<sup>19</sup> Equation 6. Panels C and D compare the classic model estimates to those produced with the fixed-effects control for cross-sectional firm mean return dispersion corresponding to Equation 7, with intercept dummy variables for market equity (ME) decile rankings of firms (Panel C) and five-by-five market equity to book-to-market (MExBM) groupings (Panel D).

Panel A of Figure 2 displays the characteristic time-series evidence of momentum so well documented in the literature, delivered with the estimation of Equation 1 for  $k = 1, \dots, 12$ . The starred estimates are significant<sup>20</sup> at the 1% two-sided level, with lags at 1, 3, 6, 9, 11 and 12 displaying strong and statistically significant evidence of time-series momentum effects.<sup>21</sup> Panel B displays very poor results for the model using past average returns to control for the cross-sectional dispersion in firm mean returns, demonstrating very little difference compared to the classic approach with no control for cross-sectional variation in firm expected returns. To the extent there is a difference, the “correction” *increases* evidence of momentum. Although this is surprising, it is consistent with simulation evidence presented in Section II B detailing the impact on the average return correction of the persistent and variable expected return model of Campbell (1991).

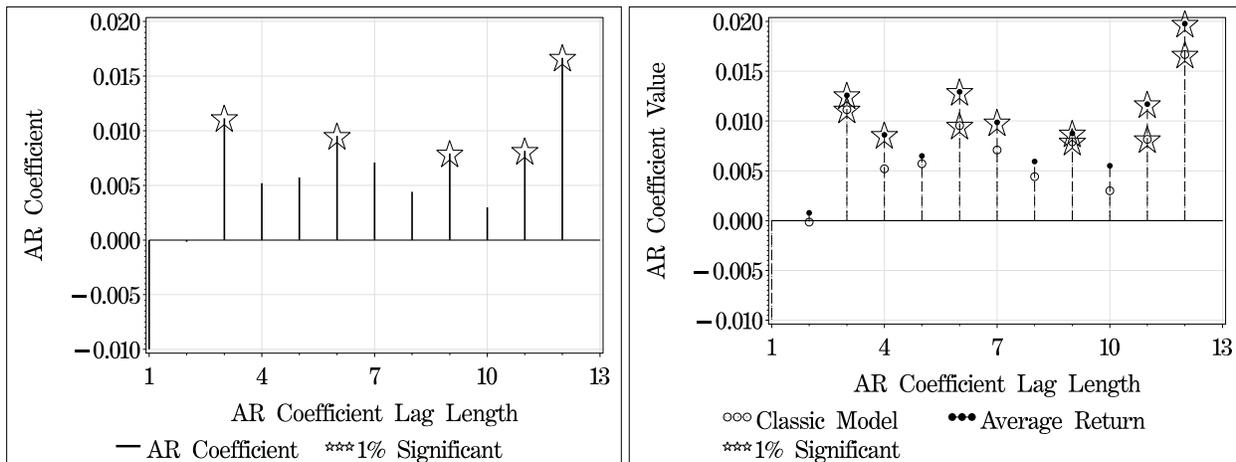
Panels C and D of Figure 2 display the autocorrelation estimates from the fixed-effects models, and here we see substantial mitigation of the momentum effect in lags 2 through 12.

---

<sup>19</sup>The firm-specific average return is calculated from as many as 20 years and no less than 5 years of past returns.

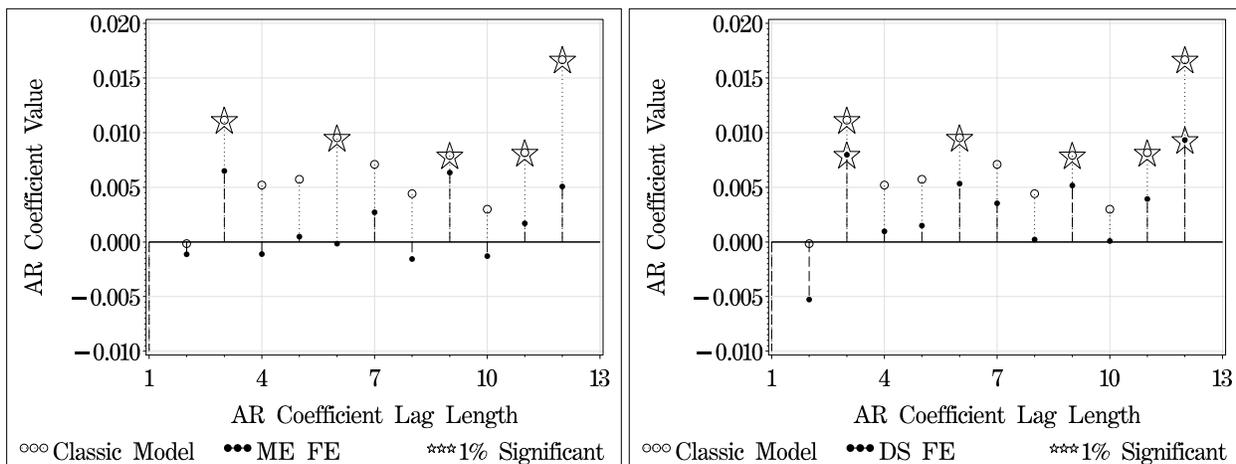
<sup>20</sup>To calculate the statistical significance of these estimates, produced through the standard Fama-MacBeth method, I follow Newey and West (1994) and estimate the standard errors with the Bartlett kernel and an automatic bandwidth parameter (autocovariance lags) equal to the integer value of  $4(T/100)^{2/9}$  where T is the sample size. Evidence of statistical significance is insensitive to choice of conventional methods with which to calculate the standard errors.

<sup>21</sup>The first lag is truncated in the figure but is also statistically significant and equal to approximately -0.04. The strong negative autocorrelation observed at lag 1 is likely an artifact of market microstructure effects. See, for instance, French and Roll (1986).



Panel A  $\hat{\rho}_k$ , Eq'n 1

Panel B  $\hat{\rho}_k$ , Eq'ns 1, 6



Panel C  $\hat{\rho}_k$  Eq'ns 1, 7 (ME Deciles)

Panel D  $\hat{\rho}_k$  Eq'ns 1, 7 (MExBM)

## Figure 2

### Raw Return Fama-MacBeth Regressions for Time Series Autocorrelation, Lags 1-36

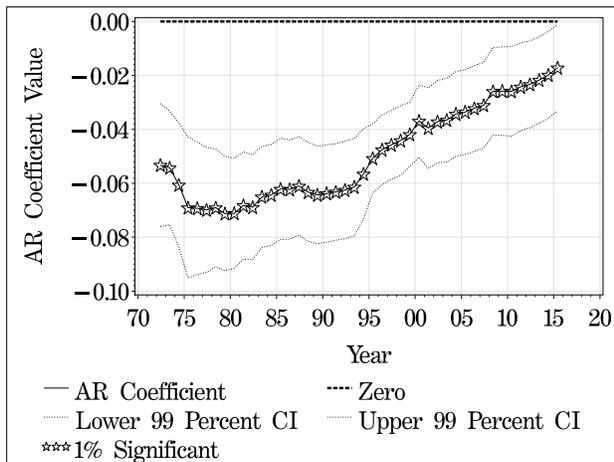
Plots of average  $\hat{\rho}_k$  from equations 1, 6, and 7, formed over 1962 through 2015. ME refers to market equity, BM refers to book-to-market equity, and MExBM refers to market equity crossed with book-to-market equity. Panel A: Equation 1, no control for cross-sectional dispersion of mean returns. Panel B: Comparison of  $\hat{\rho}_k$  from equations 1 and 6 (controls via past average return). Panel C: Comparison of  $\hat{\rho}_k$  from equations 1 and 7 (controls via intercept dummy variables for ME deciles). Panel D: Comparison of  $\hat{\rho}_k$  from equations 1 and 7 (controls via intercept dummy variables for 5x5 MExBM groupings). Values below -0.01 truncated. Summary statistics for the data used here are presented in Panel A of Table 5 and restricted to common share equity listed on the NYSE, with positive book-value and a minimum price of \$5.

None of the lag coefficients remain statistically significant for the market equity fixed effects decile model and the estimates wobble around 0 with as many negative values as positive.

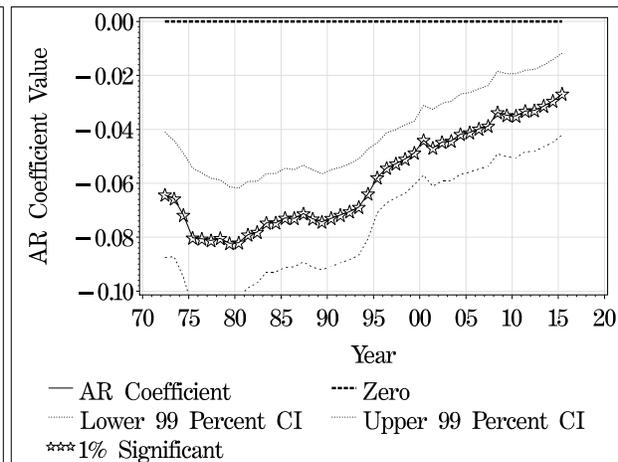
From both this and the five-by-five book-to-market and market equity fixed-effects model we see strong evidence consistent with biased estimates of momentum effects from the classic model of Equation 1.

In Figures 3 to 5, I plot the actual values of  $\rho_k$  for specific lags from the classic model, Equation 1 (Panels A and C, on the left of the figures), and the five-by-five fixed-effects model Equation 7 (Panels B and D, on the right), using a 20 year rolling window data span, for lags 1, 6, 9, 12, 24 and 36. Each point in these figures represents an average value from the previous 240 months, and each point on the confidence intervals about these values is based on this sample of 240 parameter estimates, using Newey and West (1994) robust standard errors. There are two things in particular to note from these plots. First, the estimated value from Equation 1 is larger than the estimated value from Equation 7 (the fixed-effects model) in each set of these estimated values for a given lag length, Panel A vs Panel B, and Panel C versus Panel D, of Figures 3 to 5. The difference in each case is close to 0.3%. Second, in each case the parameter estimates from the classic model with no control for cross-sectional dispersion of mean returns is always and everywhere positive except for lag 1, and is very often statistically significant. In contrast, the bias-corrected parameter estimates from Equation 7 are not only shifted downward as the analysis of Section I would imply, but are more typically statistically insignificant as well, with many instances of negative values. It is also true that many of these estimates have waned over the last 5 or 10 years of these rolling window estimations, and this is broadly consistent with the work of Robins and Smith (2016), who argue that the weekend effect has also disappeared. It is tempting to draw the conclusion that return autocorrelation is disappearing. This evidence is, however, inconsistent across lags.

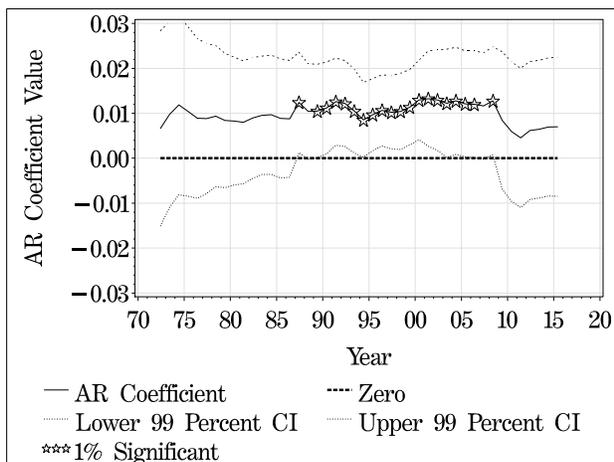
Although the use of fixed-effects controls does indeed appear to lower the autocorrelation coefficients, it is important to evaluate the statistical significance and stability of this impact. I provide evidence on the bias from estimating autocorrelation coefficients using Equation 1 (the classic model) versus Equation 7 (the fixed-effects model), in Figure 6. The bias estimate presented in Figure 6 is equal to the  $\rho_k$  estimate from the no-control model Equation 1 minus the  $\rho_k$  estimate from Equation 7 based on the five-by-five market equity value and book-to-market grouping of firms. Again, I plot 20 year rolling window average values of the quantity of interest, identical to the procedure used to form Figures 3 to 5. Given the results of Section I, I expect the differences plotted in Figure 6 to be statistically significant, and



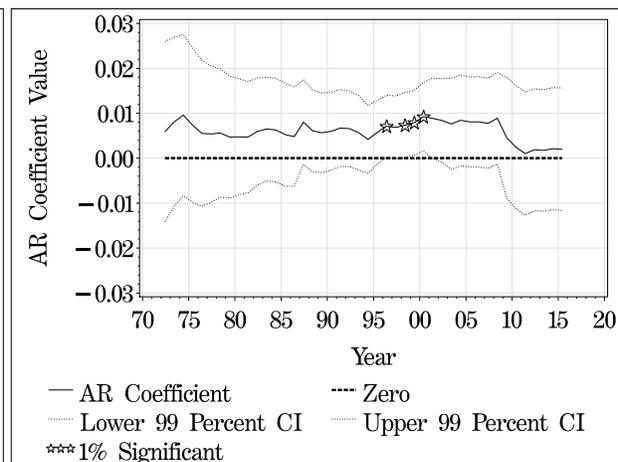
Panel A,  $\rho_1$ : Classic Model



Panel B  $\rho_1$ : MExBM



Panel C,  $\rho_6$ : Classic Model



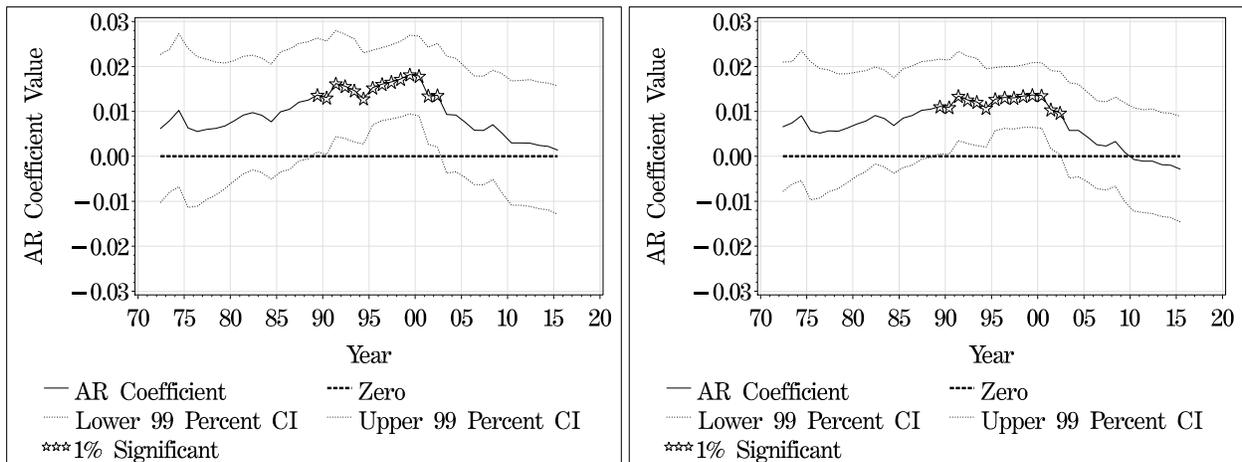
Panel D  $\rho_6$ : MExBM

### Figure 3 Parameter Estimates

Rolling Window Estimate of Average Parameter Values: Fama-MacBeth regressions on raw returns are used to produce coefficients  $\rho_k$ . Average  $\rho_k$  estimates formed with at least 10 and maximum 20 years of FM cross-sectional parameter estimates, using rolling windows starting with 1962-1972.

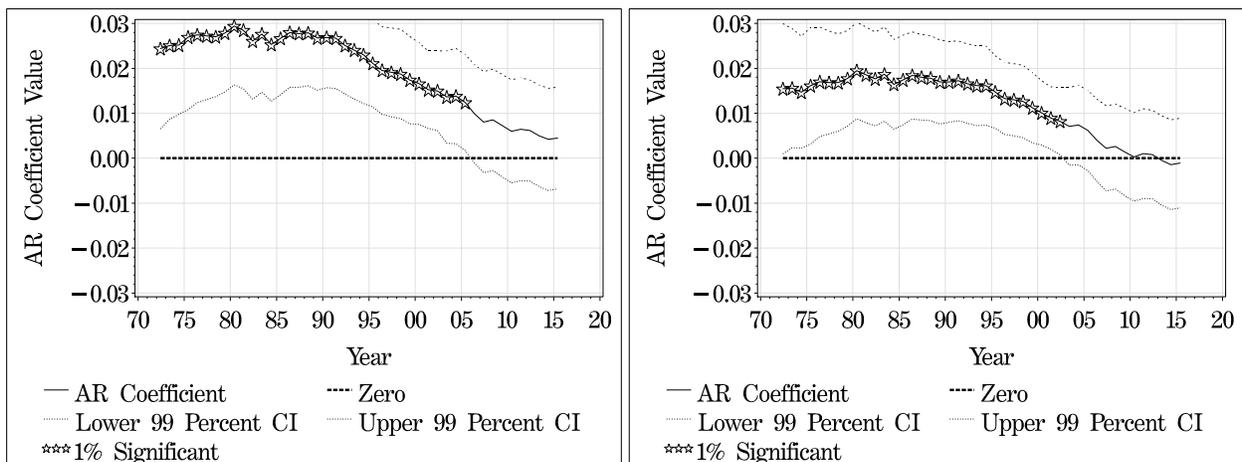
uniformly positive.

Figure 6, Panel A, is a summary plot to permit a direct comparison of the bias for several interesting collections of lags and to evaluate stability of this bias both over time and across lags. The 20 year rolling average bias over all lags (1-36) is approximately 0.3% over the 1982 to 2015 period (calculated using data from 1962-2015), shown with the solid line. The



Panel A,  $\rho_9$ : Classic Model

Panel B  $\rho_9$ : MExBM



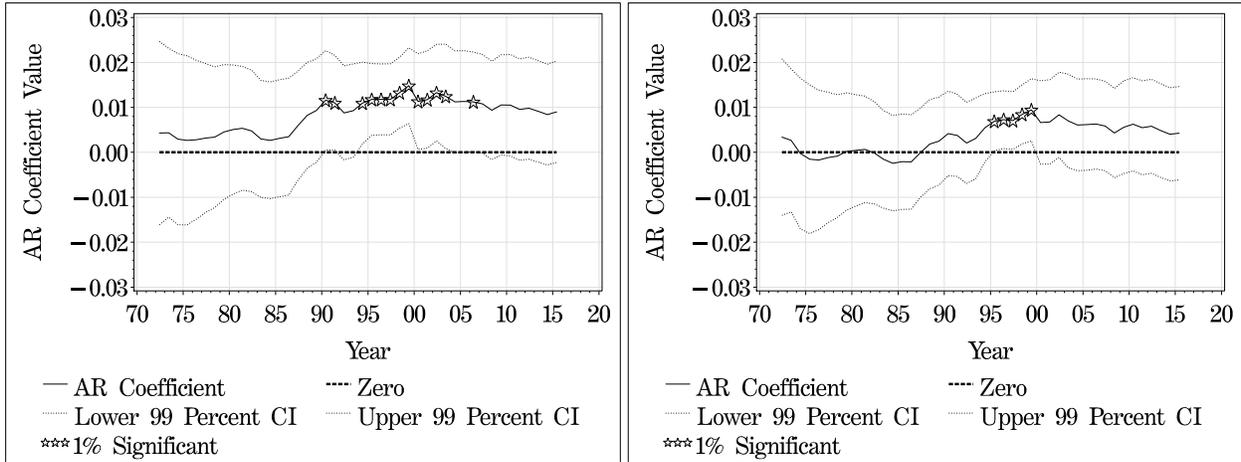
Panel C,  $\rho_{12}$ : Classic Model

Panel D  $\rho_{12}$ : MExBM

**Figure 4**  
**Parameter Estimates**

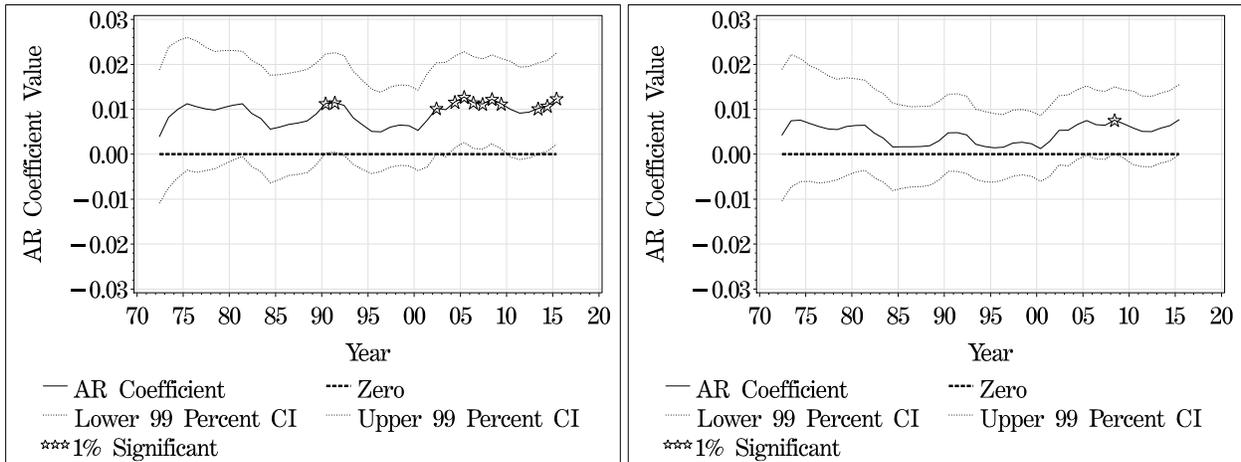
Rolling Window Estimate of Average Parameter Values: Fama-MacBeth regressions on raw returns are used to produce coefficients  $\rho_k$ . Average  $\rho_k$  estimates formed with at least 10 and maximum 20 years of FM cross-sectional parameter estimates, using rolling windows starting in with 1962-1972.

largest bias, at approximately 1%, is for the lag 1 coefficient estimate, shown with the starred line, and the bias for the lags 12, 24, and 36 coefficient estimates is also quite large at about 0.6%, shown with the line with circles. For lags 2-12 the bias lies between 0.3 and 0.5%, shown with the line with solid dots, and for longer non-annual lags 13-23 and 25-35, shown with the line with squares, the bias starts at 0 and stabilizes for much of the sample at



Panel A,  $\rho_{24}$ : Classic Model

Panel B  $\rho_{24}$ : MExBM



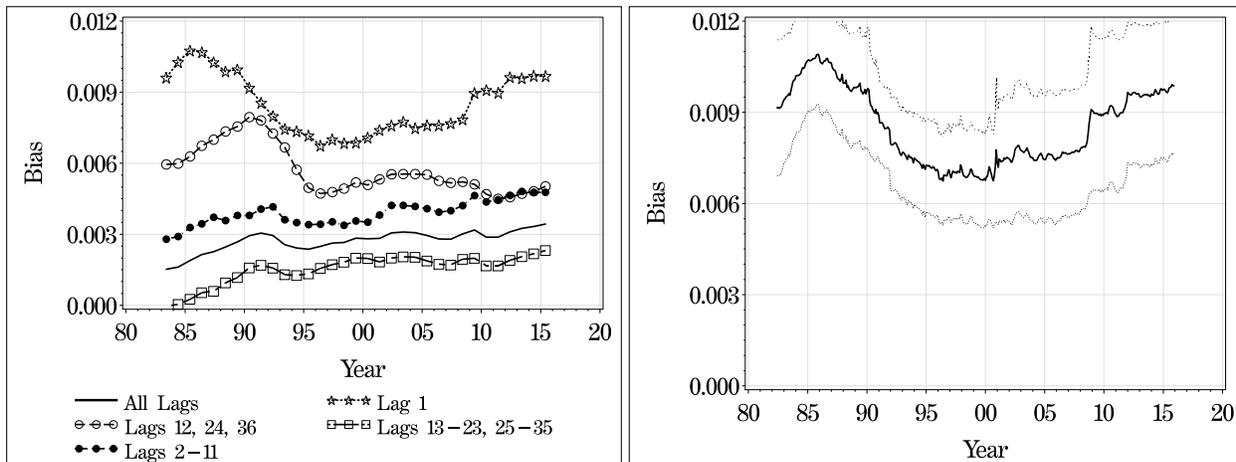
Panel C,  $\rho_{36}$ : Classic Model

Panel D  $\rho_{36}$ : MExBM

**Figure 5**  
**Parameter Estimates**

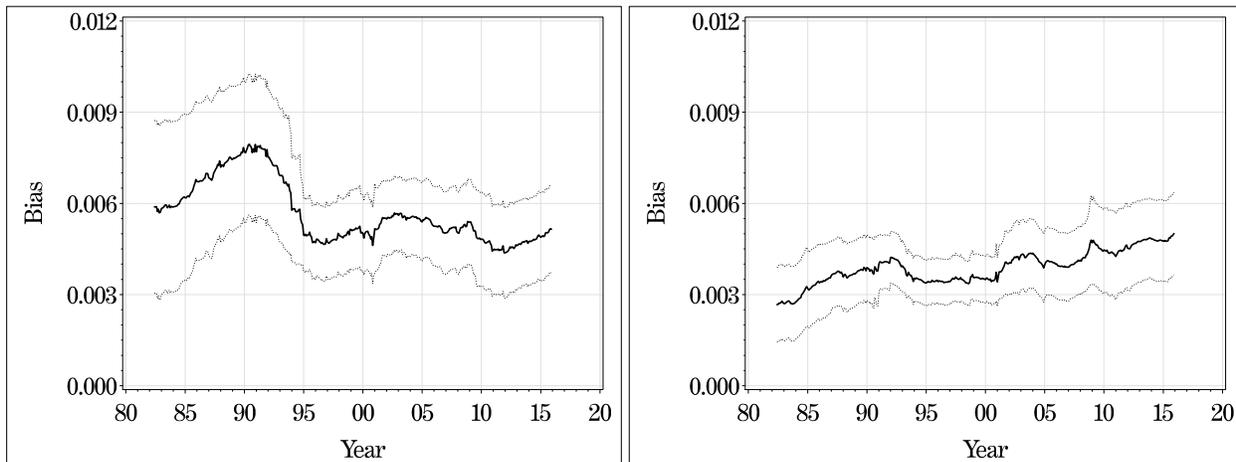
Rolling Window Estimate of Average Parameter Values: Fama-MacBeth regressions on raw returns are used to produce coefficients  $\rho_k$ . Average  $\rho_k$  estimates formed with at least 10 and maximum 20 years of FM cross-sectional parameter estimates, using rolling windows starting in 1962-1972.

about 0.25%. All of these biases move about over time, and generally move together. This movement may reflect estimation error or systematic time variation in how well the 5 by 5 double sort fixed-effects model works. Altogether, these estimates of bias from the classic no-control model of Equation 1 are of a similar magnitude to the bias that comes out of calibrated simulations, reported in Section II B.1. Figure 6, Panels B through F, present



Panel A: Bias Summary Plot

Panel B: Bias, lag 1



Panel C: Average Bias, lags 12, 24, 36

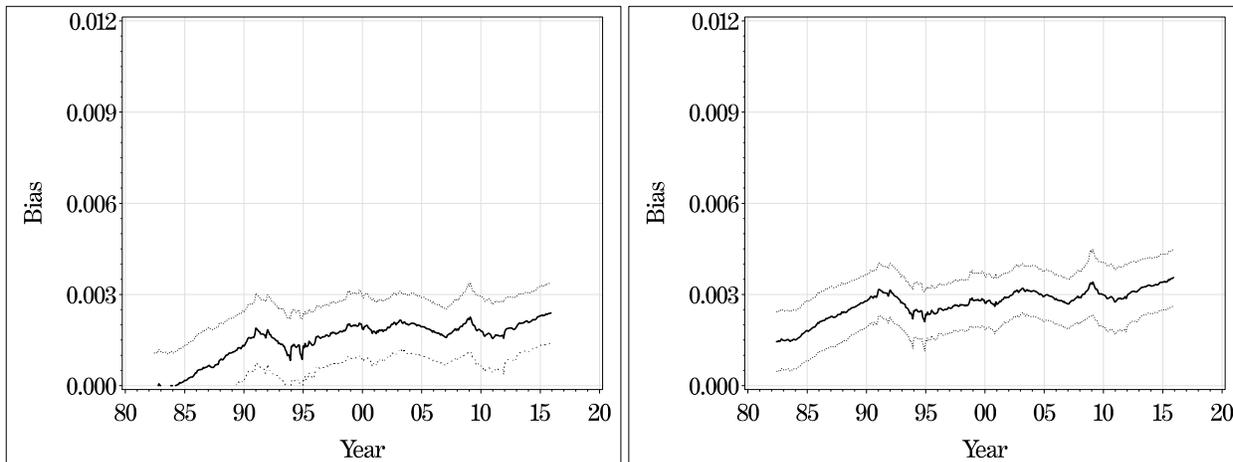
Panel D: Average Bias, lags 2-11

**Figure 6**  
**Rolling Window Estimates of the Average Bias of  $\rho_k$  Estimates, Lags  $k=1, \dots, 36$**

Fama-MacBeth regressions on raw returns are used to produce coefficients  $\rho_k$  from Equation 1 and Equation 7, with the bias estimate equal to the average of the estimate from the no-control model Equation 1 minus the estimate from Equation 7 (using the 5x5 MExBM categories). Average  $\rho_k$  and standard error estimates formed with 20 years of FM cross-sectional parameter estimates, using rolling windows. A 90% confidence interval is indicated in light dotted lines in Panels B through D. Panel A: Summary plot. Panel B: Lag 1. Panel C: Lags 12, 24, and 36. Panel D: Lags 2-12.

the average bias for these different collections of lags together with 90% confidence intervals about these estimates. This analysis suggests that the bias is not only large in magnitude

Figure 6 Continued



Panel E: Average Bias, lags 13-23, 25-35

Panel F: Average Bias, lags 1-36

Fama-MacBeth regressions on raw returns are used to produce coefficients  $\rho_k$  from Equation 1 and Equation 7, with the bias estimate equal to the average of the estimate from the no-control model Equation 1 minus the estimate from Equation 7 (using the 5x5 MExBM categories). Average  $\rho_k$  and standard error estimates formed with 20 years of FM cross-sectional parameter estimates, using rolling windows. A 90% confidence interval is indicated in light dotted lines in Panels E and F. Panel E: Lags 13-23, and 25-35. Panel F: All lags (1-36).

but also statistically significant over time, with virtually no instances of insignificance except for the collection of lags 13-23 and 25-35, and then only for the first few years of the sample.

## IV Impact of This Bias on Findings in the Literature

The use of Fama-MacBeth cross-sectional regression methodology is widespread in the empirical asset pricing literature. Even the use of this methodology applied to the study of momentum and reversals spans a great many papers, ranging in time from Jegadeesh (1990) to recent applications like Bogousslavsky (2016).

As Jegadeesh (1990) and others document, there are strong positive pulses at lags 12, 24, and 36, interrupted by reversals.<sup>22</sup> Heston and Sadka (2008) show that these positive

---

<sup>22</sup>Jegadeesh (1990) made use of a method subject to the criticism of this paper only to form forecasts of returns with which to rank stocks, and the results I have do not speak to any impact on that ranking. I am able to replicate the results of Jegadeesh (1990), and I find no evidence of bias with the model Jegadeesh

pulses at mod 12 month frequency persist for as many lags as we care to look at. Strong reversals in long run returns, from 36 to 60 months as documented in DeBondt and Thaler (1985), and at the four and five year horizons by Lee and Swaminathan (2000), are also prominent features of the data, and show up as negative coefficient estimates when  $k > 12$  in Equation 1.

I am interested in exploring the impact of a bias correction on the mod-12 pulses of strong positive autocorrelation, as well as its impact on reversals. The bias correction should reduce the mod-12 positive autocorrelation and amplify the evidence of reversals, given the results of Section I. Figure 7 provides a graphical overview of parameter estimates from Equation 1, the classic model, versus Equation 7, the fixed-effects model using the 5x5 MExBM categories. A discussion of the magnitude and statistical significance of the difference in parameter estimates of the autoregressive coefficients from Equation 1 and Equation 7 – the bias – will follow with Table 6 below.

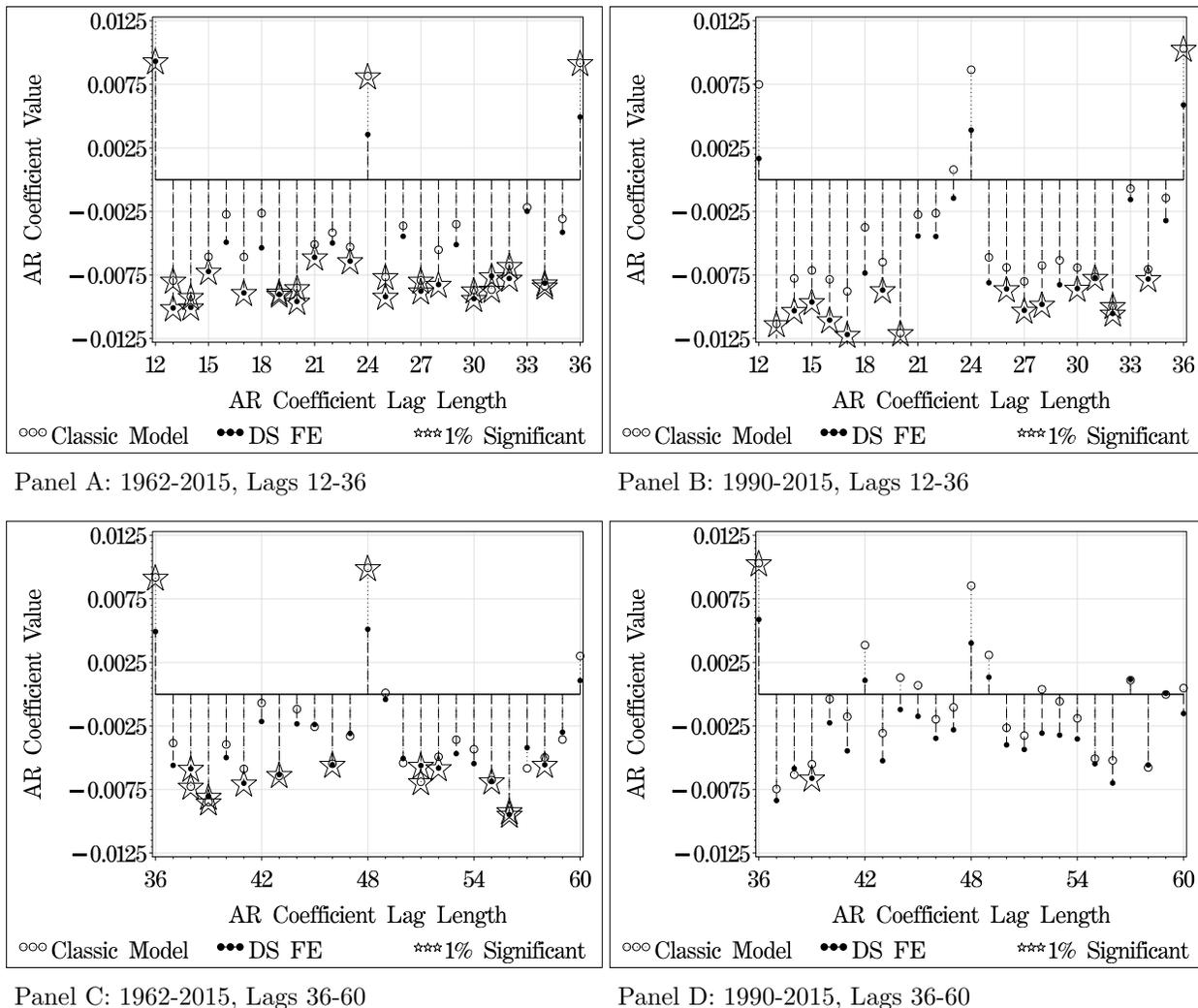
Panels A and B of Figure 7 display average autocorrelation estimates for lags 12 through 36 for two overlapping time periods, 1962-2015 and the period 1990-2015. I am interested in the subperiod 1990-2015, which follows the publication of Jegadeesh (1990), to explore if there is any impact of publication on the evidence of bias from implementing Equation 1. Panels C and D of Figure 7 display average autocorrelation estimates for lags 36 through 60.

What appears in Figure 7 is evidence of the bias produced by using the classic model versus the fixed-effects model. The biggest differences in parameter estimates from these two models is for mod-12 lags, and though the difference – the bias – is smaller for non-mod-12 lags, there is clear and strong evidence for reversals up to lag 60, strengthening with use of the fixed-effects model, Equation 7, and also strengthening for the post-publication sample 1990-2015 for lags 12 through 36. The full period results from 1962-2015 and the post-publication period 1990-2015 results are similar. As might be expected, results for the sub-sample 1962-1989 (not presented) are also similar.

Figure 8 displays only autocorrelation estimates from lags 12 through 240, mod-12, again for periods 1962-2015 and 1990-2015. Panels A and B display average autocorrelation estimates formed using the same data considered throughout this paper, data summarized in Panel A of Table 5 and restricted to common share equity listed on the NYSE, with positive

---

(1990) uses (on which he demeans the dependent variable return with mean returns from the following, not preceding, five year period).

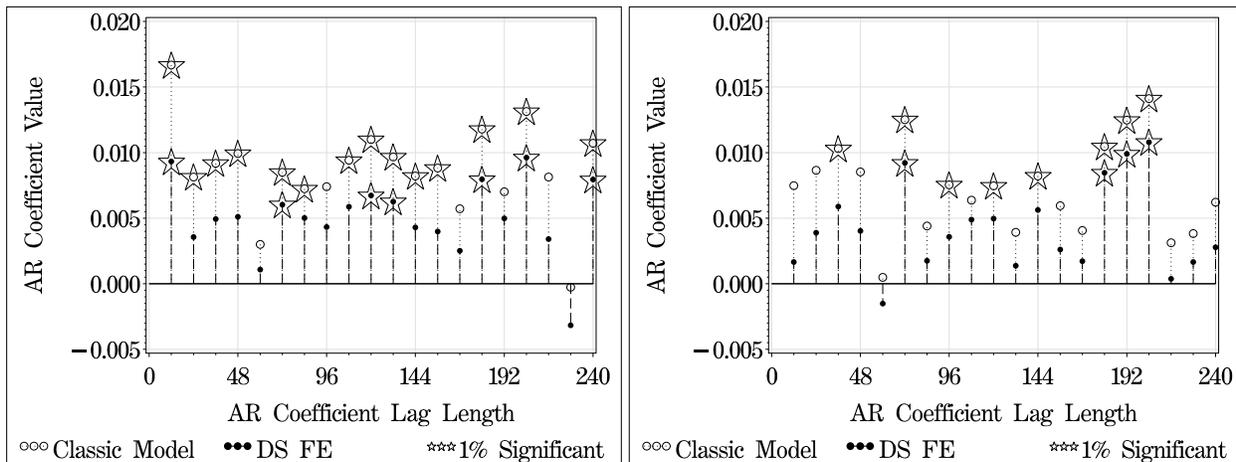


**Figure 7**  
**Raw Return Fama-MacBeth Regressions for Time Series Autocorrelation Classic Model vs. 5x5 MExBM, Lags 1-60**

Fama-MacBeth regressions on raw returns are used to produce coefficients  $\rho_k$  from Equation 1 and Equation 7. Panels A and C display average  $\rho_k$  estimates formed from cross-sectional regressions over 1962 to 2015, and Panels B and D display estimates formed over 1990 to 2015. The starred estimates are significant at the 1% two-sided level. Values below -0.0125 and above 0.0125 truncated.

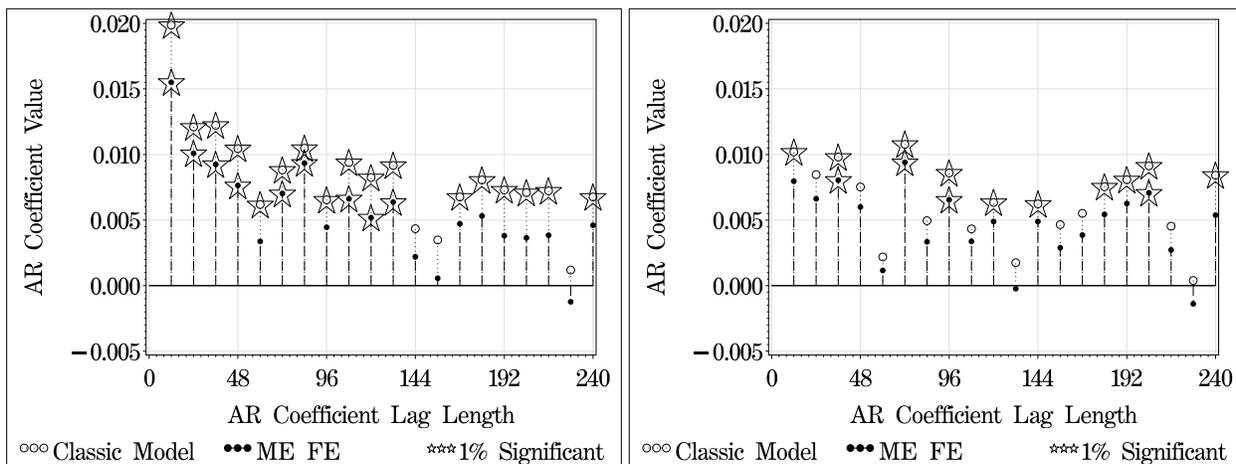
book-value and a minimum price of \$5. Based on Figure 8, the evidence of bias is strong and stable and there is little difference between the full period results from 1962-2015 and the post-publication period 1990-2015 results, Panels A and B of Figure 8.

Panels C and D of Figure 8 display average autocorrelation estimates formed using all



Panel A: 1962-2015, Lags 12-240, MExBM

Panel B: 1990-2015, Lags 12-240, MExBM



Panel C: 1925-2015, Lags 12-240

Panel D: 1990-2015, Lags 12-240

**Figure 8**  
**Raw Return Fama-MacBeth Regressions for Time Series Autocorrelation Classic Model vs. 5x5 MExBM and 10 ME, Lags 12-240**

Fama-MacBeth regressions on raw returns are used to produce coefficients  $\rho_k$  from Equation 1 and Equation 7. Only mod-12 lags are displayed, lags 12, 24, 36, .... The implementation of Equation 7 exploits a 5x5 MExBM grouping of firms. Panels C and D display average autocorrelation estimates formed using all firm data available in CRSP. As book value is not available for all firms, the implementation of Equation 7 exploits a 10 decile grouping based on lagged market value. The starred estimates are significant at the 1% two-sided level. In all cases pre-sample returns were used to form lags to 240.

firm data available, data summarized in Panel B of Table 5. Because book value is not available for all firms and not available back to 1925, the implementation of Equation 7 exploits a 10 decile grouping based on lagged market value. This different cut of the data

helps answer the question, is this evidence of bias an artifact of considering only larger, successful firms, or is it a stable feature of implementing Equation 1 with return data? The analysis of Section I should hold for any sample of return data, given we have sufficient data to precisely estimate the autocorrelation parameters. Also, of course, if the expected return of a group of firms was identical or very similar, as indeed it may be for smaller, early-stage firms, evidence of bias from using Equation 1 would not appear.

When we expand the sample period back to 1925, displayed in Panels C and D of Figure 8, and make use of all available individual firm return data, we see very similar evidence of upward biased estimates from Equation 1 versus Equation 7. Making use of the fixed-effects model, Equation 7, removes much of the evidence of statistically significant positive pulses of autocorrelation at mod-12, though in all cases and samples I have presented to this point, the mod-12 lag autocorrelation point estimates are very nearly everywhere positive.

I am also interested in possible patterns for evidence of bias in the point estimates of autocorrelation from Equation 1. Figures 7 and 8 certainly suggest patterns are likely to be observed, and Figure 6 provides evidence that various collections of lags display statistically significant estimation bias from using the classic model rather than a fixed-effects model. I will consider the bias estimate on individual autocorrelation estimates to evaluate if there is evidence of statistically significant bias that is uniform across lags. The simple analysis from Section I implies that all lags would be similarly impacted but the evidence presented so far, as well as the model of persistent and variable expected returns of Campbell (1990), leaves ambiguous the impact across lags. To the extent that some lags are helpful in predicting expected returns and some are not, we may see variable evidence of bias. For instance, the work of Heston and Sadka (2008) and Keloharju et al. (2016a) suggests that mod-12 lags are helpful to predict returns, and other lags are not. If true, I expect to see large reductions in estimated autocorrelation coefficients at the mod-12 lags and little or no impact elsewhere. The results I present in Table 6 are broadly consistent with this, but there are some interesting exceptions.

Table 6, Panels A and B, provide point estimates of the difference in estimates from Equation 1 and 7 (using the 5x5 MExBM categories) for lags 1 through 60. Bolded figures are significant at the 1% level, two-sided, and negative values are italicized. Row averages and column averages are also provided.

The row estimates inform us on the magnitude and significance of the bias for sets of 12

**Panel A**

		1962-07 - 2015-12											
Column →	1	2	3	4	5	6	7	8	9	10	11	12	Row Mean
Lags ↓													
1-12	<b>.84</b>	<b>.46</b>	<b>.31</b>	<b>.42</b>	<b>.38</b>	<b>.40</b>	<b>.36</b>	<b>.35</b>	<b>.25</b>	<b>.26</b>	<b>.44</b>	<b>.67</b>	<b>.43</b>
13-24	<b>.24</b>	.13	.15	<b>.26</b>	<b>.28</b>	<b>.25</b>	<i>-.02</i>	.09	.08	.11	.13	<b>.48</b>	<b>.18</b>
25-36	.18	.11	.12	<b>.29</b>	.17	.08	<i>-.07</i>	.07	.05	<i>-.02</i>	.14	<b>.42</b>	<b>.13</b>
37-48	.20	<i>-.08</i>	<i>-.02</i>	.11	.13	.11	.06	.17	.03	.04	.07	<b>.48</b>	.11
49-60	.11	<i>-.04</i>	<i>-.09</i>	.10	.10	.09	<i>-.07</i>	.02	<i>-.21</i>	.07	.01	.24	.03
Column Mean	<b>.32</b>	.11	.10	<b>.24</b>	<b>.21</b>	<b>.19</b>	.05	<b>.14</b>	.04	.09	<b>.16</b>	<b>.46</b>	<b>.17</b>

**Panel B**

		1990-01 - 2015-12											
Column →	1	2	3	4	5	6	7	8	9	10	11	12	Row Mean
Lags ↓													
1-12	<b>.85</b>	<b>.59</b>	<b>.41</b>	<b>.47</b>	<b>.42</b>	<b>.43</b>	<b>.49</b>	<b>.47</b>	<b>.38</b>	<b>.33</b>	<b>.39</b>	<b>.58</b>	<b>.48</b>
13-24	<b>.27</b>	<b>.25</b>	.25	<b>.32</b>	<b>.34</b>	<b>.36</b>	.22	.11	.17	.18	.22	<b>.48</b>	<b>.26</b>
25-36	.20	.17	.23	<b>.31</b>	.19	.16	.00	.06	.09	.08	.18	<b>.44</b>	<b>.18</b>
37-48	.09	<i>-.05</i>	.11	.19	<b>.27</b>	<b>.28</b>	.22	<b>.25</b>	.24	.15	.18	<b>.45</b>	<b>.20</b>
49-60	.17	.13	.11	<b>.34</b>	<b>.27</b>	.16	.04	.18	<i>-.00</i>	<i>-.02</i>	<i>-.01</i>	.20	<b>.13</b>
Column Mean	<b>.32</b>	<b>.22</b>	<b>.22</b>	<b>.33</b>	<b>.30</b>	<b>.28</b>	<b>.20</b>	<b>.22</b>	<b>.17</b>	.14	<b>.19</b>	<b>.43</b>	<b>.25</b>

**Table 6 Bias Estimates by Lag**

Values multiplied by 100. Bolded figures are 1% significant based on a two-sided test. Negative values are italicized. Summary statistics for the data used in Panels A and B are presented in Panel A of Table 5.

(monthly) lags. For instance, the bias of the autocorrelation parameter estimates for lags 1 through 12, Panel A, are individually and jointly statistically significant, with a mean bias of 0.0043 (values in Table 6 are multiplied by 100). The row averages also inform us on the size and significance of the upward bias as the lag length increases - the decline in the bias is close to monotonic, dropping from 0.0043 in the first row to a negligible and statistically insignificant value of 0.0003 in the fifth row of Panel A. Both the size and the significance of the bias declines with the length of the lag. For lags 49 through 60, Panel A, the mean bias is 0.0003. This suggests that distant returns are less helpful to predict future returns than recent returns and that a simple average of past returns is unlikely to be a useful predictor of future returns.

The column averages inform us on mod-12 effects. For instance, from column 1, lags 1,

13, 25, 37, and 49 all show evidence of a large upward bias in the autocorrelation estimates, though only the bias on the autocorrelation coefficient for lags 1 and 13 are statistically significant. This pattern of bias is strongest for the well-documented mod-12 group of annual lags (12, 24, 36, 48 and 60) displayed in column 12 of the table but is also prominent for the group of lags in column 4 (lags 4, 16, 28, 40, and 52), column 5 (lags 5, 17, 29, 41, and 53), and column 6 (lags 6, 18, 30, 42 and 54). Although the lag 12 mod-12 group looks less special when compared to other mod-12 pulses on the basis of biased estimates, it is the group of lags for which the bias is largest and most frequently statistically significant, as well as the group of lags that present the largest absolute parameter estimates.

The most recent period, post-publication of Jegadeesh (1990), 1990-2015, shows overall stronger evidence of biased estimates from using Equation 1, presented in Panel B. Many lags that indicate reversals in the 36 to 60 month window are statistically significant on the more recent 1990-2015 sample and very few estimates are biased downwards (none of these downward biases are statistically significant). The biases implied by comparing autocorrelation estimates from the classic model of Equation 1 versus 7 are largest and most significant for lags that are positive, the momentum effects. There is also no evidence of a post-publication waning of this bias, consistent with this bias issuing from a methodological error.

The long run return regressions of Lee and Swaminathan (2000), looking at annual returns regressed on lagged annual returns, may be more subject to bias than the calibrations I have performed based on monthly returns because the cross-sectional mean return dispersion could be expected to be higher with annual data while the firm-specific return total volatility would be lower, both of which would increase the bias I have identified.

This bias also extends to the panel/time-series estimation of Moskowitz et al. (2012), who focus on time-series momentum across many asset classes. Their Equation (2), p.233, normalizes returns by asset volatility, to account for variation in volatility across asset classes, but is otherwise identical in form to Equation 1. It is the restriction of the intercept to be equivalent over asset classes, even though they take advantage of a full panel/time-series estimation, that exposes them to the same bias. Similar to Jegadeesh (1990) and others, they find strong return continuation (positive autocorrelation coefficients) up to a 12 month lag, then reversals (though they find no mod-12 positive pulse autocorrelation lags with their monthly data). The implication of the bias for their regressions is that the return

continuation will be smaller and less statistically significant, and their evidence of long term reversals may understate the reversal effect.

One central takeaway from these results is that evidence for the bias from the use of Equation 1 is strong and stable over time. Another is that the evidence of positive autocorrelation pulses at mod-12 lags is much more tenuous than reported in the literature, as indeed is the support for time-series momentum for lags 12 and below. An accurate statement reflecting the corrected estimates from a fixed-effects model such as Equation 7 is that strong and statistically significant reversals are interrupted at mod-12 lags and there is little if any evidence for momentum at any lag length. That is, the evidence favors the DeBondt and Thaler (1985) reversal pattern.<sup>23</sup> Point estimates of autocorrelation produced by Equation 7, plotted in Figures 1 to 8, undermine the case for simple time-series momentum effects and support findings in the literature for reversals. These results also support, though indirectly, work arguing that momentum portfolio profits are attributable to cross-sectional differences in expected returns, not time-series dependence in firm-specific returns. See Lehmann (1990), Lo and MacKinlay (1990), Conrad and Kaul (1998), Moskowitz and Grinblatt (1999) Jegadeesh and Titman (2002), and Bogouslavsky (2016) for further discussion and debate on this issue.

## V A New Test of Asset Pricing Models

The results presented here, and my focus on biased estimates, point to a simple asset pricing test. The estimate of  $\rho_k$  from a fixed-effects model like Equation 7 is consistent but inefficient under the null of no cross-sectional dispersion in mean returns (so long as standard regularity conditions also hold) and consistent under the alternative of cross-sectional dispersion (so long as the dispersion is captured by the fixed-effects).<sup>24</sup> The estimate of  $\rho_k$  in the classic model of Equation 1 is consistent and efficient under the null of no cross-sectional dispersion in mean returns but inconsistent under the alternative of cross-sectional dispersion in mean returns (again, given standard regularity conditions). The Hausman (1978) test, applicable to two jointly asymptotically normally distributed estimators such as OLS produces when

---

<sup>23</sup>DeBondt and Thaler (1985) do not make use of the methodology I am critiquing, instead relying on cumulative abnormal returns of portfolios of stocks.

<sup>24</sup>I conjecture consistency will follow more generally if the number of fixed effects terms increases with  $n$ , depending on the nature of the cross-sectional dispersion in returns.

applied to Equations 1 (the classic model) and 7 (the fixed-effects model), allows a simple test of the null of a correctly specified model. Should the estimates of  $\rho_k$  from Equations 1 and 7 differ significantly we can reject the null of no cross-sectional dispersion in mean returns. An alternative to testing for difference in the estimates of  $\rho_k$  is a test of significant difference of the estimated intercepts,  $\rho_0$ , from the classic model of Equation 1 and the fixed-effects model of Equation 7.

If  $b_1$  (from the classic model, Equation 1) is the efficient estimate (but inconsistent under the alternative) and  $b_0$  the consistent (under both the null and alternative) but inefficient estimate (from Equation 7), then the Hausman test (H) can be written as

$$H = (b_1 - b_0)' (\text{Var}(b_0) - \text{Var}(b_1))^\dagger (b_1 - b_0) \sim \chi_p^2$$

where  $\dagger$  denotes the Moore-Penrose (generalized) inverse and  $p$  is the number of parameters,  $p = 1$  in the cases I have considered. For the  $p = 1$  case this is merely a t-test squared. The application to Fama-MacBeth regressions is straightforward, with the sample of  $\rho_{k,t}$  over  $t$ , estimated with the classic model (efficient under the null) and the fixed-effects model (consistent but inefficient under the null) permitting the estimation of  $\text{Var}(b_i)$ ,  $i = 1, 2$ . It is also possible in the context of the Fama-MacBeth method to directly estimate the variance of  $b_1 - b_0$  from the samples of ( $b_1$  and  $b_0$ ) produced by the Fama-MacBeth regressions, which is the approach I shall use here.

Of course, one could simply perform a test of the joint significance of the fixed-effects dummy variable coefficients. Fama and MacBeth (1973) perform a test similar in spirit, for linearity of expected returns in beta. An advantage of the Hausman test is that it provides a simple test of significant difference in only the parameter of interest,  $\rho_k$  (or alternatively, the intercept  $\rho_0$ ), something that cannot be tested with traditional methodologies. Hausman and Taylor (1980) show that the Hausman test is the uniformly most powerful test among invariant tests of the correct specification hypothesis, and that the Hausman test dominates the F test of the difference between the efficient and inefficient model specifications (in the situation considered here, the F-test of the fixed-effects dummy variables). See also Holly (1986). I have explored simulations to evaluate the power and size of the Hausman test in the context of Fama-MacBeth regressions, and the results (untabulated) demonstrate a very slight tendency to over-reject under the null but very good power.<sup>25</sup>

---

<sup>25</sup>At the 5% level with 25,000 replications, using the calibration of Table 1 column 1 (the null of no cross-

## A Application to the Classic Momentum Model

The Hausman test can be based on the intercept value, alpha, or the coefficient on the lagged return, although a test on the intercept, alpha, is perhaps more natural. Using deciles based on market equity categories and testing for significant difference compared to the classic model's intercept estimate (no cross-sectional dispersion in mean firm returns) with  $k = 1$  yields a t-test of 10.5.<sup>26</sup> Basing this test on the intercept and a five-by-five market equity crossed with book-to-market categories yields a t-test of 7.4. This same test, but performed on  $\rho_1$  rather than the intercept term from the five-by-five market equity crossed with book-to-market categories, yields a t-test of 12.6. This test using  $\rho_k$  rather than  $\rho_1$  gives varying results across lags  $k$  (the strongest rejections are for mod-12 values) and from different fixed-effects models. Basing the Hausman test on deciles of cashflow to price and the difference in the estimate of alpha yields a t-test of 6.1. Unsurprisingly, perhaps, the null of no cross-sectional dispersion in mean firm returns is strongly rejected.

## B Extension to Firm Characteristics Regressions: Testing the Linear Factor Structure Assumption

Common in the empirical asset pricing literature are similar models to Equation 1, substituting some firm characteristic for the lagged return term. A small collection of recent papers that employ Fama-MacBeth estimation using models with a single factor for at least some of the results reported, include Barberis et al. (2016), Bebenko et al. (2016), and Bollerslev et al. (2016).<sup>27</sup> Many more papers use Fama-MacBeth regression methods with controls like size, book-to-market, and other firm characteristics. A few recent examples include Akbas (2016) and Ge et al. (2016). The typical form of such models is similar or identical to that seen in Fama and French (1992), Table III. In Fama and French (1992) we see the results of Fama-MacBeth regressions like this:

---

sectional dispersion in mean firm returns), the Hausman test rejects 5.324% of the time. The simulation standard error is 0.1398 so that this is also beyond two standard errors from the expected value of 5%, hence statistically significant at conventional levels. Using the calibration of Table 1 columns 2 and 3 which incorporate cross-sectional dispersion in mean firm returns, the test rejects the null of no dispersion strongly. There are no failures to reject at the 5% level recorded in 25,000 replications.

<sup>26</sup>The intercept estimate from the implementation of the fixed-effects model was the average intercept across deciles.

<sup>27</sup>See Section 2.4, p. 3089, Table 5, column 1 of Barberis et al. (2016), Table 3, p. 443 of Bebenko et al. (2016), and Panel A of Table 8, p.481 of Bollerslev et al. (2016).

$$r_{t,i} = \alpha_t + \beta_{Z,t}Z_{t,i} + u_{t,i} \quad (9)$$

where  $t = 1, \dots, T$  is typically measured in months,  $i = 1, \dots, N$  indicates the cross-section of firms,  $r_{t,i}$  is the return for firm  $i$  at time  $t$ ,  $u_{t,i}$  is the firm  $i$  idiosyncratic error term at time  $t$ ,  $\alpha_t$  is the intercept, and  $\beta_{Z,t}$  the coefficient for the firm-specific variable  $Z$ . Both  $\alpha_t$  and  $\beta_{Z,t}$  are imposed to be identical for all firms in the cross-section and estimated independently for each time period  $t$ . Fama-MacBeth estimates of  $\alpha$  and  $\beta_Z$  are formed by averaging  $\alpha_t$  and  $\beta_{Z,t}$  over  $t$ , and OLS is typically used to produce estimates.

The way such regressions are motivated, with the intercept constrained to be the same for all stocks, is that such a model imposes a linear factor structure on expected returns, in keeping with Merton (1973) and Ross (1976) and the multifactor asset pricing model. See, for instance, Fama and French (1992). If this linear factor structure is correct, we have no bias from estimating Equation 9 with the Fama-MacBeth methodology.

I am interested in suggesting a new specification test of these linear factor models. A simple test of the linear factor structure is the Hausman test of model specification, based on the difference in the estimate of  $\alpha_t$  from estimating the regression Equation 9 and one augmenting this model with fixed effects, identical to that described for the classic momentum model. This test can have greater power than simply testing for a non-0 intercept in Equation 9 as the two models, Equation 9, and Equation 9 augmented with fixed effects, can have differently-signed intercepts.<sup>28</sup>

The regression results reported in Table III of Fama and French (1992) include a regression of individual firm returns on firm book-to-market ratios ( $\ln(\text{BE}/\text{ME})$ ). Making use of a fixed-effects model using deciles based on book-to-market categories, applied to the Fama and French (1992) regression of Equation 9 with the book-market ratio as the  $Z$  variable, yields a Hausman test t-value of 11.4, rejecting the null hypothesis that the two models produce identical estimates of alpha and hence indicating model misspecification. The Hausman test of equivalence of alpha estimates, from Equation 9 with the book-market ratio as the  $Z$  variable and from the fixed-effects model based on market equity deciles, leads to a t-value of 13.6, again strongly rejecting the null hypothesis that Equation 9 is correctly specified. Basing the Hausman test on cashflow deciles leads to a test value of 10.2, and using earnings-to-price ratios yields a test value of 13.8. These represent very strong rejections of a linear

---

<sup>28</sup>In practice in the applications I have investigated, this occurs frequently.

factor model such as Equation 9, using the book-market ratio as the  $Z$  variable.

Similar to the estimates of the coefficient on the lagged return in the momentum regressions, the coefficient on the book-to-market ratio also appears to be strongly biased. Untabulated results show that the coefficient on  $\ln(\text{BE}/\text{ME})$  is more likely between -0.3 and 0.3, not 0.5 as Fama and French (1992) report, and it seems that the book-to-market equity coefficient is statistically insignificant. Fama and French (1992), p.441 write:

*The FM regressions in Table III confirm the importance of book-to-market equity in explaining the cross section of average stock returns. The average slope from the monthly regressions of returns on  $\ln(\text{BE}/\text{ME})$  alone is 0.50%, with a  $t$ -statistic of 5.71.*

I am not asserting that the Hausman test results undermine the whole of Fama and French (1992). I am suggesting that at least some of the regression results they present in their Table III are not in and of themselves indicative of the impact of variables like book-to-market have on returns. Put another way, an assumption that the Equation 9 regression can be thought of as a linear factor model is poorly supported by the data, and the utility of book-to-market even as a proxy for an underlying true risk factor cannot be determined by estimating a (misspecified) model like Equation 9.

Finally, while it is not surprising that a model like that of Equation 9 is misspecified, this is simply demonstration of the Hausman test. An attractive feature of the Hausman test is that it is easily implemented for richer asset pricing models.

## VI Conclusions

A large literature explores time-series momentum making use of the Fama-MacBeth methodology. The original work by Jegadeesh (1990) implemented this method both correctly (Equation 2 of Jegadeesh 1990) and incorrectly (see footnote 14 of Jegadeesh 1990)<sup>29</sup> when estimating the autocorrelation coefficients, and unfortunately much of the work which followed adopted the version with an incorrect model structure. This oversight has misled

---

<sup>29</sup>Jegadeesh (1990) first presents a model with demeaned dependent variable returns, using this model to identify time-series momentum. This model is not subject to the criticism I outline in this paper, though the simulation evidence I have suggests that this model will not perform well at controlling for cross-sectional dispersion of firm returns. A second model that Jegadeesh (1990) uses is subject to the bias my paper identifies, however this model is used by Jegadeesh only to form forecasts of returns with which to rank stocks. The results I have do not speak to any impact on that ranking. That is to say, my results do not directly challenge or overturn Jegadeesh (1990).

researchers to believe that time-series autocorrelation is a potential driver of long-short momentum portfolio profits, because this misspecified model upward biases autocorrelation estimates and produces false evidence of positive autocorrelation, in particular at lags 2 through 12. In fact, autocorrelation patterns in firm-specific returns *is* apparent in the data, but this is largely due to *reversals* in returns. This reversal pattern was obscured by the misspecification I identify here, again because of the upward bias in parameter estimates.

The bias I document is large and strongly statistically significant, is stable over time, and does not moderate in recent data. I believe that concern over this bias is pertinent to analysis outside of the confines of time-series momentum and reversals, extending to the Fama-MacBeth methodology in application to the linear factor model. A great many papers use the Fama-MacBeth methodology when estimating the linear factor models of Merton (1973) and Ross (1976). Although it is debatable, I believe that these models are problematic to provide even summary descriptions of return predictability, because if these models are misspecified it is not just that there is mispricing, that alpha is non-zero. It is also the case that the coefficient on the factor will be strongly biased for the same reason the autocorrelation parameter is biased in the classic momentum model. Certainly it is easy to reject some simple linear factor models with a new asset pricing model test I propose here, a Hausman test.

Finally, in each of the many papers I have identified, the incorrect or at least contentious application of the Fama-MacBeth method calls into question one of the legs supporting these papers' conclusions, and results mistakenly showing strong autocorrelation in individual firm returns have even motivated entire papers, including Hong and Stein (1999) and Bogousslavsky (2016).<sup>30</sup>

---

<sup>30</sup>I am not suggesting that the results of Hong and Stein (1999) or Bogousslavsky (2016) are themselves incorrect, just that the motivation for these papers is suspect.

## References

- Akbas, F., 2016. "The calm before the storm" *The Journal of Finance* 71 (1), 225-265.
- Asparouhova, E., Bessembinder, H. and Kalcheva, I., 2010. "Liquidity biases in asset pricing tests" *Journal of Financial Economics* 96, 215-237.
- Barberis, N., Mukherjee, A., and Wang, B., 2016. "Prospect theory and stock returns: An empirical test" *Review of Financial Studies* 29 (11), 3068-3107.
- Bebenko, I., Boguth, O., and Tserlukevich, Y., 2016. "Idiosyncratic cash flows and systematic risk" *Journal of Finance* 71 (1), 425-455.
- Berk, J., 2000. "Sorting out sorts" *The Journal of Finance* 55, 407-427.
- Bogousslavsky, V., 2016. "Infrequent rebalancing, return autocorrelation, and seasonality" *The Journal of Finance* 71 (6), 2967-3006.
- Bollerslev, T., Li, S., and Todorov, V., 2016. "Roughing up beta: Continuous versus discontinuous betas and the cross-section of expected stock returns" *Journal of Financial Economics* 120, 464-490.
- Brennan, M., Chordia, T., and Subrahmanyam, A. 1998. "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns" *Journal of Financial Economics* 49, 345-373.
- Campbell, J., 1991. "A variance decomposition for stock returns" *The Economic Journal* 101 (405), 157-179.
- Campbell, J., 2001. "Why long horizons? A study of power against persistent alternatives" *Journal of Empirical Finance* 8(5), 459-491.
- Campbell, J., and Cochrane, J., 1999. "By force of habit: A consumption-based explanation of aggregate stock market behavior" *Journal of Political Economy* 107, 205-251.
- Conrad, J., and Kaul, G., 1998. "An anatomy of trading strategies" *Review of Financial Studies* 11, 489-519.
- DeBondt, W., and Thaler, R. 1985. "Does the stock market overreact?" *Journal of Finance* 40, 793-805.
- D'Souza, I., Srichanachaichok, V., Wang, G., Jiaguo and Yao, Y., 2016. "The Enduring Effect of Time-Series Momentum on Stock Returns Over Nearly 100-Years" SSRN 2720600 Working Paper.
- Fama, E., 1970. "Efficient capital markets: A review of theory and empirical Work" *Journal of Finance* 25, 383-417.
- Fama, E., 1991. "Efficient capital markets: II" *Journal of Finance* 46, 1575-1617.
- Fama, E., and French, K., 1992. The cross-section of expected stock returns. *Journal of Finance*, 47(2), 427-465.
- Ferson, W., Sarkissian, S., and Simin, T., 2003. "Is stock return predictability spurious?" *Journal of Investment Management* 1, 1-10.

- Ferson, W., Sarkissian, S., and Simin, T., 2008. "Asset pricing models with conditional betas and alphas: The effects of data snooping and spurious regression" *Journal of Financial and Quantitative Analysis* 43, 331-353.
- French, K., and Roll, R., 1986. "Stock return variances: The arrival of information and the reaction of traders" *Journal of Financial Economics* 17, 5-26.
- Gagliardini, P., Ossola, E., and Scaillet, O., 2016. "Time-varying risk premium in large cross-sectional equity data sets" *Econometrica* 84(3) 985-1046.
- Ge, L., Lin, T. and Pearson, N. 2016. "Why does the option to stock volume ratio predict stock returns?" *Journal of Financial Economics* 120, 601-622.
- Gong, Q., Liu, M., and Liu, Q., 2015. "Momentum is really short-term momentum" *Journal of Banking and Finance* 50, 169-182.
- Gospodinov, N., Kan, R. and Robotti, C., 2014. "Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors" *Review of Financial Studies* 27, 2139-2170.
- Hansen, L., 1982. "Large sample properties of generalized method of moments estimators" *Econometrica* 50, 1029-1054.
- Harvey, C., Liu, Y., and Zhu, H., 2016. "... and the cross-section of expected returns" *Review of Financial Studies* 26, 5-68.
- Hausman, J., 1978. "Specification tests in econometrics" *Econometrica* 46, 1251-1271.
- Hausman, J., and Taylor, W., 1980. "Comparing specification tests and classical tests" M.I.T. Department of Economics Working Paper No. 266, 1980.
- Heston, S. and Sadka, R. 2008. "Seasonality in the cross-section of stock returns" *Journal of Financial Economics* 87(2), 418-445.
- Heston, S., and Sadka, R., 2010. "Seasonality in the cross-section of stock returns: the international evidence" *Journal of Financial and Quantitative Analysis* 45 (5), 1133-1160.
- Heston, S. L., Korajczyk, R. A., Sadka, R., and Thorson, L. D., 2011. "Are you trading predictably?" *Financial Analysts Journal* 67(2), 36-44.
- Hirshleifer, D., Jiang, D., and Meng, Y., 2017. "Mood beta and seasonalities in stock returns" SSRN-id2880257.
- Hjalmarsson, E., 2010. "Predicting global stock returns" *Journal of Financial and Quantitative Analysis* 45, 49-80.
- Holly, A., 1986. "Some useful equivalence properties of Hausman's test" *Economics Letters* 20, 39-43.
- Hong, H., and Stein, J., 1999. "A unified theory of underreaction, momentum trading, and overreaction in asset markets" *The Journal of finance* 54(6), 2143-2184.
- Jegadeesh, N., 1990. "Evidence of predictable behavior of security returns" *Journal of Finance* 45, 881-898.

- Jegadeesh, N., Titman, S., 1993. "Returns to buying winners and selling losers: Implications for stock market efficiency" *Journal of Finance* 48, 65-91.
- Jegadeesh, N., Titman, S., 2002. "Cross-sectional and time-series determinants of momentum returns" *Review of Financial Studies* 15(1), 143-157.
- Kan, R., Robotti, C., and Shanken, J., 2013. "Pricing model performance and the two-pass cross-sectional regression methodology" *Journal of Finance* 68, 2617-2649.
- Keloharju, M., Linnainmaa J., and Nyberg, P. 2016a. "Return seasonalities" *Journal of Finance* 71 (4), 1557-1589.
- Keloharju, M., Linnainmaa J., and Nyberg, P. 2016b. "Seasonal reversals in expected stock returns" Working paper.
- Kendall, M., 1954. "Note on bias in the estimation of autocorrelation" *Biometrika* 41 (3-4), 403-404.
- Korteweg, A. and Sorensen, M., 2016. "Skill and luck in private equity performance" *Journal of Financial Economics*, forthcoming.
- Lee, C. and Swaminathan, B., 2000. "Price momentum and trading volume" *The Journal of Finance* 55, 2017-2069.
- Lehmann, B., 1990. "Fads, martingales and market efficiency" *Quarterly Journal of Economics* 105, 1-28.
- Lewellen, J., Nagel, S., and Shanken, J., 2010. "A skeptical appraisal of asset pricing tests" *Journal of Financial Economics* 96, 175-194.
- Lo, A., and MacKinlay, A., 1990. "When are contrarian profits due to stock market overreaction?" *Review of Financial Studies* 3(2), 175-205.
- Lyon, J., Barber, B., and Tsai, C., 1999. "Improved methods for tests of long-run abnormal stock returns" *Journal of Finance* 54, 165-201.
- Merton, R., 1973. "An intertemporal capital asset pricing model" *Econometrica*, 41, 867-887.
- Moskowitz, T., and Grinblatt, M., 1999. "Do industries explain momentum?" *The Journal of Finance* 54(4), 1249-1290.
- Moskowitz, T., Ooi, Y., and Pedersen, L. H., 2012. "Time series momentum" *Journal of Financial Economics* 104, 228-250.
- Newey, W. and West, K. 1994. "Automatic lag selection in covariance matrix estimation" *Review of Economic Studies* 61, 631-653.
- Novy-Marx, R., 2012. "Is momentum really momentum?" *Journal of Financial Economics* 103(3), 429-453.
- Pastor, L. and Stambaugh, R. F., 2009. "Predictive systems: Living with imperfect predictors" *The Journal of Finance* 64, 1583-1628.
- Pastor, L., Stambaugh, R., and Taylor, L., 2015. "Scale and skill in active management" *Journal of Financial Economics* 116, 23-45.

- Petersen, M., 2009. “Estimating standard errors in finance panel data sets: Comparing approaches” *Review of financial studies* 22 (1), 435-480.
- Robins, R. and Smith, G., 2016. “No more weekend effect” *Critical Finance Review* 5(2), 417-424.
- Ross, S., 1976. “The arbitrage theory of capital asset pricing” *Journal of Economic Theory* 13, 341-360.
- Shumway, T., 1997. “The delisting bias in CRSP data” *The Journal of Finance* 52(1), 327-340.
- Valkanov, R., 2003. “Long-horizon regressions: theoretical results and applications” *Journal of Financial Economics* 8(2), 201-232.

## Appendix A

### A Remark on Heston and Sadka (2008) and Feasible Portfolio Strategies as They Relate to Cross-sectional Regression Return Responses

Heston and Sadka (2008) motivate the OLS coefficient estimate  $\gamma_{k,t}$  (return responses, adopting the terminology of Heston and Sadka, 2008) from the regression of cross-sectional firm returns on lagged own-firm returns as a weighted long-short portfolio of firm returns.

$$r_{t,i} = \alpha_{k,t} + \gamma_{k,t} r_{t-k,i} + e_{t,i} \tag{10}$$

$$\hat{\gamma}_{k,t} = \sum_{i=1}^N w_{i,k,t} r_{t,i}$$

where  $r_{i,t}$  is firm  $i$ 's return at time  $t$  and we observe  $N$  firms in the cross section. The notion that the coefficient  $\hat{\gamma}_{k,t}$  can be interpreted as a long-short portfolio is made clear by writing out the weights as

$$w_{i,k,t} = \frac{r_{i,t-k} - \left( \sum_{i=1}^N r_{i,t-k}/N \right)}{\sum_{i=1}^N \left( r_{i,t-k} - \left( \sum_{i=1}^N r_{i,t-k}/N \right) \right) \left( r_{i,t-k} - \left( \sum_{i=1}^N r_{i,t-k}/N \right) \right)}$$

and noting that the numerator sums to zero over  $i$  (a natural long-short construction). Second, note that the denominator, common to all terms, can be omitted when considering the relative long-short weights, because the weight on an asset  $i$  in the long-short portfolio is proportional to

$$p_{i,t} = \left( r_{i,t-k} - \left( \sum_{i=1}^N r_{i,t-k}/N \right) \right).$$

Heston and Sadka (2008) remark: “The resulting return responses are highly correlated with results from other methodologies, including decile spreads (Jegadeesh and Titman, 1993) and weighted relative strength portfolios (Lehmann, 1990; Lo and MacKinlay, 1990).”

This is immediately apparent by observing that firms with returns  $r_{i,t-k}$  above the average return  $k$  periods ago will have positive weights, and firms below the average will have negative

weights, a simple momentum strategy. It is also immediately apparent that this long-short portfolio momentum strategy observation of Heston and Sadka (2008) is unaffected by bias in the estimate of  $\hat{\gamma}_{k,t}$ , because the bias does not change relative weightings on assets.

Heston and Sadka (2008) also remark: “We call these slope coefficient estimates return responses because they show the cross-sectional response of returns at one date to returns at a previous date”

and they use these weights to construct feasible portfolio strategies, motivating the strategies based on these weights by observing that the return responses  $\hat{\gamma}_{k,t}$  are statistically significant at long-horizon annual lag frequencies ( $k=12, 24, 36, \dots$ ). Given, however, that the portfolio strategy weights are proportional only to lagged returns, and are not changed by current returns, this motivation of the portfolio strategies with reference to  $\hat{\gamma}_{k,t}$  is incorrect.

This observation does not take away from the empirical regularities that Heston and Sadka (2008) uncover, nor does it impinge on the bulk of the evidence they provide. This observation merely highlights that leaning on the misspecified regression in Equation 10 can easily misdirect us from what is interesting. In the particular case of Heston and Sadka (2008), what is interesting is the information on cross-sectional return variation found at annual lags for current portfolio allocation purposes. The evidence provided by the pattern of  $\hat{\gamma}_{k,t}$  over  $k$  is uninformative for portfolio allocation as it does not characterize the *response* of returns at one date to returns at a previous date (it is a biased estimate of that response). Rather,  $\hat{\gamma}_{k,t}$  can be exploited (as indeed Heston and Sadka, 2008 do) for insights on the stability of the cross-sectional dispersion of returns. The pattern of  $\hat{\gamma}_{k,t}$  tells us something about the cross-sectional stability of the dispersion of returns; that this stability is interrupted following a seasonal pattern, and return dispersion patterns annually reassert themselves after an eleven month interruption. Past return dispersion tells us about coming return dispersion not because of autocorrelation, or as Heston and Sadka (2008) put it, return response at one date to a previous date. Firm returns are not *responding* to past returns, which implies a dynamic relationship. Rather, the evidence provided by Heston and Sadka (2008) informs us that returns display time-varying but seasonally stable cross-sectional expected values.