

# Asset Pricing Models with Conditional Betas and Alphas: The Effects of Data Snooping and Spurious Regression

Wayne E. Ferson, Sergei Sarkissian, and Timothy Simin\*

## Abstract

This paper studies the estimation of asset pricing model regressions with conditional alphas and betas, focusing on the joint effects of data snooping and spurious regression. We find that the regressions are reasonably well specified for conditional betas, even in settings where simple predictive regressions are severely biased. However, there are biases in estimates of the conditional alphas. When time-varying alphas are suppressed and only time-varying betas are considered, the betas become biased. Previous studies overstate the significance of time-varying alphas.

## I. Introduction

Regression models for stock or portfolio returns on market-wide factors have long been a staple of financial economics. Such factor models are used in event studies (e.g., Fama, Fisher, Jensen, and Roll (1969)), in tests of asset pricing theories such as the Capital Asset Pricing Model (CAPM) (Sharpe (1964)), and in other applications. For example, when the market return  $r_m$  is the factor, the regression model for the return  $r_{t+1}$  is:

$$(1) \quad r_{t+1} = \alpha + \beta r_{m,t+1} + u_{t+1},$$

where  $E(u_{t+1}) = E(u_{t+1}r_{m,t+1}) = 0$ . The slope coefficients are the “betas,” which measure the market factor risk. When the returns are measured in excess of a reference asset like a risk-free Treasury bill return, the intercepts are the “alphas,” which measure the expected abnormal return. For example, when  $r_m$  is the market

---

\*Ferson, ferson@marshall.usc.edu, Marshall School of Business, University of Southern California, 3670 Trousdale Parkway, Los Angeles, CA 90089; Sarkissian, sergei.sarkissian@mcgill.ca, Faculty of Management, McGill University, 1001 Sherbrooke St. West, Montreal, Quebec, H3A 1G5, Canada; and Simin, tsimin@psu.edu, Smeal College of Business, Pennsylvania State University, 345 Business Building, University Park, PA 16802. We are grateful to participants at the Federal Reserve Bank of Atlanta Conference on Topics in Financial Econometrics, McGill University, Stephen Brown (the editor), and Thomas M. Smith (the referee) for helpful comments. We are also grateful to Jeff Nucciarone at the High-Performance Computing Center at Penn. State. Sarkissian acknowledges IFM2 for financial support.

portfolio excess return, the CAPM implies that  $\alpha = 0$ , and the model is evaluated by testing that null hypothesis.

Recent work in conditional asset pricing allows for time-varying betas modeled as linear functions of lagged predictor variables following Maddala (1977). Prominent examples include Shanken (1990), Cochrane (1996), Ferson and Schadt (1996), Jagannathan and Wang (1996), and Lettau and Ludvigson (2001). The time-varying beta coefficient is  $\beta_t = b_0 + b_1 Z_t$ , where  $Z_t$  is a lagged predictor variable. In some cases, the intercept or conditional alpha is also time varying as  $\alpha_t = \alpha_0 + \alpha_1 Z_t$  (e.g., Christopherson, Ferson, and Glassman (1998)). This results in the following regression model:

$$(2) \quad r_{t+1} = \alpha_0 + \alpha_1 Z_t + b_0 r_{m,t+1} + b_1 r_{m,t+1} Z_t + u_{t+1},$$

where  $E(u_{t+1}) = E(u_{t+1} [Z_t r_{m,t+1}]) = 0$ . The conditional CAPM implies that  $\alpha_0 = 0$  and  $\alpha_1 = 0$ .

This paper studies the sampling properties of regressions like equation (2). We focus on the joint effects of two important issues. The first is naïve data mining and the second is spurious regression. Naïve data mining in this context refers to the practice of searching the data for predictor variables, perhaps collectively by a group of researchers through a series of studies, and then effectively using the same data to evaluate the model without accounting for the number of searches. Spurious regression arises when the persistence or high autocorrelation of a predictor variable tricks the standard test statistics into finding a “significant” relation where none exists. These two issues have been studied separately in many previous papers. However, in spite of the large number of studies that rely on regressions like (2), none of the work of which we are aware addresses the effects of data mining and spurious regression on these regressions. That is the goal of this paper.

Ferson, Sarkissian, and Simin (2003) study the combined effects of data mining and spurious regression in the context of simpler models where the only variable on the right-hand side is the lagged predictor:

$$(3) \quad r_{t+1} = \delta_0 + \delta_1 Z_t + v_{t+1}.$$

They focus on the slope coefficient and find that the effects of data mining and spurious regression interact and reinforce each other. The more persistent variables that generate spurious regressions are more likely to be discovered by data mining, so the spurious regression problem is worse in the presence of data mining. At the same time, standard corrections for data mining are inadequate in the presence of persistent lagged variables, so persistence magnifies the impact of data mining.

These results have profound potential implications for asset pricing regressions like (2) because the conditional asset pricing literature using regression (2) has, for the most part, used variables that were discovered based on predictive regressions like (3). This motivates our study of how data mining and spurious regression biases influence the asset pricing regressions. We make the assumption—which characterizes the existing literature well—that the lagged variables are mined on the basis of their performance in the predictive regression (3).

The lagged instruments may be mined to predict the market factor or to predict the test asset returns; we study both cases. We assume that the market factor comes from theory and is not mined from data.<sup>1</sup>

Our results indicate that the conditional betas estimated from the conditional asset pricing regressions are relatively robust to the effects of spurious regression and data mining for lagged predictors. The estimates of the average of the conditional alphas over time also are reasonably well specified. However, the estimates of time variation in alpha are subject to biases. If the model is estimated without the time-varying alpha term, the conditional betas are mildly biased. Our results help justify the use of regressions like (2) in asset pricing studies, and suggest some refinements and caveats in their interpretation.

The rest of the paper is organized as follows. Section II briefly reviews the effects of data mining and spurious regression in the simple predictive regressions. Section III further motivates our analysis of the conditional asset pricing regressions. Section IV describes our simulation design and Sections V and VI present the results. Section VII offers concluding remarks.

## II. Spurious Regression and Data Mining

The lagged predictor variables identified in the asset pricing literature frequently are highly persistent. Short-term Treasury bill yields, monthly book-to-market ratios, the market dividend yield, and some yield spreads have first-order autocorrelations of 0.97 or higher in the monthly data commonly used. This suggests that spurious regression problems may arise when these variables are the predictors.

The problem of spurious regressions was perhaps first studied by Yule (1926) who observed that two trending series that are actually independent of each other are likely to appear to be related in a given sample. Granger and Newbold (1974) study spurious regressions in the levels of macroeconomic variables. Following Granger and Newbold, we interpret a spurious regression as one in which the “*t*-ratios” in the regression are likely to indicate a significant relation when the variables are really independent. The problem may come from the numerator or the denominator of the *t*-ratio: the coefficient or its standard error may be biased. Ferson, Sarkissian, and Simin (2003) show that the problem in the predictive regression (3) lies with the standard errors. When the null hypothesis that the regression slope  $\delta_1 = 0$  is true, the error term  $v_{t+1}$  of regression (3) inherits its autocorrelation from the dependent variable. If the asset return on the left-hand side consists of a persistent expected return plus noise, the return has some degree of persistence. Assuming stationarity, the slope coefficient is consistent, but standard errors that do not account for the serial dependence correctly are biased and inconsistent.

Many studies of spurious regression examine nonstationary models, where the autocorrelation of the regressor is equal to 1.0 (e.g., Phillips (1986), Campbell and Shiller (1988), or Marmol (1998)), or where the process is local to unity, such

<sup>1</sup>There may be some debate about this if the factors are taken to be those of Fama and French (1993), (1996); see Lo and MacKinlay (1990), Ferson, Sarkissian, and Simin (1999), or Conrad, Cooper, and Kaul (2003).

that the autocorrelation approaches 1.0 as the sample size,  $T$ , gets large (e.g., Phillips (1998), Valkanov (2003), or Campbell and Yogo (2006)). In such cases, the sampling distributions for simple predictive regressions like (3) can often be described by functions of integrals of Brownian motions. It may be possible to extend this kind of analysis to regressions like (2) with data mining in future work. We focus on stationary data, allowing for realistically high autocorrelations, in our analysis. This is largely for tractability, but it also reflects a view that local to unity approximations are most appropriate when overlapping, longer horizon returns are measured. Most of the asset pricing literature that motivates this study uses monthly, non-overlapping returns data.

The theoretical foundations for spurious regression bias with stationary but close to unit root regressors is provided by Phillips (1986), (1998), Marmol (1998), Tsay and Chung (2000), Granger, Hyung, and Jeon (2001), and Jansson and Moreira (2006). Spurious regression bias is also observed in models for stock returns using dummy variables as the predictors (see Powell, Shi, Smith, and Whaley (2006)), in error-correction regressions (see Berkowitz and Giorgianni (1996)) and in predictive models for the variance of stock returns (see Paye (2006)). Phillips (2001) studies problems using bootstrap methods in the presence of spurious regressions.

Besides spurious regression bias, the financial economics literature has examined other statistical issues with predictive regressions for stock returns. Boudoukh and Richardson (1994) provide an overview. Lanne (2002) and Boudoukh, Richardson, and Whitelaw (2005) study regressions to predict long-horizon returns. In this problem, the correlation between the regression slope estimators for the different horizons is crucial. Goetzmann and Jorion (1993), Nelson and Kim (1993), Bekaert, Hodrick, and Marshall (1997), and Stambaugh (1999) study biases due to dependent stochastic regressors. Other useful studies of issues that debate the predictability of stock returns with lagged information variables include Kim, Nelson, and Startz (1991), Torous, Valkanov, and Yan (2003), Lewellen (2004), Ferson, Heuson, and Su (2005), Ang and Bekaert (2007), and Chapman, Simin, and Yan (2003).

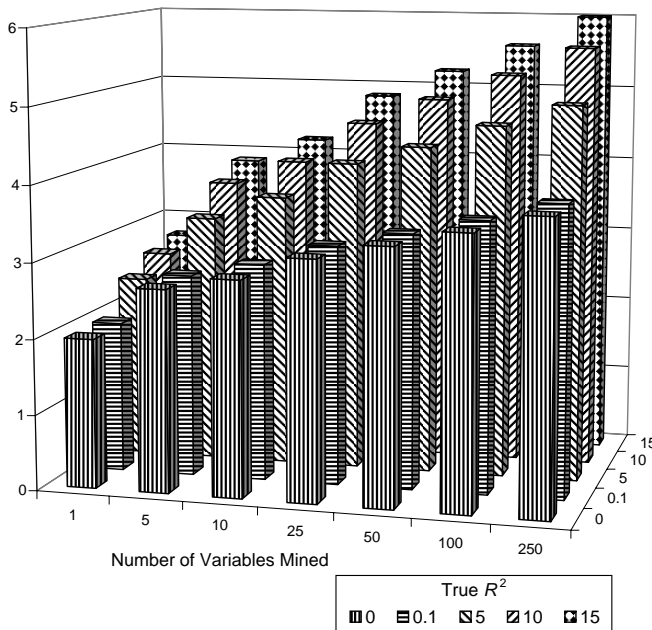
Stambaugh (1999) studies a predictive regression where the analyst observes and uses the correct regressor, but a finite sample bias arises because the future value of the regressor is correlated with the left-hand side stock return and, thus, with the regression error. Stambaugh shows that the bias is related to the finite sample bias in the autocorrelation of the predictor in a univariate setting. Amihud and Hurvich (2004) examine solutions for this bias in a multiple regression context. In our problem, the lagged variable used by the analyst is independent of the correct regressor, so the Stambaugh bias does not arise. The bias in our case is due to the joint effects of data mining and spurious regression. By assuming that the data mining searches over variables independent from the "true" predictor, we effectively assume that the Stambaugh bias cannot arise. In the real world, data mining may uncover variables that are correlated with the true predictor, so a modified version of the Stambaugh bias may arise. Under data mining, the Stambaugh bias will be modified because the relation of the mined variables to lagged returns can indirectly affect (if the miner chooses based on the biased coefficient) and be affected by data mining. The interactions between this source

of bias, spurious regression, and data mining are likely to be complex, but present an opportunity for future research.

Data mining refers to the practice of searching through the data to find predictor variables. With naïve data mining, the same data are effectively used in model estimation, prediction, or testing. Searching through 100 independent variables, one expects to find about five that are “significant” at the 5% level. Leamer (1978), Hastie, Tibshirani, and Friedman (2001), and White (2000), among others, describe statistical approaches that conduct inference, controlling for the number of searches. In financial economics, the problem is usually complicated by the fact that many researchers use the same data sets. Thus, the data are effectively mined by an unknown number of previous researchers (see, e.g., Lo and MacKinlay (1990)).

Figure 1, taken from simulations described in Ferson, Sarkissian, and Simin (2003), illustrates the interaction between data mining and spurious regression. The critical values for significant  $t$ -statistics for the  $\delta_1$  coefficient increase with the number of variables mined and with the extent of spurious regression. In the presence of spurious regression, persistent variables are likely to be mined, and the two effects reinforce each other. For example, if the expected return accounts for 10% of the stock return variance, we only have to consider mining in a set of 5 to 10 instruments to obtain critical values as high as those obtained with 50 to 100 instruments and no spurious regression. The figure illustrates that, even with a modest amount of data mining, the combined effects have a powerful impact.

FIGURE 1  
Interaction of Data Mining and Spurious Regression  
(Critical  $t$ -ratios: true persistence is 95%)



### III. Regressions with Time-Varying Alphas and Betas

The conditional asset pricing literature using regressions like (2) has evolved from the literature on pure predictive regressions. First, studies identified lagged variables that appear to predict stock returns. Later studies, beginning with Gibbons and Ferson (1985), used the same variables to study asset pricing models. Thus, it is reasonable to presume that data mining is directed at the simpler predictive regressions. The question now is: How does this affect the validity of the subsequent asset pricing research that uses these variables in regressions like (2)? That is the central question addressed by this study.

Table 1 summarizes representative studies that use regression model (2). It lists the sample period, number of observations, and the lagged instruments employed. It also indicates whether the study uses the full model (2), with both time-varying betas and alphas, or restricted versions of the model in which either the time-varying betas or time-varying alphas are suppressed. Finally, the table summarizes the largest  $t$ -statistics for the coefficients  $\alpha_1$  and  $b_1$  reported in each study. If we find that the largest  $t$ -statistics are insignificant in view of the joint effects of spurious regression and data mining, then none of the coefficients is significant. We return to this table later and revisit the evidence.

TABLE 1  
Representative Studies

Table 1 summarizes representative conditional asset pricing and performance evaluation studies. The first column indicates the published study. The second column specifies the lagged instruments used. The next two columns give the sample (Period) and the number of observations (Obs) on the stock returns. Columns 5 and 6 indicate whether the conditional model includes time-varying alpha ( $\alpha_t$ ) and the time-varying beta ( $\beta_t$ ), respectively. The last five columns summarize the regression results. Column 7 shows the ratio of an intercept (a pricing error) in the conditional model to that of the unconditional. Columns 8 and 9 report the point estimates of the time-varying alpha,  $\alpha_1$ , and their corresponding  $t$ -statistics,  $t(\alpha_1)$ . Columns 10 and 11 report the point estimates of the time-varying beta coefficient,  $b_1$ , and their corresponding  $t$ -statistics,  $t(b_1)$ . For each predictor, the table reports regression estimates corresponding to the largest in absolute value  $t$ -statistics. The abbreviations in the table are as follows. *TB1y* and *TB1vol* are the yield and volatility on the one-month Treasury bill, respectively. The variable *Cay* is the linear function of consumption, asset wealth, and labor income. *DY* is the dividend yield of the CRSP index. *Term* is a spread between long- and short-term bonds. *Default* is a spread between low- and high-grade corporate bonds. None stands for parameter values not used in the study; NA stands for results not reported.

(1) Reference	(2) Predictor	(3) Period	(4) Obs	(5) $\alpha_t$	(6) $\beta_t$	(7) $ \alpha_0/\alpha $	(8) $\alpha_1$	(9) $t(\alpha_1)$	(10) $b_1$	(11) $t(b_1)$		
Shanken (90, T3 & 4)	<i>TB1y</i>	5301–8212	360	Yes	Yes	NA	−0.48	−1.17	1.42	5.92		
	<i>TB1vol</i>										−5.70	−3.56
Cochrane (96, T8)	<i>DY</i>	47Q1–93Q4	188	No	Yes	NA	None	None	−0.53	−4.74		
	<i>Term</i>										−0.31	−1.76
Ferson & Schadt (96, T2)	<i>TB1y</i>	6801–9012	276	No	Yes	0.72	None	None	NA	NA		
	<i>DY</i>											
	<i>Term</i>											
	<i>Default</i>											
Jagannathan & Wang (96, T2)	<i>Default</i>	6307–9012	330	Yes	No	1.53	−65.7	−3.10	None	None		
Christopherson et al. (98, T1 & 2)	<i>TB1y</i>	7901–9012	144	Yes	Yes	0.77	−0.21	−2.01	NA	NA		
	<i>DY</i>										1.22	3.72
	<i>Term</i>										−0.21	−1.85
	<i>Cay</i>											
Lettau & Ludvigson (01, T3)	<i>Cay</i>	63Q3–98Q3	144	Yes	Yes	0.84	NA	NA	NA	NA		
Petkova & Zhang (05, T3)	<i>TB1y</i>	2701–0112	900	No	Yes	0.97	None	None	NA	NA		
	<i>DY</i>											
	<i>Term</i>											
	<i>Default</i>											

Using regression models like equation (2), the literature has produced a number of “stylized facts.” First, studies typically find that the intercept is smaller in the “conditional” model (2) than in the “unconditional” model (1):  $|\alpha| > |\alpha_0|$ . The interpretation of these studies is that the conditional CAPM does a better job of “explaining” average abnormal returns than the unconditional CAPM. Examples with this finding include Cochrane (1996), Ferson and Schadt (1996), Ferson and Harvey (1997), (1999), Lettau and Ludvigson (2001), and Petkova and Zhang (2005). Second, studies typically find evidence of time-varying betas: The coefficient estimate for  $b_1$  is statistically significant. Third, studies typically find that the conditional models fail to completely explain the dynamic properties of returns: The coefficient estimate for  $\alpha_1$  is significant, indicating a time-varying alpha. Our objective is to study the reliability of such inferences in the presence of persistent lagged instruments and data mining.

#### IV. The Simulation Design

The data in our simulations are generated according to:

$$(4) \quad \begin{aligned} r_{t+1} &= \beta_t r_{m,t+1} + u_{t+1}, \\ \beta_t &= 1 + Z_t^*, \\ r_{m,t+1} &= \mu_{sp} + kZ_t^* + w_{t+1}. \end{aligned}$$

We use the simulated data to run regression model (2), focusing on the  $t$ -statistics for the coefficients  $\{\alpha_0, \alpha_1, b_0, b_1\}$ . The variable  $Z_t^*$  in equation (4) is an unobserved latent variable that drives both expected market returns and time-varying betas. The term  $\beta_t$  in equation (4) is a time-varying beta coefficient. As  $Z_t^*$  has mean equal to zero, the expected value of beta is 1.0. When  $k \neq 0$ , there is an interaction between the time variation in beta and the expected market risk premium. A common persistent factor drives the movements in both expected returns and conditional betas. Common factors in time-varying betas and expected market premiums are important in asset pricing studies such as Chan and Chen (1988), Ferson and Korajczyk (1995), and Jagannathan and Wang (1996), and in conditional performance evaluation, as in Ferson and Schadt (1996). There is a zero intercept, or “alpha,” in the data generating process (4), consistent with asset pricing theory.

Because the spurious regression problem is driven by biased estimates of the standard error, at least in the context of regression (3), the choice of the standard error estimator is crucial. In a simulation exercise, it is possible to find an efficient unbiased estimator, since we know the “true” model that describes the simulated regression error. Of course, this will not be known in practice. To mimic the practical reality, the analyst in our simulations uses the popular heteroskedasticity and autocorrelation consistent (HAC) standard errors from Newey and West (1987) with an automatic lag selection procedure. The number of lags is chosen by computing the autocorrelations of the estimated residuals, and truncating

the lag length when the sample autocorrelations become “insignificant” at longer lags.<sup>2</sup>

### A. Modeling the Market Return

The market return data,  $r_{m,t+1}$ , are generated as follows. We set the parameter  $\mu_{sp} = 0.0071$  to equal the monthly average return of the S&P 500 stock index in excess of the one-month Treasury bill. The variance of the error is  $\sigma_w^2 = \sigma_{sp}^2 - k^2 \text{var}(Z^*)$ , where  $\sigma_{sp} = 0.057$  matches the S&P 500 return and  $\text{var}(Z^*) = 0.055$  is the estimated average monthly variance of the market betas on 58 randomly selected stocks from CRSP over the period 1926–1997.<sup>3</sup> The predictor variables follow an autoregressive process:

$$(5) \quad (Z_t^*, Z_t)' = \begin{Bmatrix} \rho^* & 0 \\ 0 & \rho \end{Bmatrix} (Z_{t-1}^*, Z_{t-1})' + (\varepsilon_t^*, \varepsilon_t)'$$

The assumption that the true expected return is autoregressive follows studies such as Lo and MacKinlay (1988), Conrad and Kaul (1988), Fama and French (1988), and Huberman and Kandel (1990).

The errors  $(\varepsilon_t^*, \varepsilon_t)$  are drawn randomly as a normal vector with mean zero and covariance matrix,  $\Sigma$ . The covariance matrix is diagonal,  $Z_t$  and  $Z_t^*$  are independent, and the true value of  $\delta_1$  in regression (3) is zero. We build up the time series of the  $Z$  and  $Z^*$  through the vector autoregression (5), where the initial values are drawn from a normal with mean zero and variances,  $\text{var}(Z)$  and  $\text{var}(Z^*)$ . The other parameters that calibrate the simulations,  $\{\mu, \sigma_u^2, \rho, \rho^*, \text{and } \Sigma\}$ , are described below.

### B. Modeling Data Mining

Our simulations capture the interaction between spurious regression and data mining, where the instruments to be mined are independent as in Foster, Smith, and Whaley (1997). There are  $L$  measured instruments over which the analyst searches for the “best” predictor of the test asset return, based on their Newey-West  $t$ -ratios in univariate regression (3). In equation (5),  $Z_t$  becomes a vector of length  $L$ , where  $L$  is the number of instruments through which the analyst sifts. The error terms  $(\varepsilon_t^*, \varepsilon_t)$  become an  $L+1$  vector with a diagonal covariance matrix; thus,  $\varepsilon_t^*$  is independent of  $\varepsilon_t$ . Following Ferson, Sarkissian, and Simin (2003), we

<sup>2</sup>Specifically, we compute 12 monthly sample autocorrelations and compare the values with a cutoff at two approximate standard errors:  $2/\sqrt{T}$ , where  $T$  is the sample size. The number of lags chosen is the minimum lag length at which no higher order autocorrelation is larger than two standard errors.

<sup>3</sup>We calibrate the variance of the betas to actual monthly data by randomly selecting 58 stocks with complete CRSP data for January 1926 through December 1997. Following Fama and French (1997), we estimate simple regression betas for each stock’s excess return against the S&P 500 excess return, using a series of rolling five-year windows. For each window, we also compute the standard error of the beta estimate. This produces a series of 805 beta estimates and standard error estimates of beta for each firm. We calibrate the variance of the true beta for each firm to equal the sample variance of the rolling beta estimates minus the average estimated variance of the estimator. Averaging the result across firms, the value of  $\text{var}(Z^*)$  is 0.0550. Repeating this exercise with firms that have data from January 1926 through the end of 2004 increases the number of months used from 864 to 948 but decreases the number of firms from 58 to 46. The value of  $\text{var}(Z^*)$  in this case is 0.0549.

compile a randomly selected sample of 500 potential instruments through which our simulated analyst sifts to mine the data. We select the 500 series randomly from a much larger sample of 10,866 potential variables as described in the Appendix. The 500 series are randomly ordered, and then permanently assigned numbers between 1 and 500. When a data miner in our simulations searches through, say 50 series, we use the sampling properties of the first 50 series to calibrate the simulations.

### C. Modeling Persistence

We use our sample of potential instruments to calibrate the amount of persistence in the “true” expected market returns,  $\rho^*$ . Ferson, Sarkissian, and Simin (2003) argue that if the instruments we see in the literature rise spuriously from data mining, they are likely to be more highly autocorrelated than the underlying “true” expected returns. However, if the instruments in the literature are a realistic representation of expected stock returns, their autocorrelations may be a good proxy for the persistence of the true expected returns. The mean autocorrelation of our 500 variables is 15% and the median is 2%. Of the 13 popular instruments from the literature surveyed by Ferson, Sarkissian, and Simin ((2003), Table 1), the median autocorrelation is 95%. We use  $\rho^* = 95\%$  for the results reported in the tables. We also experiment with smaller values of  $\rho^*$  as described below. With larger values of  $\rho^*$ , the biases we document become even more severe.

The autocorrelations of the observed instruments, denoted by the  $L$ -vector,  $\rho$ , are set equal to the sample autocorrelations of the first  $L$  instruments in our 500 potential instruments, rescaled around the value of 0.95.<sup>4</sup> The rescaling allows us to center the distribution of autocorrelations at 0.95 while preserving the range in the original data.<sup>5</sup> The simulations match the unconditional variances of the instruments,  $\text{var}(Z)$ , to the data. The first element of the covariance matrix  $\Sigma$  is equal to  $\sigma_w^2$ . For a typical  $i$ -th diagonal element of  $\Sigma$ , denoted by  $\sigma_i$ , the elements of  $\rho(Z_i)$  and  $\text{var}(Z_i)$  are given by the data, and we set  $\sigma_i^2 = [1 - \rho(Z_i)^2] \text{var}(Z_i)$ .

### D. Coefficients of Determination

In the asset pricing regressions, there are two ways to think about  $R$ -squared. The *predictive*  $R$ -squared,  $R_p^2$ , measures the proportion of the variance of the mar-

<sup>4</sup>We calibrate the true autocorrelations in the simulations to the sample autocorrelations, adjusted for first-order finite sample bias as:  $r + (1 + 3r)/T$ , where  $r$  is the OLS estimate of the autocorrelation and  $T$  is the sample size.

<sup>5</sup>The transformation is as follows. In the 500 instruments, the minimum bias-adjusted autocorrelation is  $-0.571$ , the maximum is  $0.999$ , and the median is  $0.02$ . We center the transformed distribution at about  $0.95$ . If the original autocorrelation  $\rho$  is less than  $0.95$ , we transform it to:

$$0.95 + (\rho - 0.02)\{(0.95 + 0.571)/(0.02 + 0.571)\}.$$

If the value is above  $0.95$ , we transform it to:

$$0.95 + (\rho - 0.02)\{(0.999 - 0.95)/(0.999 - 0.02)\}.$$

ket return that could be predicted if  $Z^*$  were observed:  $R_p^2 = \text{var}\{E(r_{t+1}|Z_t^*)\} / \text{var}(r_{t+1})$ . We choose the scale parameter,  $k$ , to match the values of  $R_p^2$  and  $\text{var}(Z^*)$ , which imply  $k^2 = \sigma_{sp}^2 R_p^2 / \text{var}(Z^*)$ .

The  $R$ -squares observed when the regressors include a contemporaneous market return, as in (2), will be higher than those of pure predictive regressions. Hence, we define the *contemporaneous*  $R$ -squared,  $R_c^2 = \text{var}\{\beta_t r_{m,t+1}\} / \text{var}(r_{t+1})$ . This is the  $R^2$  that could in principle be observed by using the contemporaneous market return as the regressor if the true value of the time-varying beta were known. The two versions of  $R$ -squared are monotonically related in our experiments and we report only the predictive  $R_p^2$  in the tables.

The final parameters of the simulations are chosen to match the values of  $\rho$  and  $R_p^2$  as follows. We set  $\sigma_\varepsilon^2 = \text{var}(Z^*)(1 - \rho^2)$  and we set  $\sigma_u^2$  so that the predictive  $R$ -squared of the return  $r_{t+1}$  is equal to the predictive  $R$ -squared of the market return,  $r_{m,t+1}$ .

In summary, the analyst in the simulations estimates regression model (2). He uses the lagged instrument,  $Z_t$ , which is independent of  $Z_t^*$ . He identifies the lagged instrument as the one that maximizes the absolute  $t$ -ratio in predictive regression (3). The true values of the coefficients in regression (2) are  $\alpha_0 = 0$ ,  $\alpha_1 = 0$ ,  $b_0 = 1$ , and  $b_1 = 0$ . Thus, the true alpha is zero as predicted by the asset pricing model and there is no time variation in the beta that is related to the measured instrument. Because of the interaction terms in the data generating process, the returns data will be conditionally heteroskedastic with an unobservable form. The analyst forms  $t$ -ratios for the coefficients using the Newey-West procedure as described above.<sup>6</sup>

## V. Results

### A. Cases with Small Amounts of Persistence

We first consider a special case of the model where we set  $\rho^* = 0$  in the data generating process for the market return and true beta, so that  $Z^*$  is white noise and  $\sigma^2(\varepsilon^*) = \text{var}(Z^*)$ . In this case, the predictable (but unobserved by the analyst) component of the stock market return and the betas follow a white noise process. We allow a range of values for the autocorrelation,  $\rho$ , of the measured instrument,  $Z_t$ , including values as large as 0.99. For a given value of  $\rho$ , we choose  $\sigma^2(\varepsilon) = \text{var}(Z^*)(1 - \rho^2)$ , so the measured instrument and the unobserved beta have the same variance. We find in this case that the critical values for all of the coefficients are well behaved. Thus, when the true expected returns and betas are not persistent, the use of even a highly persistent regressor does not create a spurious regression bias in the asset pricing regressions.

<sup>6</sup>Substituting the three equations of (4) together, we can express the data generating process for the return  $r_{t+1}$  as:

$$r_{t+1} = (1 + Z_t^*)(\mu_{sp} + kZ_t^* + w_{t+1}) + u_{t+1}.$$

Because of the interaction between the two  $Z_t^*$  terms, we transform the data generated from this expression to obtain the desired true parameter values in the model, which are  $\alpha_0 = \alpha_1 = b_1 = 0$ ,  $b_0 = 1$ . The transformed return is given by  $a + br_{t+1}$ , where the constants are  $b = [1 + 3k\mu_{sp} \text{var}(Z^*) / \text{var}(r_m)]^{-1}$ ,  $a = \mu_{sp} - b\{\mu_{sp} + k \text{var}(Z^*)\}$ . This transformation makes the means of  $r_{t+1}$  and  $r_{m,t+1}$  equal to each other.

It seems intuitive that there should be no spurious regression problem when there is no persistence in  $Z^*$ . Since the true coefficient on the measured instrument,  $Z$ , is zero, the error term in the regression is unaffected by the persistence in  $Z$  under the null hypothesis. When there is no spurious regression problem, there can be no interaction between spurious regression and data mining. Thus, standard corrections for data mining (e.g., White (2000)) can be used without concern in these cases.

In our second experiment, the measured instrument and the true beta have the same degree of persistence, but their persistence is not extreme. We fix  $\text{var}(Z) = \text{var}(Z^*)$  and choose, for a given value of  $\rho^* = \rho$ ,  $\sigma^2(\varepsilon) = \sigma^2(\varepsilon^*) = \text{var}(Z^*)(1 - \rho^2)$ . For values of  $\rho < 0.95$  and all values of  $R_p^2$ , the regressions seem generally well specified, even at sample sizes as small as  $T = 66$ . These findings are similar to the findings of Ferson, Sarkissian, and Simin (2003) for the case of the pure predictive regression (3). Thus, the regressions appear to be well specified when the autocorrelation of the true predictor is below 0.95.

## B. Cases with Persistence

Table 2 summarizes simulation results for a case that allows data mining and spurious regression. In this experiment, the true persistence parameter  $\rho^*$  is set equal to 0.95. The table summarizes the results for time-series samples of  $T = 66$ ,  $T = 350$ , and  $T = 960$ . The number of variables over which the artificial agent searches in mining the data ranges from 1 to 250. We focus on the two abnormal return coefficients,  $\{\alpha_0, \alpha_1\}$  and on the time-varying beta coefficient,  $b_1$ .

Table 2 shows that the means of the coefficient  $\alpha_0$ , the fixed part of the alpha, are close to zero, and they get closer to zero as the number of observations increases as expected of a consistent estimator. The 5% critical  $t$ -ratios for  $\alpha_0$  are reasonably well specified at the larger sample sizes although there is some bias at  $T = 66$ , where the critical values rise with the extent of data mining. Data mining has little effect on the intercepts at the larger sample sizes. Since the lagged instrument has a mean of zero, the intercept is the average conditional alpha. Thus, the issue of data mining for predictive variables appears to have no serious implications for measures of average abnormal performance in the conditional asset pricing regressions provided  $T > 66$ . This justifies the use of such models for studying the cross section of average equity returns.

The coefficients  $\alpha_1$ , which represent the time-varying part of the conditional alphas, present a different pattern. We would expect a data mining effect, given that the data are mined based on the coefficients on the lagged predictor in the simple predictive regression. The presence of the interaction term, however, would be expected to attenuate the bias in the standard errors, compared with the simple predictive regression. Table 2 shows only a small effect of data mining on the  $\alpha_1$  coefficient, but a large effect on its  $t$ -ratio. The overall effect is greatest at the smaller sample size ( $T = 66$ ), where the critical  $t$ -ratios for the intermediate  $R_p^2$  values (10% predictive  $R^2$ ) vary from about 2.4 to 5.2 as the number of variables mined increases from 1 to 250. The bias diminishes with  $T$ , especially when the number of mined variables is small, and for  $L = 1$  there is no substantial bias at  $T = 360$  or  $T = 960$  months.

TABLE 2  
 Simulating a Conditional Asset Pricing Model

Table 2 shows the results of 10,000 simulations of the estimates from the conditional asset pricing model, allowing for possible data mining of the lagged instruments. The regression model is:

$$r_{t+1} = \alpha_0 + \alpha_1 \bar{Z}_t + b_0 r_{m,t+1} + b_1 r_{m,t+1} \bar{Z}_t + u_{t+1}$$

$T$  is the number of time-series observations,  $L$  is the number of lagged instruments mined, and  $R_p^2$  is the true predictive  $R^2$  in the artificial data generating process.

$R_p^2$	$T = 66$			$T = 350$			$T = 960$		
	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$
<i>Panel A. Means: <math>\alpha_0</math></i>									
0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.01	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.05	-0.001	-0.001	-0.001	0.000	-0.001	0.000	0.000	0.000	0.000
0.1	-0.002	-0.002	-0.002	0.000	-0.001	-0.001	0.000	0.000	0.000
0.15	-0.002	-0.002	-0.002	-0.001	-0.001	-0.001	0.000	0.000	0.000
<i>Panel B. Critical 5% t-Statistics for <math>\alpha_0</math></i>									
0.001	2.280	2.603	2.855	1.999	2.061	2.146	1.996	2.005	2.115
0.005	2.266	2.540	2.792	1.994	2.058	2.135	2.013	2.002	2.104
0.01	2.253	2.508	2.759	2.000	2.045	2.125	2.016	2.002	2.098
0.05	2.153	2.408	2.728	1.974	1.998	2.094	2.021	1.991	2.100
0.1	2.088	2.388	2.652	1.977	2.000	2.030	2.058	2.008	2.073
0.15	2.065	2.382	2.597	1.968	1.960	1.987	2.069	2.031	2.041
<i>Panel C. Means: <math>\alpha_1</math></i>									
0.001	0.001	0.007	-0.003	-0.001	-0.002	0.003	-0.001	-0.002	-0.001
0.005	0.001	0.005	-0.001	-0.001	-0.001	0.002	-0.001	-0.002	-0.001
0.01	0.001	0.005	0.000	-0.001	-0.001	0.002	-0.001	-0.002	-0.001
0.05	0.001	0.005	-0.002	-0.001	-0.001	0.002	-0.001	-0.001	-0.002
0.1	0.001	0.005	0.000	-0.001	-0.001	0.002	-0.001	-0.001	0.000
0.15	0.001	0.004	0.000	-0.001	0.000	0.002	-0.001	-0.001	0.001
<i>Panel D. Critical 5% t-Statistics for <math>\alpha_1</math></i>									
0.001	2.392	3.992	5.305	2.023	3.240	3.891	1.910	3.097	3.748
0.005	2.390	3.961	5.252	2.024	3.206	3.874	1.905	3.092	3.719
0.01	2.387	3.912	5.198	2.025	3.198	3.872	1.902	3.091	3.712
0.05	2.412	3.924	5.237	2.039	3.172	3.855	1.902	3.062	3.706
0.1	2.426	3.912	5.163	2.036	3.159	3.837	1.912	3.040	3.690
0.15	2.423	3.913	5.086	2.024	3.155	3.785	1.909	3.024	3.629
<i>Panel E. Means: <math>b_1</math></i>									
0.001	-0.041	-0.017	0.026	-0.003	-0.002	0.011	0.010	0.008	-0.009
0.005	-0.038	-0.019	0.061	-0.003	0.001	0.002	0.009	0.007	-0.005
0.01	-0.038	-0.012	0.055	-0.003	0.000	0.004	0.008	0.004	-0.006
0.05	-0.039	-0.014	0.077	-0.003	0.006	-0.003	0.006	0.001	0.001
0.1	-0.040	-0.018	0.058	-0.003	0.010	0.003	0.005	0.000	0.004
0.15	-0.041	-0.015	0.062	-0.003	0.014	0.007	0.004	0.001	0.000
<i>Panel F. Critical 5% t-Statistics for <math>b_1</math></i>									
0.001	2.576	2.534	2.634	2.122	2.098	2.175	1.996	2.013	2.075
0.005	2.574	2.579	2.611	2.116	2.138	2.210	2.022	2.036	2.075
0.01	2.583	2.574	2.597	2.114	2.133	2.219	2.027	2.071	2.126
0.05	2.603	2.588	2.597	2.149	2.212	2.336	2.027	2.121	2.219
0.1	2.612	2.614	2.596	2.157	2.297	2.451	2.024	2.188	2.475
0.15	2.610	2.601	2.657	2.156	2.361	2.614	2.018	2.322	2.722

The results on the  $\alpha_1$  coefficient are interesting in three respects. First, the critical  $t$ -ratios vary by only small amounts across the rows of Table 2. This indicates very little interaction between the spurious regression and data mining effects. Second, Table 2 shows a smaller data mining effect than that observed on the pure predictive regression (see Ferson, Sarkissian, and Simin (2003) Table 3). Thus, data mining corrections for predictive regressions will overcompensate in this setting. Third, the critical  $t$ -ratios for  $\alpha_1$  become smaller in Table 2 as the sample size is increased. This is just the opposite of what is found for the simple

predictive regressions, where the inconsistency in the standard errors makes the critical  $t$ -ratios larger at larger sample sizes. Thus, the sampling distributions for time-varying alpha coefficients are not likely to be well approximated by simple corrections.<sup>7</sup>

Table 2 does not report the  $t$ -statistics for  $b_0$ , the constant part of the beta estimate. These are generally unbiased across all of the samples, except that the critical  $t$ -ratios are slightly inflated at the smaller sample size ( $T = 66$ ) when data mining is not at issue ( $L = 1$ ).

Finally, Table 2 shows results for the  $b_1$  coefficients and their  $t$ -ratios, which capture the time-varying component of the conditional betas. Here, the average values and the critical  $t$ -ratios are barely affected by the number of variables mined. When  $T = 66$ , the critical  $t$ -ratios stay in a narrow range from about 2.5 to 2.6, and they cluster more closely around a value of 2.0 at the larger sample sizes. There are no discernible effects of data mining on the distribution of the time-varying beta coefficients except when the  $R^2$  values are very high. This is an important result in the context of the conditional asset pricing literature, which we characterize as having mined predictive variables based on regression (3). Our results suggest that the empirical evidence in this literature for time-varying betas based on regression model (2) is relatively robust to the data mining.

### C. Suppressing Time-Varying Alphas

Some studies in the conditional asset pricing literature use regression models with interaction terms, but without the time-varying alpha component (e.g., Cochrane (1996), Ferson and Schadt (1996), and Ferson and Harvey (1999)). Since the time-varying alpha component is the most troublesome term in the presence of spurious regression and data mining effects, it is interesting to ask if regressions that suppress this term may be better specified. Table 3 presents results for models in which the analyst runs regressions without the  $\alpha_1$  coefficient. The results suggest that the average alpha coefficient,  $\alpha_0$ , and its  $t$ -statistic are well specified regardless of data mining and potential spurious regression. Thus, once again we find little cause for concern about the inferences on average abnormal returns using the conditional asset pricing regressions even though they use persistent, data mined lagged regressors.

The distribution of the average beta estimate,  $b_0$ , is not shown in Table 3. The results are similar to those obtained in a factor model regression where no lagged instrument is used. The coefficients and standard errors generally appear well specified. However, we find that the coefficient measuring the time-varying beta is somewhat more susceptible to bias than in the regression that includes  $\alpha_1$ . The  $b_1$  coefficient is biased, especially when  $T = 66$ , and its mean varies with the number of instruments mined. The critical  $t$ -ratios are inflated at the higher values of  $R_p^2$  and when more instruments are mined.

Including the time-varying alpha in the regression helps “soak up” the bias so that it will not adversely effect the time-varying beta estimate. These experiments

---

<sup>7</sup>We conducted some experiments in which we applied a simple local to unity correction to the  $t$ -ratios dividing by the square root of the sample size. We found that this correction does not result in a  $t$ -ratio that is approximately invariant to the sample size.

TABLE 3  
 Simulating a Conditional Asset Pricing Model with No Time-Varying Alpha

Table 3 shows the results of 10,000 simulations of the estimates from the conditional asset pricing model with no time-varying alpha, allowing for the possibility of data mining for the lagged instruments. The regression model is:

$$r_{t+1} = \alpha_0 + b_0 r_{m,t+1} + b_1 r_{m,t+1} Z_t + u_{t+1}.$$

$T$  is the number of time-series observations,  $L$  is the number of lagged instruments mined, and  $R_D^2$  is the true predictive  $R^2$  in the artificial data generating process.

$R_D^2$	$T = 66$			$T = 350$			$T = 960$		
	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$
<i>Panel A. Means: <math>\alpha_0</math></i>									
0.001	0.000	0.000	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
0.005	-0.001	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
0.01	-0.001	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
0.05	-0.002	-0.002	-0.002	0.000	0.000	-0.001	0.000	0.000	0.000
0.1	-0.002	-0.002	-0.003	0.000	-0.001	-0.001	0.000	0.000	0.000
0.15	-0.003	-0.003	-0.003	-0.001	-0.001	-0.001	0.000	0.000	-0.001
<i>Panel B. Critical 5% t-Statistics for <math>\alpha_0</math></i>									
0.001	2.165	2.146	2.123	1.951	1.988	1.947	1.990	1.965	1.934
0.005	2.138	2.114	2.089	1.949	1.981	1.929	1.983	1.961	1.922
0.01	2.132	2.102	2.054	1.944	1.966	1.918	1.992	1.956	1.931
0.05	2.042	2.020	1.986	1.907	1.921	1.917	1.983	1.970	1.921
0.1	1.984	1.956	1.929	1.896	1.910	1.885	1.988	1.962	1.904
0.15	1.949	1.922	1.893	1.900	1.884	1.829	1.993	1.965	1.869
<i>Panel C. Means: <math>b_1</math></i>									
0.001	0.007	-0.006	-0.043	-0.008	0.004	-0.001	-0.003	0.013	-0.012
0.005	0.008	-0.017	-0.060	-0.006	-0.003	0.000	-0.003	0.013	-0.008
0.01	0.008	-0.015	-0.064	-0.006	-0.003	-0.004	-0.003	0.010	-0.007
0.05	0.010	-0.031	-0.047	-0.003	-0.003	-0.003	-0.002	0.007	-0.001
0.1	0.010	-0.020	-0.035	-0.002	-0.005	-0.001	-0.002	0.000	0.001
0.15	0.010	-0.029	-0.042	0.000	-0.003	-0.003	-0.002	0.002	-0.002
<i>Panel D. Critical 5% t-Statistics for <math>b_1</math></i>									
0.001	2.630	2.605	2.639	2.157	2.128	2.218	1.987	2.136	2.147
0.005	2.636	2.646	2.631	2.156	2.145	2.246	1.991	2.162	2.170
0.01	2.661	2.665	2.643	2.163	2.150	2.256	1.987	2.154	2.216
0.05	2.656	2.748	2.739	2.146	2.257	2.441	1.988	2.267	2.476
0.1	2.629	2.811	2.861	2.175	2.378	2.618	1.994	2.395	2.639
0.15	2.607	2.857	3.001	2.201	2.466	2.755	2.008	2.479	2.828

suggest that if one is interested in obtaining good estimates of conditional betas, then in the presence of data mining and persistent lagged instruments, the time-varying alpha term should be included in the regression.

#### D. Suppressing Time-Varying Betas

There are examples in the literature where the regression is run with a linear term for a time-varying conditional alpha but no interaction term for a time-varying conditional beta (e.g., Jagannathan and Wang (1996)). Table 4 considers this case. First, the coefficient for the average beta in the regression with no  $b_1$  term (not shown in Table 4) is reasonably well specified and largely unaffected by data mining on the lagged instrument. We find that the coefficients for alpha,  $\alpha_0$  and  $\alpha_1$ , behave similarly to the corresponding coefficients in the full model. The estimates of the average alpha are reasonably well behaved and only mildly affected by the extent of data mining at smaller sample sizes. The bias in  $\alpha_1$  is severe. The bias leads the analyst to overstate the evidence for a time-varying alpha, and the bias is worse as the amount of data mining increases. Thus, the evidence

in the literature for time-varying alphas, based on these asset pricing regressions, is likely to be overstated.

TABLE 4  
 Simulating a Conditional Asset Pricing Model with No Time-Varying Beta

Table 4 shows the results of 10,000 simulations of the estimates from the conditional asset pricing model with no time-varying beta, allowing for the possibility of data mining for the lagged instruments. The regression model is:

$$r_{t+1} = \alpha_0 + \alpha_1 Z_t + b_0 r_{m,t+1} + u_{t+1}.$$

$T$  is the number of time-series observations,  $L$  is the number of lagged instruments mined, and  $R_p^2$  is the true predictive  $R^2$  in the artificial data generating process.

$R_p^2$	$T = 66$			$T = 350$			$T = 960$		
	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$
<i>Panel A. Means: <math>\alpha_0</math></i>									
0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.01	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.05	-0.001	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
0.1	-0.002	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
0.15	-0.002	-0.002	-0.002	-0.001	-0.001	0.000	0.000	0.000	0.000
<i>Panel B. Critical 5% t-Statistics for <math>\alpha_0</math></i>									
0.001	2.239	2.592	2.794	1.991	2.080	2.182	2.004	2.047	2.092
0.005	2.219	2.533	2.764	1.982	2.070	2.147	2.014	2.036	2.084
0.01	2.206	2.511	2.731	1.988	2.069	2.150	2.013	2.018	2.089
0.05	2.124	2.429	2.656	1.970	2.056	2.148	2.019	2.012	2.130
0.1	2.065	2.389	2.607	1.968	2.034	2.158	2.035	2.010	2.160
0.15	2.015	2.356	2.546	1.978	2.033	2.177	2.072	2.010	2.167
<i>Panel C. Means: <math>\alpha_1</math></i>									
0.001	0.001	0.000	0.002	-0.001	0.001	0.000	0.000	-0.001	0.000
0.005	0.001	0.000	0.001	-0.001	0.001	0.001	0.000	-0.001	0.001
0.01	0.001	0.000	0.001	-0.001	0.001	0.001	0.000	-0.001	0.001
0.05	0.000	0.001	0.003	-0.001	0.001	0.001	0.000	-0.001	0.002
0.1	0.000	0.001	0.005	-0.001	0.000	0.000	0.000	0.000	0.001
0.15	0.000	0.000	0.003	-0.001	0.000	-0.001	0.000	-0.001	0.000
<i>Panel D. Critical 5% t-Statistics for <math>\alpha_1</math></i>									
0.001	2.307	4.015	5.298	2.068	3.172	3.912	2.020	3.073	3.794
0.005	2.311	4.005	5.169	2.066	3.149	3.890	2.017	3.061	3.761
0.01	2.312	3.998	5.142	2.061	3.148	3.891	2.017	3.058	3.754
0.05	2.322	3.968	5.040	2.054	3.156	3.885	2.003	3.029	3.739
0.1	2.323	3.908	5.041	2.056	3.138	3.849	2.012	3.001	3.713
0.15	2.323	3.901	5.075	2.056	3.124	3.783	2.005	3.003	3.659

### E. A Cross Section of Asset Returns

We extend the simulations to study a cross section of asset returns. We use five book-to-market (BM) quintile portfolios, equally weighted across the size dimension, as an illustration. The data are courtesy of Kenneth French. In these experiments, the cross section of assets features cross-sectional variation in the true conditional betas. Instead of  $\beta_t = 1 + Z_t^*$ , the betas are  $\beta_t = \beta_0 + \beta_1 Z_t^*$ , where the coefficients  $\beta_0$  and  $\beta_1$  are the estimates obtained from regressions of each quintile portfolio's excess return on the market portfolio excess return and the product of the market portfolio with the lagged value of the dividend yield. The set of  $\beta_0$ 's is  $\{1.259, 1.180, 1.124, 1.118, 1.274\}$ , and the set of  $\beta_1$ 's is  $\{-1.715,$

1.000, 3.766, 7.646, 8.970}.<sup>8</sup> The true predictive  $R$ -squared in the artificial data generating process is set to 0.5%. This value matches the smallest  $R$ -squared from the regression of the market portfolio on the lagged dividend yield with a window of 60 months.

Table 5 shows simulation results for the conditional model with time-varying alphas and betas. The means of the  $b_0$  and  $b_1$  coefficients are shown in excess of their true values in the simulations. The critical  $t$ -statistics for both  $\alpha_1$  and  $b_1$  are generally similar to the case where  $R_p^2 = 0.5\%$  in Table 2. As before, there is a large bias in the  $t$ -statistic for  $\alpha_1$  that increases with data mining but decreases somewhat with the sample size. The  $t$ -statistics for the time-varying betas are generally well specified.

We conduct additional experiments using the cross section of asset returns, where the conditional asset pricing regression suppresses either the time-varying alphas or the time-varying betas. The results are similar to those in Table 5. When the time-varying betas are suppressed, there is severe bias in  $\alpha_1$  that diminishes somewhat with the sample size. When time-varying alphas are suppressed, there is a mild bias in  $b_1$ .

## F. Revisiting Previous Evidence

In this section, we explore the impact of the joint effects of data mining and spurious regression bias on the asset pricing evidence. First, we revisit the studies listed in Table 1. Consider the models with both time-varying alphas and betas. If the data mining searches over 250 variables predicting the test asset return and  $T = 350$ , the 5% cut-off value to apply to the  $t$ -statistic for  $\alpha_1$  is larger than 3.8 in absolute value. For smaller sample sizes, the cut-off value is higher. Note from Table 1 that the largest  $t$ -statistic for  $\alpha_1$  in Shanken (1990) with a sample size of 360 is  $-3.57$  on the T-bill volatility, while the largest  $t$ -statistic for  $\alpha_1$  in Christopherson, Ferson, and Glassman (1998) with a sample size of 144 is 3.72 on the dividend yield. This means that the significance of the time-varying alphas in both studies is questionable. However, the largest  $t$ -statistic for  $b_1$  in Shanken (1990) exceeds the empirical 5% cutoff, irrespective of spurious regression and data mining adjustments. This illustrates that the evidence for time-varying beta is robust to the joint effects of data mining and spurious regression bias, while the evidence for time-varying alphas is fragile.

Now consider the model with no time-varying alpha. If the data mining searches over 250 variables to predict the test asset return, the 5% cut-off value to apply to the  $t$ -statistic on  $b_1$  is less than 3.5 in absolute value. Cochrane (1996) reports a  $t$ -statistic of  $-4.74$  on the dividend yield in a time-varying beta with a sample of  $T = 186$ . Thus, we find no evidence to doubt the inference that there is a time-varying beta. (However, the significance at a 10% level of the term premium in the time-varying beta with a  $t$ -statistic of  $-1.76$  is in doubt.)

---

<sup>8</sup>The  $\beta_1$  coefficient for the BM2 portfolio is 1.0, replacing the estimated value of 0.047. When the  $\beta_1$  coefficient is 0.047, the simulated return becomes nearly perfectly correlated with  $r_m$  and the simulation is uninformative. The dividend yield is demeaned and multiplied by 10. The dividend yield has the largest average sample correlation with the five BM portfolios among the standard instruments we examine.

TABLE 5  
Conditional Asset Pricing Models with a Cross Section of Returns

Table 5 shows the results of 10,000 simulations from a conditional asset pricing model, allowing for the possibility of data mining for the lagged instruments. The dependent variables are book-to-market quintile portfolios. The regression model is:

$$r_{t+1} = \alpha_0 + \alpha_1 Z_t + b_0 r_{m,t+1} + b_1 r_{m,t+1} Z_t + u_{t+1}.$$

$T$  is the number of time-series observations and  $L$  is the number of lagged instruments mined. The true predictive  $R^2$  in the artificial data generating process is 0.005.

BM Quintile	$T = 66$			$T = 350$			$T = 960$		
	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$	$L = 1$	$L = 25$	$L = 250$
<i>Panel A. Means: <math>\alpha_0</math></i>									
BM1 (low)	-0.002	-0.001	-0.001	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
BM2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BM3	0.000	0.002	0.001	0.002	0.002	0.002	0.002	0.002	0.003
BM4	0.001	0.003	-0.001	0.004	0.005	0.006	0.005	0.006	0.007
BM5 (high)	0.004	0.002	0.006	0.006	0.006	0.005	0.008	0.007	0.006
<i>Panel B. Critical 5% t-Statistics for <math>\alpha_0</math></i>									
BM1 (low)	2.157	2.593	2.705	1.691	1.847	1.914	1.526	1.618	1.651
BM2	2.297	2.523	2.742	1.916	2.067	2.156	1.960	2.034	2.056
BM3	2.296	2.681	2.916	2.093	2.218	2.380	2.143	2.206	2.301
BM4	2.343	2.686	3.060	2.059	2.253	2.399	2.132	2.233	2.368
BM5 (high)	2.369	2.659	3.099	2.135	2.233	2.326	2.269	2.284	2.359
<i>Panel C. Means: <math>\alpha_1</math></i>									
BM1 (low)	0.002	0.000	0.001	-0.001	-0.004	0.002	0.000	0.000	0.000
BM2	0.000	0.000	0.002	0.001	0.000	0.001	0.000	0.000	0.001
BM3	0.009	0.006	-0.008	-0.001	-0.005	-0.008	0.001	0.007	0.001
BM4	0.019	-0.027	-0.004	0.007	-0.003	0.012	-0.003	-0.003	0.005
BM5 (high)	0.021	0.028	-0.068	0.005	0.028	0.027	0.002	-0.006	-0.009
<i>Panel D. Critical 5% t-Statistics for <math>\alpha_1</math></i>									
BM1 (low)	2.381	4.088	5.382	2.037	3.243	3.917	1.962	3.115	3.813
BM2	2.390	3.884	4.956	2.025	3.145	3.793	1.971	3.044	3.637
BM3	2.418	4.146	5.720	1.972	3.264	3.999	1.952	3.148	3.804
BM4	2.403	4.263	5.705	2.078	3.240	3.934	1.999	3.179	3.807
BM5 (high)	2.417	4.227	5.594	2.005	3.271	4.076	2.021	3.134	3.817
<i>Panel E. Means: <math>b_1</math></i>									
BM1 (low)	0.018	0.007	-0.054	-0.007	0.002	0.051	0.021	-0.001	0.004
BM2	-0.003	0.032	-0.015	0.000	-0.004	-0.009	-0.010	0.000	-0.002
BM3	0.108	0.050	0.128	0.066	-0.033	-0.042	-0.041	-0.004	0.034
BM4	-0.230	-0.050	-0.062	-0.016	-0.054	0.087	0.028	-0.068	0.032
BM5 (high)	-0.479	0.075	-0.389	0.041	0.113	-0.006	-0.136	0.032	0.058
<i>Panel F. Critical 5% t-Statistics for <math>b_1</math></i>									
BM1 (low)	2.612	2.522	2.548	2.065	2.168	2.181	2.045	2.115	2.061
BM2	2.521	2.591	2.559	2.146	2.126	2.149	2.056	2.056	2.067
BM3	2.573	2.501	2.585	2.083	2.159	2.089	2.103	2.023	2.061
BM4	2.624	2.556	2.552	2.048	2.114	2.091	2.130	2.004	2.035
BM5 (high)	2.628	2.640	2.536	2.177	2.149	2.068	2.084	2.050	1.982

Finally, consider the model with no time-varying beta. If the data mining searches over 25 variables to predict the test asset return, then the 5% cut-off value to apply to the  $t$ -statistic on  $\alpha_1$  is larger than 3.1 in absolute value. The largest  $t$ -statistic in Jagannathan and Wang (1996) with a sample size of 330 is 3.1. Therefore, their evidence for a time-varying alpha does not survive even with a modest amount of data mining.

An empirical example further illustrates the effects of using the correct cut-offs to evaluate the evidence for time-varying alphas. We use the value-growth spread portfolio excess return as the test asset and a conditional CAPM as the asset pricing model. The lagged conditioning variable is the term spread, measured as the lagged difference between 10- and one-year constant maturity yields from

the Federal Reserve Data Base.<sup>9</sup> The regressions include both time-varying betas and alphas, and the monthly returns cover the period from April 1953 through the end of 2005. We summarize results for samples of length  $T = 66$  and  $T = 350$ . We roll the regressions through the overall sample and use the average of the rolling regression coefficients and the average absolute  $t$ -ratios as proxies for the expected results, given a randomly chosen sample of each size. The average absolute  $t$ -ratios for the two sample sizes are shown below the regression equation:

$$(6) \quad r_{t+1} = \alpha_0 + \alpha_1 Z_t + b_0 r_{m,t+1} + b_1 r_{m,t+1} Z_t + u_{t+1}.$$

$T = 66$	1.559	2.199	21.32	2.897
$T = 350$	0.815	1.978	21.34	0.952

Figure 2 presents time-series plots of the absolute  $t$ -ratios for the coefficients,  $\alpha_1$ , from the rolling regressions. We evaluate the significance of the  $t$ -ratios using the simulations previously described, assuming  $\rho^* = 0.95$  and the true predictive  $R$ -squared is 10%. The cut-off levels for the absolute  $t$ -statistics with different amounts of data mining are depicted as horizontal lines in the figure.

Figure 2 shows that the expected  $t$ -ratio for a time-varying alpha is just above the 5% critical cut-off level for  $T = 66$ , and very close to the cutoff for  $T = 350$  when there is no adjustment for data mining. However, the results vary with the sample period. When we sort the sample  $t$ -statistics from high to low, we find that about  $\frac{1}{3}$  are above the cutoff with no data mining when  $T = 66$  and 53% are above the cutoff when  $T = 350$ . Thus, we are likely to find evidence of time-varying alphas when there is no adjustment for data mining.

Figure 2 also shows as horizontal bars the 5% critical cut-off levels when the lagged instrument is mined to predict the test asset, searching over  $L = 25$  and  $L = 250$  instruments. When  $L = 25$  and  $T = 66$ , only 15% of the  $t$ -statistics lie above the cutoff, and less than 10% exceed the cutoff when  $L = 250$ . None of the  $t$ -statistics appear significant when  $T = 350$  and  $L = 25$  or  $L = 250$ . Thus, even modest amounts of data mining are likely to change the inferences about time-varying alphas. In particular, given a typical sample with  $T = 350$ , there is no evidence for a time-varying alpha related to the term premium.

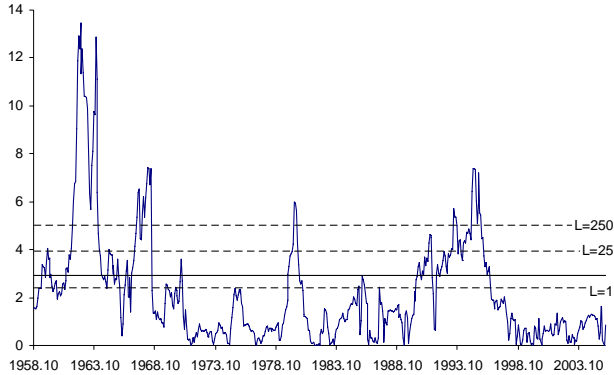
## VI. Robustness

This section summarizes the results of a number of additional experiments. We extend the simulations to consider examples with more than a single lagged instrument. We also consider asset pricing models with multiple factors, motivated by Merton's (1973) multiple beta model. We also examine models where the data mining to select the lagged instruments focuses on predicting the market portfolio return instead of the test asset returns. Tables with the results of these additional experiments are available from the authors.

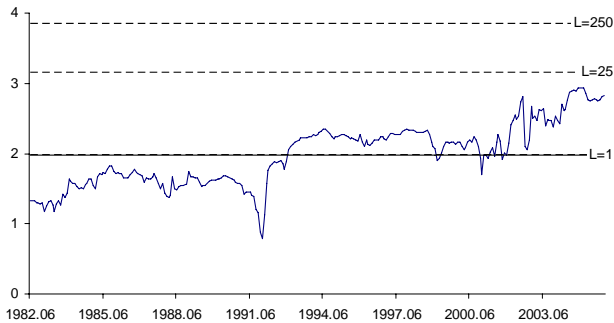
<sup>9</sup>We also examined the dividend yield and a short-term Treasury yield and the overall impressions are similar.

FIGURE 2  
Rolling Window Absolute Values of Conditional Alpha  $t$ -Statistic  
 $Z = \text{Term Spread}$

Graph A.  $T = 66$



Graph B.  $T = 350$



### A. Multiple Instruments

The experiments summarized above focus on a single lagged instrument, while many studies in the literature use multiple instruments. We modify the simulations, assuming that the researcher mines two independent instruments with the largest absolute  $t$ -statistics and then uses both of them in the conditional asset pricing model with time-varying betas and alphas. (Thus, there are two  $a_1$  coefficients and two  $b_1$  coefficients.) These simulations reveal that the statistical behavior of both coefficients is similar to each other and similar to our results as reported in Table 2.

### B. Multiple Beta Models

We extend the simulations to study models with three state variables or factors. In building the three-factor model, we make the following assumptions.

All three risk premiums are linear functions of one instrument,  $Z^*$ . The factors differ in their unconditional means and their disturbance terms, which are correlated with each other. The variance-covariance matrix of the disturbance terms matches that of the residuals from regressions of the three Fama-French (1993), (1996) factors on the lagged dividend yield. The true coefficients for the asset return on all three factors and their interaction terms with the correct lagged instrument,  $Z^*$ , are set to unity. Thus, the true conditional betas on each factor are equal to  $1 + Z^*$ . We find that the bias in the  $t$ -statistic for  $\alpha_1$  remains and is similar to the simulation in Table 2. There are no biases in the  $t$ -statistics associated with the  $b_1$ 's for the larger sample sizes.

### C. Predicting the Market Return

Much of the previous literature looked at more than one asset to select predictor variables. For the examples reported in the previous tables, the data mining is conducted by attempting to predict the excess returns of the tests' assets. But a researcher might also choose instruments to predict the market portfolio return. We examine the sensitivity of the results to this change in the simulation design. The results for the conditional asset pricing model with both time-varying alphas and betas are reexamined. Recall that when the instrument is mined to predict the test asset return, there is an upward bias in the  $t$ -statistic for  $\alpha_1$ . The bias increases with data mining and decreases somewhat with  $T$ . When the instruments are mined to predict the market, the bias in  $\alpha_1$  is small and is confined to the smaller sample size,  $T = 66$ . Mining to predict the market return has little impact on the sampling distribution of  $b_1$ .

## VII. Conclusions

We study regression models for conditional asset pricing models in which lagged variables are used to model conditional betas and alphas. We focus on the finite sample properties of the models in view of two important issues. The first issue is data mining and the second is spurious regression caused by predictor variables that are persistent time series. These two issues have been studied separately in many previous papers, and their combined effects for simple predictive regressions are studied by Ferson, Sarkissian, and Simin (2003). This study addresses the joint effects of data mining and spurious regression on models with time-varying alphas and betas.

The conditional asset pricing literature has, for the most part, used the same variables that were discovered based on simple predictive regressions, and our analysis characterizes the problem by assuming the data mining occurs in this way. Our results relate to several stylized facts that the literature on conditional asset pricing has produced.

Previous studies find evidence that the intercept, or average alpha, is smaller in a conditional model than in an unconditional model, suggesting for example that the conditional CAPM does a better job of explaining average abnormal returns. Our simulation evidence finds that the estimates of the average alphas in the conditional models are reasonably well specified in the presence of spurious

regression and data mining, at least for samples larger than  $T = 66$ . Some caution should be applied in interpreting the common 60-month rolling regression estimator, but otherwise we take no issue with the stylized fact that conditional models deliver smaller average alphas.

Studies typically find evidence of time-varying betas based on significant interaction terms. Here again we find little cause for concern. The coefficient estimator for the interaction term is well specified in larger samples, and largely unaffected by data mining in the presence of persistent lagged regressors. There is an exception when the model is estimated without a linear term in the lagged instrument. In this case, the coefficient measuring the time-varying beta is slightly biased. Thus, when the focus of the study is to estimate accurate conditional betas, we recommend that a linear term be included in the regression model.

Studies typically find that even conditional models fail to completely explain the dynamic properties of stock returns. That is, the estimates indicate time-varying conditional alphas. We find that this result is the most problematic. The estimates of time variation in alpha inherit biases similar to, if somewhat smaller than, the biases in predictive regressions. We use our simulations to revisit the evidence of several prominent studies. Our analysis suggests that the evidence for time-varying alphas in the current literature should be viewed with some suspicion. Perhaps, the current generation of conditional asset pricing models does a better job of capturing the dynamic behavior of asset returns than existing studies suggest.

## Appendix. The Sample of 500 Instruments

All the data come from the web site [economagic.com](http://economagic.com): economic time-series page, maintained by Ted Bos. The sample consists of all monthly series listed on the main home page of the site, except under the headings of LIBOR, Australia, Bank of Japan, and Central Bank of Europe. From the Census Bureau, we exclude building permits by region, state, and metro areas (more than 4,000 series). From the Bureau of Labor Statistics, we exclude all non-civilian labor force data and state, city, and international employment (more than 51,000 series). We use the CPI measures from the city average listings, but include no finer subcategories. The PPI measures include the aggregates and the two-digit subcategories. From the Department of Energy, we exclude data in Section 10, the international energy series.

We first randomly select (using a uniform distribution) 600 out of the 10,866 series that were left after the above exclusions. From these 600, we eliminated series that mixed quarterly and monthly data and extremely sparse series, and took the first 500 from what remained.

Because many of the data are reported in levels, we tested for unit roots using an augmented Dickey-Fuller test (with a zero order time polynomial). We could not reject the hypothesis of a unit root for 361 of the 500 series and we replaced these series with their first differences.

We estimate a sample correlation matrix of the 500 instruments as follows. We take each pair of instruments and compute the sample correlation between the two series, using all of the observations for which our data for the two series overlap. For some pairs, there is no overlapping data. For these cases, we substitute the average of all the sample correlations that we can compute.

## References

- Amihud, Y., and C. M. Hurvich. "Predictive Regressions: A Reduced-Bias Estimation Method." *Journal of Financial and Quantitative Analysis*, 39 (2004), 813–841.
- Ang, A., and G. Bekaert. "Stock Return Predictability: Is It There?" *Review of Financial Studies*, 20 (2007), 651–707.
- Bekaert, G.; R. Hodrick; and D. Marshall. "On Biases in Tests of the Expectations Hypothesis of the Term Structure." *Journal of Financial Economics*, 44 (1997), 309–348.
- Berkowitz, J., and L. Giorgianni. "Long-Horizon Exchange Rate Predictability." *Review of Economics and Statistics*, 83 (1996), 81–91.
- Boudoukh, J., and M. Richardson. "The Statistics of Long-Horizon Regressions." *Mathematical Finance*, 4 (1994), 103–120.
- Boudoukh, J.; M. P. Richardson; and R. F. Whitelaw. "The Myth of Long-Horizon Predictability." Working Paper, New York University (2005).
- Campbell, J. Y., and R. Shiller. "The Dividend Ratio and Small Sample Bias." *Economics Letters*, 29 (1988), 325–331.
- Campbell, J. Y., and M. Yogo. "Efficient Tests of Stock Return Predictability." *Journal of Financial Economics*, 81 (2006), 27–60.
- Chan, C. K., and N. F. Chen. "An Unconditional Asset Pricing Test and the Role of Size as an Instrumental Variable for Risk." *Journal of Finance*, 43 (1988), 309–325.
- Chapman, D. A.; T. Simin; and H. Yan. "Stock Returns and Dividend Yields: Some New Evidence." Working Paper, Pennsylvania State University (2003).
- Christopherson, J. A.; W. Ferson; and D. Glassman. "Conditioning Manager Alpha on Economic Information: Another Look at the Persistence of Performance." *Review of Financial Studies*, 11 (1998), 111–142.
- Cochrane, J. H. "A Cross-Sectional Test of an Investment-Based Asset Pricing Model." *Journal of Political Economy*, 104 (1996), 572–621.
- Conrad, J.; M. Cooper; and G. Kaul. "Value Versus Glamour." *Journal of Finance*, 58 (2003), 1969–1995.
- Conrad, J., and G. Kaul. "Time-Variation in Expected Returns." *Journal of Business*, 61 (1988), 409–425.
- Fama, E. F.; L. Fisher; M. C. Jensen; and R. Roll. "The Adjustment of Stock Prices to New Information." *International Economic Review*, 10 (1969), 1–21.
- Fama, E. F., and K. R. French. "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy*, 96 (1988), 246–273.
- Fama, E. F., and K. R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.
- Fama, E. F., and K. R. French. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance*, 51 (1996), 55–87.
- Fama, E. F., and K. R. French. "Industry Cost of Equity." *Journal of Financial Economics*, 43 (1997), 153–194.
- Ferson, W. E., and C. R. Harvey. "The Fundamental Determinants of International Equity Returns: A Perspective on Country Risk and Asset Pricing." *Journal of Banking and Finance*, 21 (1997), 1625–1665.
- Ferson, W. E., and C. R. Harvey. "Conditioning Variables and the Cross-Section of Stock Returns." *Journal of Finance*, 54 (1999), 1325–1360.
- Ferson, W. E.; A. Heuson; and T. Su. "Weak and Semi-Strong Form Predictability Revisited." *Management Science*, 51 (2005), 1582–1592.
- Ferson, W. E., and R. Korajczyk. "Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?" *Journal of Business*, 68 (1995), 309–349.
- Ferson, W. E.; S. Sarkissian; and T. Simin. "The Alpha Factor Asset Pricing Model: A Parable." *Journal of Financial Markets*, 2 (1999), 49–68.
- Ferson, W. E.; S. Sarkissian; and T. Simin. "Spurious Regressions in Financial Economics?" *Journal of Finance*, 58 (2003), 1393–1414.
- Ferson, W. E., and R. Schadt. "Measuring Fund Strategy and Performance in Changing Economic Conditions." *Journal of Finance*, 51 (1996), 425–462.
- Foster, F. D.; T. Smith; and R. E. Whaley. "Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R-Squared." *Journal of Finance*, 52 (1997), 591–607.
- Gibbons, M. R., and W. E. Ferson. "Testing Asset Pricing Models with Changing Expectations and an Unobservable Market Portfolio." *Journal of Financial Economics*, 14 (1985), 216–236.
- Goetzmann, W., and P. Jorion. "Testing the Predictive Power of Dividend Yields." *Journal of Finance*, 48 (1993), 663–679.

- Granger, C. W. J.; N. Hyung; and Y. Jeon. "Spurious Regressions with Stationary Series." *Applied Economics*, 33 (2001), 899–904.
- Granger, C. W. J., and P. Newbold. "Spurious Regressions in Economics." *Journal of Econometrics*, 4 (1974), 111–120.
- Hastie, T.; R. Tibshirani; and J. Friedman. *The Elements of Statistical Learning*. New York, NY: Springer (2001).
- Huberman, G., and S. Kandel. "Market Efficiency and Value Line's Record." *Journal of Business*, 63 (1990), 187–216.
- Jagannathan, R., and Z. Wang. "The Conditional CAPM and the Cross-Section of Expected Returns." *Journal of Finance*, 51 (1996), 3–54.
- Jansson, M., and M. J. Moreira. "Optimal Inference in Regression Models with Nearly Integrated Regressors." *Econometrica*, 74 (2006), 681–714.
- Kim, M. J.; C. R. Nelson; and R. Startz. "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence." *Review of Financial Studies*, 58 (1991), 515–528.
- Lanne, M. "Testing the Predictability of Stock Returns." *Review of Economics and Statistics*, 84 (2002), 407–415.
- Leamer, E. E. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York, NY: John Wiley & Sons (1978).
- Lettau, M., and S. Ludvigson. "Consumption, Aggregate Wealth and Expected Stock Returns." *Journal of Finance*, 56 (2001), 815–849.
- Lewellen, J. "Predicting Returns with Financial Ratios." *Journal of Financial Economics*, 74 (2004), 209–235.
- Lo, A. W., and A. C. MacKinlay. "Stock Prices Do Not Follow Random Walks: Evidence From a New Specification Test." *Review of Financial Studies*, 1 (1988), 41–66.
- Lo, A. W., and A. C. MacKinlay. "Data Snooping in Tests of Financial Asset Pricing Models." *Review of Financial Studies*, 3 (1990), 431–467.
- Maddala, G. S. *Econometrics*. New York, NY: McGraw-Hill (1977).
- Marmol, F. "Spurious Regression Theory with Non-Stationary Fractionally Integrated Processes." *Journal of Econometrics*, 84 (1998), 233–250.
- Merton, R. C. "An Intertemporal Capital Asset Pricing Model." *Econometrica*, 41 (1973), 867–887.
- Nelson, C. R., and M. J. Kim. "Predictable Stock Returns: The Role of Small Sample Bias." *Journal of Finance*, 48 (1993), 641–661.
- Newey, W. K., and K. D. West. "A Simple, Positive Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.
- Petkova, R., and L. Zhang. "Is Value Riskier Than Growth?" *Journal of Financial Economics*, 78 (2005), 187–202.
- Paye, B. S. "Do Macroeconomic Variables Predict Aggregate Stock Market Volatility?" Working Paper, Rice University (2006).
- Phillips, P. C. B. "Understanding Spurious Regressions in Econometrics." *Journal of Econometrics*, 33 (1986), 311–340.
- Phillips, P. C. B. "New Tools for Understanding Spurious Regressions." *Econometrica*, 66 (1998), 1299–1326.
- Phillips, P. C. B. "Bootstrapping Spurious Regressions." Working Paper, Yale University (2001).
- Powell, J. G.; J. Shi; T. Smith; and R. E. Whaley. "Political Regimes, Business Cycles, Seasonalities, and Returns." Working Paper, Australian National University (2006).
- Shanken, J. "Intertemporal Asset Pricing: An Empirical Investigation." *Journal of Econometrics*, 45 (1990), 99–120.
- Sharpe, W. F. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance*, 19 (1964), 425–442.
- Stambaugh, R. S. "Predictive Regressions." *Journal of Financial Economics*, 54 (1999), 315–421.
- Torous, W.; R. Valkanov; and S. Yan. "On Predicting Stock Returns with Nearly Integrated Explanatory Variables." *Journal of Business*, 7 (2005), 937–966.
- Tsay W.-J., and C.-F. Chung. "The Spurious Regression of Fractionally Integrated Processes." *Journal of Econometrics*, 96 (2000), 155–182.
- Valkanov, R. "Long-Horizon Regressions: Theoretical Results and Applications." *Journal of Financial Economics*, 68 (2003), 201–232.
- White, H. "A Reality Check for Data Snooping." *Econometrica*, 68 (2000), 1097–1126.
- Yule, G. U. "Why Do We Sometimes Get Nonsense Correlations between Time Series? A Study in Sampling and the Nature of Time Series." *Journal of the Royal Statistical Society*, 89 (1926), 1–64.

