JETEM

Journal of Economic Theory and Econometrics

The Journal of the Korean Econometric Society

(계 량 경 제 학 보)

December 2009

Volume 20. Issue 4. pp.50-73

Test of Return Predictability: A New Two-Step Procedure*

Jinho Bae†

Abstract Predictive regressions are subject to finite sample bias. We propose a new twostep procedure to make correct inference for predictive regressions. Simulation results show that our procedure performs better than the existing two-step procedure in eliminating the bias and size distortion of the conventional t-test. We apply this procedure and find that dividend yield has lost predictive power for return since the early 1990s.

Keywords Predictive Regression; Bias Correction; Two-Step Estimation; Structural Break

JEL Classification G10

* The author wishes to thank three anonymous referees and Chang-Jin Kim for their helpful comments. This work was supported by the faculty research fund of Konkuk University in 2007.

† Assistant Professor, Department of Economics, Konkuk University, E-mail address: jinhbk@konkuk.ac.kr

Received March 25, 2009, Revised July 7, 2009, Accepted July 13, 2009

1. Introduction

Since Stambaugh (1986) and Mankiw and Shapiro (1986), the system of equations below has been widely employed in examining the ability of dividend yield to predict future stock returns:

$$y_t = \alpha + \beta x_{t-1} + u_t, \quad t = 1, \cdots, T, \tag{1}$$

$$x_t = \delta + \phi x_{t-1} + v_t, \quad |\phi| < 1, \tag{2}$$

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim i.i.d. \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{pmatrix} \right), \tag{3}$$

where y_t is the stock return, x_t is the dividend yield, and ρ is the correlation between u_t and v_t. [1]

Fama and French (1988) and many others report evidence of predictability based on the least squares estimation of equation (1). [2] However, $\hat{\beta}$, the least squares estimator of β, is subject to an upward bias in finite samples (see Stambaugh 1999):

$$E(\hat{\beta} - \beta) = \gamma E(\hat{\phi} - \phi) > 0, \tag{4}$$

where $\hat{\phi}$ is the least squares estimator of ϕ and $\gamma = \rho\frac{\sigma_u}{\sigma_v}$. The bias gives rise to severe size distortion of the conventional t-test and we are likely to erroneously reject no predictability. Nelson and Kim (1993), Stambaugh (1999), Amihud and Hurvich (2004), and

[1] Financial ratios such as the book-to-market ratio and the price-earnings ratio are also considered as potential predictive variables in the literature, but they are less commonly considered. Meanwhile, in Boudoukh et al. (2007), payout (measured as dividends plus repurchases) and net payout (measured as dividends plus repurchases minus issuances) yields are used instead of the dividend yield.

[2] This paper is concerned only with one-period ahead in-sample return predictability. Other studies focus on out-of-sample predictability (e.g., Campbell and Thompson 2008, Guo 2006) and long horizon predictability (e.g., Valkanov 2003, Boudoukh et al. 2005). For a recent survey, refer to Goyal and Welch (2008).

Lewellen (2004), among others, present various procedures for eliminating the bias. [3] In particular, Amihud and Hurvich (2004) provide a two-step procedure, where the first step is to estimate the ϕ parameter in equation (2) by the ordinary least squares principle and then correct for the bias with $\hat{\phi}$ using the approach proposed by Kendall (1954). The second step is to estimate the predictive regression of equation (1) using the residuals taken from the first step as an additional regressor. This additional regressor acts as a bias-correcting term by controlling the part of u_t that is correlated with v_t.

We propose a new two-step procedure that improves on that of Amihud and Hurvich (2004). As equation (4) suggests, an appropriate adjustment for the bias in $\hat{\phi}$ in the first step is crucial for correct inference of β. From this point of view, Kendall's (1954) approach used in Amihud and Hurvich (2004) leaves room for improvement. The approach has a nontrivial error in reducing the bias in $\hat{\phi}$, as documented by Stambaugh (1999). Furthermore, while the historical dividend yield is highly persistent, this problem is aggravated when ϕ is close to one. The new two-step procedure provides a remedy for this potential problem. In the first step, the new procedure makes use of the simulations-based methodology of MacKinnon and Smith (1998) which has better performance in reducing the bias in $\hat{\phi}$ than Kendall's (1954), especially when ϕ is close to one.

We use this new methodology to examine the predictability of historical stock return, with focus on how the forecasting relationship has changed during the early 1990s. Since

[3] High persistence in x_t is a source of bias in $\hat{\beta}$ but it is controversial whether or not x_t is stationary. Studies such as Stambaugh (1999), Amihud and Hurvich (2004), and Lewellen (2004) view x_t as highly persistent but stationary. This paper also takes this view, as indicated in equation (1). Alternatively, studies like Torous et al. (2004) and Campbell and Yogo (2006) allow x_t to follow an autoregression with a root within $1/T$-neighborhood of one where T denotes the sample size, which is supported by conventional unit root tests. From this local-to-unity viewpoint, they derive asymptotic distributions for $\hat{\beta}$.

Viceira (1996), the literature has documented instability in β and recently Paye and Timmermann (2006) find statistical evidence for U.S. data that a structural break in β occurred in the early 1990s. [4] Using two subsamples separated by the break point, they find that the predictive power of dividend yield has diminished since the break point. However, their coefficient estimates are not adjusted for the finite sample bias. In this paper, we apply the new bias-adjustment procedure to each of the two subintervals to more accurately assess how return predictability has changed since the break point.

This paper is organized as follows. Section 2 describes the two-step procedure and discusses Monte Carlo experiments. In Section 3, we discuss empirical results. Section 4 concludes the paper.

2. Two-Step Estimation Procedure of Predictive Regressions

In this section, we review Amihud and Hurvich's (2004) two-step procedure, discuss the new two-step procedure, and present Monte Carlo experiment results on the bias-adjustment performance of the procedure.

2.1. A Review of Amihud and Hurvich's (2004) Two-Step Procedure

As in Stambaugh (1999), under the normality assumption, u_t can be decomposed into two components, one correlated and one uncorrelated with v_t:

$$u_t = \gamma v_t + \epsilon_t, \tag{5}$$

[4] Ang and Bekaert (2007) also find that return predictability of both interest rate and dividend related variables have deteriorated in the second half of the 1990s. However, Lettau and van Nieuwerburgh (2008) find no evidence of changes in predictability when the predictive regression model takes into account the mean shifts in the predictor variables.

where $\epsilon_t \sim i.i.d.N(0, \sigma_\epsilon^2)$. Then, equation (1) can be rewritten as

$$y_t = \alpha + \beta x_{t-1} + \gamma v_t + \epsilon_t, \tag{6}$$

where ϵ_t has a zero mean conditional on $X_{-1} = [x_0 \cdots x_{T-1}]'$ and $V = [v_1 \cdots v_T]'$. This implies that if ϕ were known and thus v_t were observable, the regression would be standard and the resulting least squares estimator of β would be unbiased. In practice, however, v_t is not available since ϕ is unknown. A natural alternative is to use an estimate of v_t as a proxy. Hence, bias-correction can be achieved in two steps as given below.

Step 1: Run the least squares regression of equation (2) and then correct for the bias in $\hat{\phi}$.

For the bias-correction of $\hat{\phi}$, Amihud and Hurvich (2004) employ Kendall's (1954) approximation, i.e.,

$$E(\hat{\phi} - \phi) = -\frac{(1 + 3\phi)}{T} + O(T^{-2}), \tag{7}$$

which leads to a second-order biased-corrected estimator,

$$\check{\phi} = \hat{\phi} + \frac{1 + 3\hat{\phi}}{T} + \frac{3(1 + 3\hat{\phi})}{T^2}, \tag{8}$$

which can be used to compute the residuals

$$\check{v}_t = x_t - \check{\delta} - \check{\phi} x_{t-1}, \tag{9}$$

where, following Andrews (1993a), $\check{\delta}$ is obtained from the regression of $x_t - \check{\phi} x_{t-1}$ on a constant.

Step 2: Run the least squares regression of the predictive regression with the bias-adjusted residuals of v_t as a bias-correcting regressor, i.e.,

$$y_t = \alpha - \gamma(\delta - \check{\delta}) + (\beta - \gamma(\phi - \check{\phi}))x_{t-1} + \gamma \check{v}_t + \epsilon_t. \tag{10}$$

Since ϵ_t has zero expectation conditional on X_{-1} and $\check{V} = [\check{v}_1 \cdots \check{v}_T]'$, the resulting least squares estimator of β, referred to as $\check{\beta}$, is unbiased. The variance of $\check{\beta}$ is of the form:

$$Var(\check{\beta}) = \gamma^2 Var(\check{\phi}) + \sigma_\epsilon^2 (X'_{-1} \check{M} X_{-1})^{-1}, \tag{11}$$

where, following Pagan (1984), $\gamma^2 Var(\check{\phi})$ is the uncertainty associated with using the generated regressor $\check{v}_t$ and, as implied by equation (8),

$$Var(\check{\phi}) = (1 + \frac{3}{T} + \frac{9}{T^2})^2 Var(\hat{\phi}), \tag{12}$$

and $\check{M} = I_T - [i_T \ \ \check{V}]([i_T \ \ \check{V}]'[i_T \ \ \check{V}])^{-1}[i_T \ \ \check{V}]'$ where I_T is an identity matrix of order T, and i_T is a $T \times 1$ vector of ones.

2.2. A New Two-Step Procedure

The new two-step procedure differs from Amihud and Hurvich's (2004) in the first step. We employ the simulations-based methodology of MacKinnon and Smith (1998) instead of Kendall's (1954) approximation given in equation (7) for two reasons. First, Stambaugh (1999) points out that the approximation error $O(T^{-2})$ in equation (7) can be nontrivial. Second, MacKinnon and Smith (1998) document that the simulations-based methodology exhibits better performance in reducing the bias in than the approximation methodology, especially when the true value of ϕ is close to one. We briefly review the simulations-based methodology and then summarize our two-step procedure below.

The key element of the simulation-based bias-correction of $\hat{\phi}$ is to estimate the following bias function, which is linear in ϕ,

$$E(\hat{\phi} - \phi) = a + b\ \phi, \tag{13}$$

where a and b are the intercept and the slope of the bias function, respectively. [5] Estimators of these parameters, denoted $\tilde{a}$ and $\tilde{b}$, are obtained using computer simulations. Evaluating the bias function at two different points of ϕ and solving the simultaneous equations for ϕ yields an expression for $\tilde{\phi}$, the resulting bias-adjusted estimator of ϕ,

$$\tilde{\phi} = \frac{1}{1+\tilde{b}}(\hat{\phi} - \tilde{a}). \tag{14}$$

Details of the methodology are discussed in the Appendix.

Our new two-step procedure is summarized as below.

Step 1: Run the least squares regression of equation (2) and then correct for the bias in $\hat{\phi}$ using the simulation-based methodology. Then we obtain the bias-adjusted residuals as

$$\tilde{v}_t = x_t - \tilde{\delta} - \tilde{\phi} x_{t-1}, \tag{15}$$

which is used as a bias-correction term in the second step below. [6]

Step 2: Use least squares to estimate the predictive regression with $\tilde{v}_t$ as a bias-correcting regressor

$$y_t = \alpha - \gamma(\delta - \tilde{\delta}) + (\beta - \gamma(\phi - \tilde{\phi}))x_{t-1} + \gamma\tilde{v}_t + \epsilon_t, \tag{16}$$

The resulting least squares estimator of β, referred to as $\tilde{\beta}$, is unbiased and has a variance of the form:

$$Var(\tilde{\beta}) = \gamma^2 Var(\tilde{\phi}) + \sigma_\epsilon^2 (X'_{-1}\tilde{M}X_{-1})^{-1}, \tag{17}$$

[5] Andrews (1993a) documents the dependence of the bias in $\hat{\phi}$ on ϕ. Unaware of the exact functional form, MacKinnon and Smith (1998) also consider flat and nonlinear functional forms. They find that the linear bias function performs as well as the nonlinear one and is better than the flat one.

[6] $\tilde{\delta}$ is obtained in the same manner as $\check{\delta}$.

where, from equation (14),

$$Var(\tilde{\phi}) = \frac{1}{(1+\tilde{b})^2} Var(\hat{\phi}), \tag{18}$$

and $\tilde{M} = I_T - [i_T \;\; \tilde{V}]([i_T \;\; \tilde{V}]'[i_T \;\; \tilde{V}])^{-1}[i_T \;\; \tilde{V}]'$ where $\tilde{V} = [\tilde{v}_1, \cdots, \tilde{v}_T]'$. [7]

2.3. Monte Carlo Experiments

Here we carry out Monte Carlo experiments to assess the ability of the new two-step procedure to reduce the bias in $\hat{\beta}$ and size distortions of the conventional t-test. Each simulation generates 10,000 sets of artificial series of y and x, following equations (1)–(3). We set α=−0.3411, δ=−0.0033, σ_u=4.0861, σ_v=0.0426, and ρ=−0.9642 on the basis of the least squares regression of returns on the value-weighted NYSE stocks against lagged dividend yield. The values of σ_u, σ_v, and ρ imply γ=−92.484. β is set to zero to examine the empirical sizes of t-test. As dividend yield in the data is highly persistent, three values of ϕ, larger than 0.99, are considered: 0.999, 0.995, and 0.991. Four sample sizes are examined: T=50, 100, 250, and 600.

Table 1 reports the means and standard deviations of 10,000 estimates of β in Panel A and ϕ in Panel B. The column labeled *None* indicates regressions without bias-correction. *New* and *AH* signify the two-step procedures proposed in this paper and in Amihud and Hurvich (2004), respectively. As the results are similar with different values of ϕ and T, we discuss only the results for ϕ=0.995 and T=600.

<Insert Table 1 here.>

We find a nonnegligible bias in $\hat{\beta}$, 0.7811, as measured by the difference between the average given in the column labeled *None* and the true value β=0. The two-step

[7] A proof of unbiasedness and the derivation of equation (17) are available upon request.

procedure proposed in Amihud and Hurvich (2004) eliminates most of this bias but less successfully than our two-step procedure. The average of $\check{\beta}$ from *AH's* procedure is 0.1677 but three times larger than the average of $\tilde{\beta}$, 0.0548, from the our procedure. We find that this better performance in the bias-correction of $\hat{\beta}$ is mainly attributable to better performance in the bias-adjustment of $\hat{\phi}$. As shown in Panel B, the average of $\tilde{\phi}$, 0.9944, is closer to the true value of 0.995 than the average of $\check{\phi}$, 0.9932. Though not reported here, the γ estimates for the three procedures are almost identical.

Table 2 reports the empirical sizes of the t-test of β=0 against β> 0 at various significance levels. First of all, when the bias is not corrected for, the null hypothesis is rejected too often because of the large bias in $\hat{\beta}$. For example, for ϕ=0.995 and T=600, out of 10,000 trials, the null is rejected 43.5%, 59.0%, and 66.9% at the 1%, 5%, and 10% significance levels, respectively. In contrast, when our two-step bias-correction procedure is employed, most of the size distortion is eliminated. The null is rejected 1.2%, 5.3%, and 9.9% at the 1%, 5%, and 10% significance levels, respectively. This indicates that the sampling distribution of the $t-$statistic from our procedure approximates quite successfully the t distribution.

<Insert Table 2 here.>

In addition, the two-step procedure of this paper performs better than that of Amihud and Hurvich (2004) in eliminating the size distortion of the t test. For all values of ϕ and T, the $t-$statistic from *AH's* procedure tends to over-reject the null compared to that from our procedure. For example, for $\phi = 0.995$ and $T = 600$, empirical sizes are 1.6%, 6.8%, and 12.5% at the 1%, 5%, and 10% significance levels, respectively. Hence, these results, taken together, suggest that the two-step procedure of this study allows researchers to make more reliable inference than using that of Amihud and Hurvich (2004).

3. An Application

In this section, we apply the new two-step procedure to examine whether historical stock returns are predictable. The data used is monthly returns and dividend yield on the value-weighted portfolio of NYSE stocks. Returns are continuously compounded. Dividend yield for month t is measured as the natural logarithm of $(\frac{\mathrm{d_t}}{\mathrm{p_t}} \times 100)$, where d_t is the sum of the dividend payments during months t-11 through t and p_t is the value of the portfolio at the end of month t. All the data is taken from the Center for Research in Security Prices database. We examine the period from July 1952 to December 2005. The starting date is chosen following Paye and Timmermann (2006). Table 3 shows the descriptive statistics for the data over the entire period and the subperiods before and after December 1994, the break date for β suggested by Paye and Timmermann (2006). The two series are plotted in Figure 1.

<Insert Table 3 here.>

Given a structural change in the forecasting relationship found by Paye and Timmermann (2006), it would be interesting to explore how dividend yield's predictive power has changed across the break date. Hence, we test for return predictability for the pre- and post-break periods, using the new two-step procedure. For comparison purposes, we also carry out predictability tests, without adjustment for bias and adjusting for bias using the two-step procedure of Amihud and Hurvich (2004). Table 4 reports the results of the three regressions for each subperiod. The labels *None*, *AH*, and *New* indicate no bias adjustment, bias-adjustment using the two-step procedure of Amihud and Hurvich (2004) and the new two-step procedure, respectively. The main results are summarized as follows.

<Insert Table 4 here.>

First, when no bias-adjustment is made, there seems to be strong evidence of predictability for both periods. The p-value for the β estimate, given in the column labeled *None*, is 0.0006 for the pre-break period and 0.0010 for the post-break period, and the null hypothesis of no return predictability $\beta = 0$ is rejected at the 1% significance level for both sample periods. That there is a slight decline in the significance for the later period is consistent with Paye and Timmermann (2006).

Second, when the two-step procedure of Amihud and Hurvich (2004) is used for bias-adjustment, there is evidence of return predictability for both subperiods. The null of no return predictability is rejected at the 1% level for 1952.M7-1994.M11 and it is rejected at 5% level for 1994.M12–2005.M12.

Third, when bias-adjustment is made using our two-step procedure, evidence of predictability is found only for the pre-break period. For the period of 1952.M7-1994.M11, the p-value for $\tilde{\beta}$ is 0.0176, which indicates rejection of the null at the 5% level. In contrast, evidence of predictability disappears when we move to the period 1994.M12–2005.M12. The estimate of $\tilde{\beta}$ for the post-break period is not significant even at the 10% significance level with the associated p-value being 0.1475. This finding is in contrast with the result from the other two tests that dividend yield has predictive power for return for the post-WWII period. Henceforth, this finding suggests that 1) evidence of predictability for the post-break period found by Paye and Timmermann (2006) may be a consequence of failure to correct for the finite sample bias, 2) the procedure of Amihud and Hurvich (2004) overrejects the null hypothesis, which is consistent with the results from Monte Carlo experiments in Section 2.

Why has the dividend yield's predictive power has disappeared since the break? [8] Goyal and Welch (2003) provide an answer to the question. Based on the dividend ratio model of Campbell and Shiller (1988), Goyal and Welch (2003) analytically show that β can change when ϕ, the AR(1) coefficient in Equation (2), changes. This suggests that a sudden change in the persistence of dividend yield may explain the absence of return predictability in the 1990s. To examine this possibility, we conduct a structural break test for the dividend yield process of equation (2), after modifying the equation such that the parameters are allowed to change permanently at an unknown changepoint:

$$x_t = \delta_{S_t} + \phi_{S_t} x_{t-1} + v_t, \tag{19}$$

$$S_t = \begin{cases} 0, & \text{for } \; t = 1, 2, \cdots, \tau_S \\ 1, & \text{for } \; t = \tau_S + 1, \cdots, T, \end{cases} \tag{20}$$

where τ_S is the break point so that the AR(1) coefficient is ϕ_0 up until τ_S and it becomes ϕ_1 afterwards. The Sup Wald test of Andrews (1993b) gives strong evidence of a one-time shift in the dividend yield process: the Sup Wald statistic is 21.381 while the 1% critical value is 16.04. Interestingly, the estimated break point is November 1994, which almost coincides with the date of the break in the predictive regression, December 1994, found

[8] Paye and Timmermann (2006) view major changes in market sentiments and regime switches in monetary policies as responsible for the diminished predictability, but provide no empirical evidence.

by Paye and Timmermann (2006). [9]

4. Summary and Conclusions

Regressions of returns on lagged dividend yield are subject to the well-known finite sample bias. This paper presents a two-step estimation procedure to make correct inference of the regressions. Our Monte Carlo experiments show that the procedure proposed here is more effective in eliminating the bias and size distortions of the t-test than the two-step procedure of Amihud and Hurvich (2004). This new methodology is employed to investigate predictability in returns on the value-weighted portfolio of NYSE stocks. Two sample periods, 1952.M7-1994.M11 and 1994.M12–2005.M12, are examined, following the structural break found by Paye and Timmermann (2006). We find that dividend yield has predictive power of returns for the earlier period but none for the later period.

[9] See the first row of Table 6 of Paye and Timmermann (2006).

References

Amihud, Y. and C.M. Hurvich, 2004, Predictive Regressions: A Reduced-Bias Estimation Method, Journal of Financial and Quantitative Analysis 39, 813-841.

Andrews, D.W.K., 1993a, Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models, Econometrica 61, 139-165.

Andrews, D.W.K., 1993b, Test for Parameter Instability and Structural Change with Unknown Change Point, Econometrica 61, 821-856.

Ang, A. and G. Bekaert, 2007, Stock Return Predictability: Is it There? Review of Financial Studies 20, 651-707.

Boudoukh, J., R. Michaely, M.P. Richardson and M.R. Roberts, 2007, On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing, Journal of Finance 62, 877-915.

Boudoukh, J., M. Richardson and R.F. Whitelaw, 2008, The Myth of Long-Horizon Predictability, Review of Financial Studies 21, 1577-1605.

Campbell, J.Y. and R.J. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, Review of Financial Studies 1, 195-227.

Campbell, J.Y. and S.B. Thompson, 2008, Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average? Review of Financial Studies 21, 1509-1531.

Campbell, J.Y. and M. Yogo, 2006, Efficient Tests of Stock Return Predictability, Journal of Financial Economics 81, 27-60.

Fama, E. and K. French, 1988, Dividend Yields and Expected Stock Returns, Journal of Financial Economics 22, 3-25.

Goyal, A. and I. Welch, 2003, Predicting the Equity Premium with Dividend Ratios, Management Science 49, 639-654.

Goyal, A. and I. Welch, 2008, A Comprehensive Look at the Empirical Performance of the Equity Premium Prediction, Review of Financial Studies 21, 1455-1508.

Guo, H., 2006, On the Out-of-Sample Predictability of Stock Market Returns, Journal of Business 79, 645-670.

Kendall, M.G., 1954, Note on Bias in the Estimation of Autocorrelation, Biometrika 41, 403-404.

Lettau, M. and S. van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence, Review of Financial Studies 21, 1607-1652.

Lewellen, J., 2004, Predicting Returns with Financial Ratios, Journal of Financial Economics 74, 209-235.

MacKinnon, J.G. and A.A. Smith, 1998, Approximate Bias Correction in Econometrics, Journal of Econometrics 85, 205-230.

Mankiw, N.G. and M. Shapiro, 1986, Do We Reject Too Often: Small Sample Properties of Tests of Rational Expectations Models, Economics Letters 20, 139-145.

Nelson, C.R. and M. Kim, 1993, Predictable Stock Returns: The Role of Small Sample Bias, Journal of Finance 48, 641-661.

Pagan, A., 1984, Econometric Issues in the Analysis of Regressions with Generated Regressors, International Economic Review 25, 221-247.

Paye, B. and A. Timmermann, 2006, Instability of Return Prediction Models, Journal of Empirical Finance 13, 274-315.

Stambaugh, R., 1986, Bias in Regressions with Lagged Stochastic Regressors, Unpublished Manuscript, University of Chicago.

Stambaugh, R., 1999, Predictive Regressions, Journal of Financial Economics 54, 375-

421.

Torous, W., R. Valkanov and S. Yan, 2004, On Predicting Stock Returns with Nearly Integrated Explanatory Variables, Journal of Business 77, 937-966.

Valkanov, R., 2003, Long-Horizon Regressions: Theoretical Results and Applications, Journal of Financial Economics 68, 201-232.

Viceira, L., 1996, Testing for Structural Change in the Predictability of Asset Returns, Unpublished manuscript, Harvard University.

Appendix

This appendix presents details of bias-correction in $\hat{\phi}$ based on MacKinnon and Smith (1998). The linear bias function takes the form:

$$E(\hat{\phi} - \phi) = a + b\ \phi. \qquad (A-1)$$

The key to bias-correction is to acquire a and b. They can be estimated by two sets of simulations. Evaluating equation (A-1) at any two points, $\phi = \phi_1$ and $\phi = \phi_2$, we can solve for a and b. We generate N sets of artificial data at $\phi = \phi_1$ and run least squares for each data set, which gives us N estimates of ϕ, i.e., $\hat{\phi}^j, j = 1, \cdots, N$ where the superscript j indicates the j-th data set. Then, the bias is calculated as

$$bias_1 = \frac{1}{N}\sum_{j=1}^{N}(\hat{\phi}^j - \phi_1).$$

Since the bias is linear in the true value, we let

$$bias_1 = a + b\phi_1. \qquad (A-2)$$

Likewise, we generate another N sets of artificial data at $\phi = \phi_2$, and

$$bias_2 = a + b\phi_2. \qquad (A-3)$$

The bias-adjusted estimator $\tilde{\phi}$ is the solution to the equation:

$$\tilde{\phi} = \hat{\phi} - \tilde{a} - \tilde{b}\ \tilde{\phi},$$

where $\tilde{a}$ and $\tilde{b}$ denote the solutions to equations (A-2) and (A-3), and $\hat{\phi}$ is the initial estimate from the actual data. Solving for $\tilde{\phi}$ yields

$$\tilde{\phi} = \frac{1}{1+\tilde{b}}(\hat{\phi} - \tilde{a}).$$

Even though any two arbitrary points can be used for ϕ_1 and ϕ_2, MacKinnon and Smith (1998) recommend using the initial estimate $\hat{\phi}$ and the bias-adjusted estimate of ϕ under the assumption of a flat bias function:

$$E(\hat{\phi} - \phi) = c.$$

We denote the bias-adjusted estimate of ϕ as $\check{\phi}$. Thus, $\phi_1 = \hat{\phi}$ and $\phi_2 = \check{\phi}$. For bias-adjustment under the flat bias function, one set of simulations is enough. Since the bias is independent of the true value, it can be evaluated at any arbitrary point. MacKinnon and Smith (1998) recommend using the initial estimate $\hat{\phi}$. Thus, we generate N sets of artificial data at $\phi = \hat{\phi}$. Letting $\bar{\phi}$ be the average of N estimates of ϕ, the bias estimate $\check{c}$ is

$$\check{c} = \bar{\phi} - \hat{\phi}.$$

The bias-adjusted estimator $\check{\phi}$ is, then,

$$\check{\phi} = \hat{\phi} - \check{c} = \hat{\phi} - \bar{\phi} + \hat{\phi} = 2\hat{\phi} - \bar{\phi}.$$

Table 1. Monte Carlo Experiment Results: Bias Correction

ϕ		*None*	*AH*	*New*	*None*	*AH*	*New*
			Panel A: β				
			T = 50			**T = 100**	
0.999	Mean	9.1741	1.8920	1.8436	4.8020	1.1405	0.9796
	S.D.	7.6445	8.1150	7.3111	4.0210	4.1412	3.7179
0.995	Mean	9.2528	2.0060	1.8697	4.9220	1.2757	1.0071
	S.D.	7.7485	8.2261	7.5014	4.1975	4.3046	3.9770
0.991	Mean	9.2646	2.0397	1.7898	4.8389	1.1988	0.8624
	S.D.	7.7280	8.1991	7.5419	4.1457	4.2705	3.9767
			T = 250			**T = 600**	
0.999	Mean	1.9940	0.5213	0.4131	0.8334	0.2185	0.1551
	S.D.	1.6985	1.7184	1.5575	0.7231	0.7266	0.6689
0.995	Mean	1.9964	0.5290	0.3365	0.7811	0.1677	0.0548
	S.D.	1.7593	1.7796	1.6847	0.7893	0.7931	0.7858
0.991	Mean	1.9396	0.4756	0.2434	0.7394	0.1275	-0.0071
	S.D.	1.8462	1.8676	1.8210	0.8673	0.9002	0.8715

Table 1. (Cont'd)

ϕ		*None*	*AH*	*New*	*None*	*AH*	*New*
			Panel B: ϕ				
			T = 50			**T = 100**	
0.999	Mean	0.8999	0.9786	0.9791	0.9471	0.9867	0.9884
	S.D.	0.0801	0.0852	0.0762	0.0423	0.0437	0.0390
0.995	Mean	0.8944	0.9728	0.9743	0.9417	0.9812	0.9841
	S.D.	0.0813	0.0865	0.0783	0.0439	0.0453	0.0416
0.991	Mean	0.8909	0.9690	0.9717	0.9388	0.9782	0.9818
	S.D.	0.0804	0.0856	0.0782	0.0436	0.0449	0.0417
			T = 250			**T = 600**	
0.999	Mean	0.9775	0.9934	0.9946	0.9900	0.9966	0.9973
	S.D.	0.0179	0.0181	0.0163	0.0075	0.0077	0.0071
0.995	Mean	0.9734	0.9892	0.9913	0.9865	0.9932	0.9944
	S.D.	0.0185	0.0187	0.0177	0.0083	0.0084	0.0083
0.991	Mean	0.9700	0.9859	0.9884	0.9830	0.9896	0.9910
	S.D.	0.0195	0.0197	0.0192	0.0091	0.0092	0.0095

The simulation experiments assess the ability of the new two-step procedure to reduce the finite-sample bias in the OLS slope coefficient of the predictive regression. The data generating process is $y_t = \alpha + \beta x_{t-1} + u_t$, $x_t = \delta + \phi x_{t-1} + v_t$ where y is the return and x is the dividend yield. The entries are means and standard deviations, denoted as S.D., of the β (Panel A) and ϕ (Panel B) estimates from 10,000 Monte Carlo simulations. *None* indicates that no bias-adjustment is made, *AH* refers to the two-step procedure of Amihud and Hurvich (2004), and *New* signifies the new two-step procedure. The true parameters are chosen, based on the OLS regression of returns on the value-weighted NYSE stocks against lagged dividend yield over the period 1952.M7–2005.M12. β is set to zero, and three values of ϕ are considered as indicated in the table. Different sample sizes, T, are examined.

Table 2. Monte Carlo Experiment Results: Empirical Sizes

ϕ	*None*	*AH*	*New*	*None*	*AH*	*New*
		Panel A: 1% level				
		T = 50			**T = 100**	
0.999	0.6118	0.0152	0.0123	0.6397	0.0191	0.0142
0.995	0.5940	0.0167	0.0128	0.6099	0.0219	0.0174
0.991	0.5817	0.0147	0.0115	0.5722	0.0161	0.0122
		T = 250			**T = 600**	
0.999	0.6426	0.0217	0.0167	0.6160	0.0222	0.0173
0.995	0.5607	0.0197	0.0159	0.4353	0.0156	0.0123
0.991	0.4853	0.0184	0.0159	0.3641	0.0142	0.0120
		Panel B: 5% level				
		T = 50			**T = 100**	
0.999	0.7668	0.0722	0.0537	0.7756	0.0839	0.0565
0.995	0.7495	0.0742	0.0563	0.7580	0.0854	0.0636
0.991	0.7395	0.0686	0.0514	0.7257	0.0762	0.0567
		T = 250			**T = 600**	
0.999	0.7790	0.0881	0.0618	0.7584	0.0845	0.0608
0.995	0.7127	0.0773	0.0595	0.5899	0.0676	0.0530
0.991	0.6390	0.0726	0.0566	0.5004	0.0618	0.0526

Table 2. (Cont'd)

ϕ	*None*	*AH*	*New*	*None*	*AH*	*New*
			Panel C: 10% level			
		T = 50			**T = 100**	
0.999	0.8243	0.1328	0.0965	0.8319	0.1494	0.1049
0.995	0.8137	0.1359	0.1019	0.8193	0.1515	0.1124
0.991	0.8034	0.1320	0.0974	0.7919	0.1380	0.1024
		T = 250			**T = 600**	
0.999	0.8330	0.1594	0.1095	0.8146	0.1531	0.1073
0.995	0.7808	0.1465	0.1073	0.6691	0.1246	0.0989
0.991	0.7151	0.1365	0.1052	0.5816	0.1116	0.0962

The simulation experiments assess the empirical sizes of the t-tests for predictability. *None* indicates no bias-adjustment, and *AH* and *New*, respectively, refer to the two-step procedure of Amihud and Hurvich (2004) and that of this paper. The data generating process is $y_t = \alpha + \beta x_{t-1} + u_t$, $x_t = \delta + \phi x_{t-1} + v_t$ where y is the return and x is the dividend yield. The entries are the rejection rates of β=0, out of 10,000 Monte Carlo simulations, at the 1% (Panel A), 5% (Panel B), and 10% (Panel C) significance levels. The true parameters are chosen, based on the OLS regression of returns on the value-weighted NYSE stocks against lagged dividend yield over the period 1952.M7–2005.M12. β is set to zero, and three values of ϕ are considered as indicated in the table. Various sample sizes, T, are examined.

Table 3. Descriptive Statistics

Sample periods	*Variables*	*Mean*	*S.D.*	*Median*	*Autocorrelation*
1952.M7–2005.M12	Returns	0.986	4.060	1.223	0.040
	Div. Yield	1.227	0.370	1.250	0.992
1952.M7–1994.M11	Returns	0.974	4.099	1.184	0.041
	Div. Yield	1.362	0.250	1.339	0.984
1994.M12–2005.M12	Returns	1.040	3.897	1.309	0.034
	Div. Yield	0.630	0.171	0.637	0.946

This table presents summary statistics for monthly stock returns and dividend yield from July 1952 to December 2005. Stock returns are continuously compounded returns, expressed in percent, on the value-weighted portfolio of NYSE stocks. Dividend yield on the portfolio is the natural log of the sum of the dividend payments during a year divided by the value of the portfolio at the end of the year and multiplied by 100. All the data is taken from the CRSP database. The entries in the last column are first-order autocorrelations.

Table 4. Parameter Estimates

Parameters	*None*	*AH*	*New*
Panel A: Sample 1952.M7–1994.M11			
β	2.7051 (0.8279) [0.0006]	1.9949 (0.8325) [0.0085]	1.8589 (0.8800) [0.0176]
ϕ	0.9788 (0.0086)	0.9865 (0.0087)	0.9880 (0.0092)
ρ	-0.9639	-0.9639	-0.9639
$\frac{\sigma_u}{\sigma_v}$	95.828	95.830	95.827
Panel B: Sample 1994.M12–2005.M12			
β	5.8967 (1.8692) [0.0010]	3.1665 (1.9109) [0.0500]	2.2779 (2.1666) [0.1475]
ϕ	0.9497 (0.0195)	0.9793 (0.0200)	0.9889 (0.0228)
ρ	-0.9638	-0.9641	-0.9646
$\frac{\sigma_u}{\sigma_v}$	95.671	95.638	95.594

This table presents estimation results for the predictive regression $y_t = \alpha + \beta x_{t-1} + u_t$ and the AR(1) regression for dividend yield $x_t = \delta + \phi x_{t-1} + v_t$ for two periods, 1952.M7-1994.M11 and 1994.M12-2005.M12. Returns are on the value-weighted NYSE stocks. *None* reports the standard OLS estimates, and *AH* and *New* report the bias-adjusted estimates from the two-step procedures of Amihud and Hurvich (2004) and this paper, respectively. The numbers in parentheses and squared brackets indicate standard errors and p-values, respectively.