

THE MYTH OF LONG-HORIZON PREDICTABILITY

Jacob Boudoukh,^a Matthew Richardson^b and Robert F. Whitelaw^{b*}

This Version: October 26, 2005

* ^a Arison School of Business, IDC and NBER; ^b Stern School of Business, New York University and NBER. We would like to thank Jeff Wurgler and seminar participants at Yale University and New York University for helpful comments. Contact: Prof. M. Richardson, NYU, Stern School of Business, 44 W. 4th St., New York, NY 10012; mrichar0@stern.nyu.edu.

THE MYTH OF LONG-HORIZON PREDICTABILITY

Abstract

The prevailing view in finance is that the evidence for long-horizon stock return predictability is significantly stronger than that for short horizons. We show that for persistent regressors, a characteristic of most of the predictive variables used in the literature, the estimators are almost perfectly correlated across horizons under the null hypothesis of no predictability. For example, for the persistence levels of dividend yields, the analytical correlation is 99% between the 1- and 2-year horizon estimators and 94% between the 1- and 5-year horizons, due to the combined effects of overlapping returns and the persistence of the predictive variable. Common sampling error across equations leads to ordinary least squares coefficient estimates and R^2 s that are roughly proportional to the horizon under the null hypothesis. This is the precise pattern found in the data. The asymptotic theory is corroborated, and the analysis extended by extensive simulation evidence. We perform joint tests across horizons for a variety of explanatory variables, and provide an alternative view of the existing evidence.

I. Introduction

Over the last two decades, the finance literature has produced growing evidence of stock return predictability, though not without substantive debate. The strongest evidence cited so far comes from long-horizon stock returns regressed on variables such as dividend yields, term structure slopes, and credit spreads, among others. A typical view is expressed in Campbell, Lo, and MacKinlay's standard empirical textbook for financial economics, *The Econometrics of Financial Markets* (1997, p.268):

At a horizon of one month, the regression results are rather unimpressive: The R^2 statistics never exceed 2%, and the t-statistics exceed 2 only in the post-World War II subsample. The striking fact about the table is how much stronger the results become when one increases the horizon. At a two-year horizon the R^2 statistic is 14% for the full sample ... at a four-year horizon the R^2 statistic is 26% for the full sample.

However, there is an alternative interpretation of this evidence: Researchers should be equally impressed by the short- and long-horizon evidence for the simple reason that the regressions are almost perfectly correlated. For an autocorrelation of 0.953 for annual dividend yields, we show analytically that the 1-year and 2-year predictive estimators are 98.8% correlated under the null hypothesis of no predictability. For longer horizons, the correlations are even higher, reaching 99.6% between the 4- and 5-year horizon estimators. This degree of correlation manifests itself in multiple-horizon regressions in a particularly unfortunate way.

Since the sampling error that is almost surely present in small samples shows up in each regression, both the estimator and R^2 are proportional to the horizon.

This paper provides analytical expressions for the correlations across multiple-horizon estimators, and then shows, through simulations, that these expressions are relevant in small samples. The analytical expressions relate the correlations across these estimators to both the degree of overlap across the horizons and the level of persistence of the predictive variable. Our findings relate to an earlier literature looking at joint tests of the random walk hypothesis for stock prices using multiple-horizon variance ratios and autocorrelations, among other estimators (see, e.g., Richardson and Smith (1991, 1994) and Richardson (1993)). This earlier literature stresses accounting for the degree of overlap. The problem here is much more severe. In the previous univariate framework, the predictive variable—past stock returns—is approximately independently and identically distributed (IID). In this paper’s framework, the predictive variable, such as dividend yields, is highly persistent.

Our simulations show that any sampling error in the data under the null hypothesis of no predictability appears in the same manner in every multiple-horizon regression when the predictive variable is highly persistent. Using box plots and tables describing the relation across the multiple horizon estimates and R^2 s, we show the exact pattern one should expect under the null hypothesis: The multiple-horizon estimates are monotonic in the horizon approximately two-thirds of the time, and the mean ratios of the 2- to 5-year estimators to the 1-year estimator are 1.93, 2.78, 3.56, and 4.28, respectively. Consider the actual estimated coefficients for the regression of 1- to 5-year stock returns on dividend yields over the 1926–2004 sample period: 0.117, 0.236, 0.355, 0.424, and 0.469. These correspond to monotonically increasing estimates with corresponding ratios of 2.01, 3.03, 3.62, and 4.01. We show that

these estimates lie in the middle of the distribution of possible outcomes under the null hypothesis.

The theoretical and simulation analyses stress the importance of interpreting the evidence jointly across horizons. We develop an analytical expression for a joint test based on the Wald statistic. While the level of high persistence implies that it can be dangerous to document or interpret regressions over multiple horizons, the joint tests show that this persistence may lead to powerful tests for economies in which predictability exists. Such predictability may take a particular form, in which the multiple-horizon coefficients are much less tied together than the null hypothesis implies. Applying the joint tests to commonly used predictive variables, we point out various anomalies and show a contrast to the conclusions of the existing literature. For example, we find that the book-to-market ratio generates some of the strongest evidence of predictability. Interestingly, this variable is often cited as having the greatest cross-sectional explanatory power. It is well known that the factors in a multi-factor capital asset pricing model (CAPM) that capture changes in the investment opportunity set should both explain the cross-section of stock returns and predict future returns. As an additional example, joint tests confirm the importance of the net payout yield for forecasting stock returns across all horizons, consistent with the recent finding of Boudoukh, Michaely, Richardson, and Roberts (2005) for short-horizon stock returns (see also Baker and Wurgler (2000) for related work using the market share of equity issues).

The paper proceeds as follows. In Section 2, we provide the expressions for analyzing multiple-horizon regressions and show that the basic findings carry through to small samples. The small sample results are especially alarming in the context of the existing literature.

Section 3 applies the results to a number of data series and evaluates existing evidence using joint tests of predictability. Section 4 concludes.

II. Multiple Horizon Regressions

A. The Existing Literature

Fama and French (1988) is the first paper to document evidence of multivariate stock return predictability over multiple horizons.¹ In brief, they regress overlapping stock returns of one month to four years on dividend yields, reporting coefficients and R^2 s that increase somewhat proportionately with the horizon. As it documents what has become one of the dominant stylized facts in empirical finance, the paper has over 250 citations to date. To illustrate the consensus view, consider John Cochrane's description of the three most important facts in finance in his survey, "New Facts in Finance" (1999, p.37).

Now, we know that ...

[Fact] 2. Returns are predictable. In particular: Variables including the dividend/price (d/p) ratio and term premium can predict substantial amounts of stock return variation. This phenomenon occurs over business cycle and longer horizons. Daily, weekly, and monthly stock returns are still close to unpredictable...

This fact is emphasized repeatedly in other surveys (see, e.g., Fama (1998, p.1578), Campbell (2000, p.1522 and 2003, p.5) and Barberis and Thaler (2003, p.21), among others),

¹ See also Campbell and Shiller (1988).

is often used to calibrate theoretical models (see, among others, Campbell and Cochrane (1999, p.206), Campbell and Viceira (1999, p.434), Barberis (2000, p.225), Menzly, Santos, and Veronesi (2004, p.2), and Lettau and Ludvigson (2005, p.584)), and in motivation for new empirical tests (e.g., Ferson & Korajczyk (1995, p.309), Patelis (1997, p.1951), Lettau & Ludvigson (2001, p.815), and Ait-Sahalia and Brandt (2001,p.1297)).

It is fairly well known since Fama and French (1988), and in particular from Campbell (2001), that the key determinants of long-horizon predictability are

- (i) The extent of predictability at short horizons, and
- (ii) The persistence of the regressor.

The R^2 s at long horizons relative to a single-period R^2 are a function of (ii). Holding everything else constant—single-period predictability in particular—higher persistence results in a higher fraction of explainable long-horizon returns. As a function of the horizon, the R^2 first rises with the horizon, but eventually decays, due to the exponential decline in the informativeness of the predictive variable. As we show below, (ii) also matters in the case of no predictability, but in the presence of sampling error. Nevertheless, this important fact has not been used as the main line of attack against evidence supporting the multivariate predictability of stock returns.

Three principal alternative arguments, however, have been put forward in the literature. The first involves data snooping, which is perhaps best described by Foster, Smith, and Whaley (1997). The idea is that the levels of predictability found at short horizons are not surprising, given the number of variables from which researchers can choose. A variety of papers support these findings somewhat, including Bossaerts and Hillion (1999), Cremers (2002), and Goyal and Welch (2003).

A second approach looks at the small sample biases of the estimators. Stambaugh (1999) shows that the bias can be quite severe, given the negative correlation between contemporaneous shocks to returns and the predictive variable, which usually involves some type of stock price deflator. His findings suggest much less predictability once the estimators are adjusted for this bias. However, Lewellen (2004) argues that the effect of the bias may be much smaller if one takes the persistence of the predictive variable into account. Lewellen's approach is similar to Stambaugh's (1999) Bayesian analysis of the predictability problem. While both of these papers certainly question the magnitude of the predictability, they do not address long-horizon predictability *per se*.

The third line of criticism, first explored by Richardson and Stock (1989) in a univariate setting, uses an alternative asymptotic theory, in which the horizon increases with the sample size. Valkanov (2003) argues that long-horizon regressions have poor properties relative to standard asymptotics. He shows that the estimators may no longer be consistent, and have limiting distributions that are functionals of Brownian motions; in fact, the distributions are not normal, and are not centered around the true coefficient. Valkanov then shows that this alternative asymptotic theory works better in small samples. His results can be viewed as the theoretical foundation for earlier simulated distributions by Kim and Nelson (1993) and Goetzmann and Jorion (1993), and for the intuition put forward by Kirby (1997), who uses standard asymptotics.

These three criticisms of the stock return predictability literature are valid and important. Our paper focuses on a different aspect of predictability, examining the joint properties of the regression estimators across horizons. The conclusions here closely resemble those of Richardson and Smith (1991) and Richardson (1993) regarding long-horizon evidence against

the random walk in Fama and French (1998) and Poterba and Summers (1988). In many ways, the arguments here are more damaging because we show that the degree of correlation across the multiple-horizon estimators is much higher than it is in the case of long horizon tests for the random walk. In fact, the null hypothesis of no predictability implies the exact pattern in coefficients and R^2 s found in papers presenting evidence in favor of predictability. We show these results in the next two subsections.

B. Multiple Horizon Regressions: Statistical Properties

We consider regression systems of the following type:

$$\begin{aligned}
 R_{t,t+1} &= \alpha_1 + \beta_1 X_t + \varepsilon_{t,t+1} \\
 &\vdots \\
 R_{t,t+J} &= \alpha_J + \beta_J X_t + \varepsilon_{t,t+J} \\
 &\vdots \\
 R_{t,t+K} &= \alpha_K + \beta_K X_t + \varepsilon_{t,t+K},
 \end{aligned} \tag{1}$$

where $R_{t,t+J}$ is the J -period stock return, X_t is the predictor, e.g., the dividend yield, and $\varepsilon_{t,t+J}$ is the error term over J periods. As is well known from Hansen and Hodrick (1980) and Hansen (1982), among others, the error terms are serially correlated due to overlapping observations. Using the standard generalized method of moments calculations under the null hypothesis of no predictability and under conditional homoskedasticity (e.g., Richardson and Smith (1991)), Appendix A derives the covariance matrix for any two horizons, J and J^* , for $\hat{\beta}_J$ and $\hat{\beta}_{J^*}$:

$$T\text{Var}(\hat{\beta}_J, \hat{\beta}_{J^*}) = \frac{\sigma^2}{\sigma_X^2} \begin{pmatrix} J+2 \sum_{l=1}^{J-1} \binom{J-1}{J-l} \rho_l & J + \left[\sum_{l=1}^{J-1} \binom{J-1}{J-l} (\rho_l + \rho_{l+(J^*-J)}) \right] + \sum_{l=1}^{J^*-J} J \rho_l \\ \dots & J^*+2 \sum_{l=1}^{J^*-1} \binom{J^*-1}{J^*-l} \rho_l \end{pmatrix}, \quad (2)$$

where $J^* > J$ and ρ_l is the l^{th} -order autocorrelation of X_t . The above expression for the covariance matrix of the estimators is not particularly intuitive, though it is immediately apparent that, for J close to J^* , the estimators are almost perfectly correlated. Less obvious is the fact that, for $\text{cov}(X_t, X_{t-1}) \approx \sigma_X^2$, the estimators are also almost perfectly correlated irrespective of the horizon. Intuitively, the persistence of X_t acts in much the same way overlapping horizons do regarding independent information across multiple horizons.

A popular simplification is to assume that X_t follows an AR(1) (see, among others, Campbell (2001), Boudoukh and Richardson (1994), Stambaugh (1993), and Cochrane (2001)). Under the AR(1) model, $\text{cov}(X_t, X_{t-1}) = \rho \sigma_X^2 \approx \rho^l \sigma_X^2$ where ρ is the autoregressive parameter for X_t . The covariance matrix in (2) reduces to a much simpler form:

$$T\text{Var}(\hat{\beta}_J, \hat{\beta}_{J^*}) = \frac{\sigma_R^2}{\sigma_X^2} \begin{pmatrix} J + \frac{2\rho}{(1-\rho)^2} [(J-1) - \rho(J-\rho^{J-1})] & J + \frac{\rho}{(1-\rho)^2} \{2[(J-1) - \rho(J-\rho^{J-1})] + (1-\rho^J)(1-\rho^{J^*-J})\} \\ \dots & J^* + \frac{2\rho}{(1-\rho)^2} [(J^*-1) - \rho(J^*-\rho^{J^*-1})] \end{pmatrix}. \quad (3)$$

For the special case $J = 1$, the correlation between J and J^* is

$$\frac{(1-\rho)^2 + \rho(1-\rho)(1-\rho^{J^*-1})}{\sqrt{(1-\rho)^2} \sqrt{J^*(1-\rho)^2 + 2\rho[(J^*-1) - \rho(J^*-\rho^{J^*-1})]}}. \quad (4)$$

For example, for $J = 1$ and $J^* = 2$, we get $\sqrt{\frac{1+\rho}{2}}$. In our sample, the autocorrelation of the dividend yield is 0.953, which yields a correlation of 0.988 between the 1- and 2-year estimators. It does not get much better as the horizon J^* increases to 3, 4, and 5 years,

producing correlations of 0.974, 0.959, and 0.945, respectively.² Even at a 10-year horizon, the correlation is over 87%. With the types of sample sizes faced by researchers in the field of empirical finance, these results suggest that one has to be extremely cautious in interpreting the coefficients separately (as has been the case in the existing literature).

For the typical 1- through 5-year horizons examined in the literature, we provide the analytical covariance matrix of the estimators under the null hypothesis of no predictability and the dividend yield's ρ of 0.953:

$$T\text{Var}(\hat{\beta}_{1-5}) = \frac{\sigma_R^2}{\sigma_X^2} \begin{pmatrix} 1 & 0.988 & 0.974 & 0.959 & 0.945 \\ & 1 & 0.993 & 0.982 & 0.970 \\ & & 1 & 0.995 & 0.986 \\ & & & 1 & 0.996 \\ & & & & 1 \end{pmatrix}. \quad (5)$$

Several observations are in order. First, the high degree of correlation across the multi-period estimators implies that, under the null hypothesis, the regressions are essentially redundant. This is important because there is little doubt that the literature has not taken this view. Second, under the null hypothesis, the estimators are asymptotically distributed as multivariate normal with a mean of zero. While this is clearly not true in small samples,³ consider using this distribution to understand the effect of sampling error across the equations. Specifically, conditional on $\hat{\beta}_1 = \bar{\beta}_1$, what do we expect $\hat{\beta}_{J^*}$ to equal? Using properties of a bivariate normal, we can write

$$E[\hat{\beta}_{J^*} | \hat{\beta}_1 = \bar{\beta}_1] = \left(1 + \frac{\rho(1 - \rho^{J^* - 1})}{1 - \rho} \right) \bar{\beta}_1. \quad (6)$$

² Of course, these correlations are lower bounds, as the magnitude increases as J increases for a fixed J*.

³ See Stambaugh (1999) for small sample bias and Valkanov (2003) for non-normality of the distributions of the estimators.

For ρ close to 1, the coefficients should basically be proportional to the horizon. As an example, for $\rho=0.953$, $\hat{\beta}_{J^*}$ is $1.953 \hat{\beta}_1$, $2.861 \hat{\beta}_1$, $3.727 \hat{\beta}_1$, and $4.552 \hat{\beta}_1$ for the 2-, 3-, 4-, and 5-year horizon regressions, respectively. Similarly, for the R^2 of the regression,

$$E[R_{J^*}^2 | R_1^2 = \bar{R}_1^2] = \frac{\left(1 + \frac{\rho(1 - \rho^{J^*-1})}{1 - \rho}\right)^2}{J^*} \bar{R}_1^2. \quad (7)$$

For ρ close to 1, the R^2 s also increase significantly with the horizon. The ratios of the R^2 s are 1.907, 2.729, 3.472, and 4.143 for the 2-, 3-, 4-, and 5-year horizon regressions, respectively.

The intuition is straightforward. Compare the regression of $R_{t,t+1}$ on X_t to that of $R_{t,t+K}$ on X_t . The former regression involves summing the cross product of the sequence of $R_{t,t+1}$ and X_t for all t observations. Note that for a persistent series X_t , there is very little information across the sequence of X_t values. Thus, when an unusual draw from $R_{t,t+1}$ occurs (denote it R_{t^*,t^*+1}), and this observation happens to coincide with the most recent value of the predictive variable, X_{t^*} , it will also coincide with all the surrounding X_t variables such as X_{t^*-1} , X_{t^*-2} , and X_{t^*-3} . Since R_{t^*,t^*+1} shows up K times in $R_{t,t+K}$ within the sample period (i.e., R_{t^*+1-K,t^*+1} , R_{t^*+2-K,t^*+2} , \dots , R_{t^*,t^*+K}), the impact of the unusual draw of R_{t^*,t^*+1} for $R_{t,t+K}$ will be roughly K times than that of $R_{t,t+1}$.

C. Multiple Horizon Regressions: Joint Tests

At first glance, the results in II.B provide a fairly devastating critique of the strategy of running multiple long-horizon regressions. However, this view is not necessarily accurate.

Because the regressions are linked so closely under the null hypothesis, joint tests may have considerable power under alternative models.

What are these alternatives? The models must be such that the long horizons pick up information not contained in short horizons. The standard model, in which short-horizon returns are linear in the current predictor and that predictor follows an ARMA process, is clearly not a good candidate. It would be better to focus on estimating the short horizon and the ARMA process directly in this case (see, e.g., Campbell (2001), Hodrick (1992), and Boudoukh and Richardson (1994), among others). It should be noted, though, that the standard model is often chosen for reasons of parsimony rather than on an underlying theoretical basis.

Consider testing the null of no predictability in the regression system given in equation (1), i.e., $\beta_1 = \dots = \beta_J = \dots = \beta_K = 0$. The corresponding Wald Test statistic for this hypothesis is $T\hat{\beta}'V(\hat{B})^{-1}\hat{B}$ where $\hat{\beta}' = (\hat{\beta}_1 \dots \hat{\beta}_J \dots \hat{\beta}_K)$ and $V(\hat{B})$ is the variance-covariance matrix of the \hat{B} estimators with typical element of \hat{B}_J and \hat{B}_{J^*} given by $\text{Var}(\hat{\beta}_J, \hat{\beta}_{J^*})$ as described in equation (2). The statistic follows an asymptotic chi-squared distribution with degrees of freedom given by the number of multiple horizons used in estimation. Note that $V(\hat{B})$ is a function of the autocorrelation structure of the X_t variable (i.e., its persistence) as well as the degree of overlap between horizons, i.e., J versus J^* . Aside from the magnitude of the \hat{B} estimators, what matters is whether the pattern in \hat{B} across horizons is consistent with the correlation implied by $V(\hat{B})$.

To see this, consider performing a Wald Test of the hypothesis $\beta_1 = \beta_2 = 0$. It is possible to show that the corresponding Wald statistic is given by $T \frac{\sigma_X^2}{\sigma_R^2} \left[\frac{2\beta_1^2 + \beta_2^2 / (1+\rho) - 2\beta_1\beta_2}{1-\rho} \right]$. For a given

sample size T and estimated coefficient $\hat{\beta}_1 = \bar{\beta}_1$, it is possible to show that this statistic is minimized at $\hat{\beta}_2 = (1 + \rho)\bar{\beta}_1$. Since a low value of the statistic implies less evidence against the null, this result means that we not only expect a nonzero $\hat{\beta}_2$ under the null but that it should be of a magnitude greater than the $\hat{\beta}_1$ estimate. In fact, for a highly persistent regressor, the Wald statistic is minimized when the 2-period coefficient is almost double the one-period coefficient.

These results provide important clues in searching for powerful tests against the null of no predictability. If the alternative hypothesis does not imply coefficient estimates that increase at the same rate across horizons or that are not as heavily tied to the predictive variable's persistence, one can find evidence of predictability even with modestly sized coefficients. But the fact that the no predictability null and standard ARMA predictive model imply similar coefficients patterns (and thus low power) does not mean the null is false. On the contrary, it explains why the literature shifted away from the no predictability paradigm.

This is an important point. Treating the individual coefficient estimates separately in a joint setting can lead to very misleading conclusions. The null hypothesis of no predictability as described by the Wald Test is most supported in the data when we observe monotonically increasing/decreasing coefficient estimates that can be described by the horizon and persistence of the predictive variable. This is the exact pattern documented in the original Fama and French (1988) and Campbell and Shiller (1988) papers. It is worth wondering how the finance literature may have treated these papers if armed with this fact. This is especially true given the weak evidence of predictability at short horizons, especially in the context of the previously mentioned data snooping arguments (e.g., Foster, Smith, and Whaley (1997) and

small sample bias (Stambaugh (1999)), both of which tend to overstate short horizon significance.

D. Multiple Horizon Regressions: Simulation Evidence

The theoretical results in II.B and II.C are based on asymptotic properties of fixed-horizon estimators. A priori there is reason to be wary of these results in small samples, particularly because of the considerable evidence of a small sample bias in the coefficient estimators and of non-normality in small samples. As mentioned in Section II.A, Stambaugh (1999) derives a small sample bias of potentially large magnitude. Therefore, it would be useful to evaluate some of the above conclusions in the context of this bias. In addition, Valkanov (2003) shows that, under alternative asymptotics, in which the horizon length increases with the sample size, the estimators are no longer consistent and can have limiting distributions that deviate significantly from normality. This alternative asymptotic theory seems to behave better in small samples than the fixed-horizon one, at least for long-horizon estimators. As a result, in this section, we look at the small sample properties of the estimators in general, and the patterns in sampling error across equations in particular. Previewing the results to come, the basic tenet of equations (2) and (3), namely, the dependence across equations, carries through to small samples.

We simulate the model in equation (1) under the assumption of no predictability, an AR(1) process on X_t , and 75 years of annual data. The analysis is performed over 1- to 5-year horizons with the AR parameter ρ , the standard deviation of X_t and $\varepsilon_{t,t+1}$, and the correlation

between $\varepsilon_{t,t+1}$ and $u_{t,t+1} \equiv (X_{t+1} - \rho X_t)$ chosen to match the data.⁴ The simulations involve 100,000 replications each.

Table 1A reports the simulated 5x5 correlation matrix of the multiple horizon estimators. Consistent with the theoretical analytical calculations in Section II.B, the correlations tend to be high, even for the most distant horizons. The simulated correlations between the 1-year and 2- to 5-year horizon estimators are 0.966, 0.925, 0.883, and 0.841, respectively, showing that the correlation calculations under the fixed-horizon asymptotics in II.B hold in small samples. Thus, the estimators' almost perfect cross-correlation leads to little independent information across equations, and the sampling error that is surely present in small samples shows up in every equation in (1).

As shown in II.B, the persistence of ρ is an important determinant of the magnitude of the correlation matrix of the multiple-horizon estimators. Figure 1A graphs the correlation between the 1-year and 2- to 5-year horizon estimators as a function of ρ , in particular, for values $\rho = 0.953, 0.750, 0.500, 0.250,$ and 0.000 . Consistent with the theory in II.B, the correlations decrease as ρ falls. The drop-off can be quite large as the horizon increases. As a function of the above ρ values, the 1- and 2-year estimators have correlations of 0.966, 0.917, 0.849, 0.776, and 0.698, respectively. In contrast, the 1- and 5-year estimators have correlations of 0.841, 0.684, 0.545, 0.465, and 0.429. Even when the predictive variable has no persistence, the correlation can still be quite high due to the overlapping information across the multiple-horizon returns.

⁴ Specifically, for the regression of annual stock returns on the most commonly used predictive variable, namely, dividend yields, we estimate $\rho = 0.953, \sigma_\varepsilon = 0.196, \sigma_u = 0.166$ and $\sigma_{\varepsilon u} = -0.683$. While the magnitudes of

However, the staggering result in Table 1A is that 65.2% of all the replications produce monotonic estimates as a function of the horizon. That is, almost two-thirds of the time, the data produce coefficients increasing or decreasing with the horizon. This result coincides with the predictions from the asymptotic theory given in II.B. To understand how unlikely monotonicity is, suppose that the five different multiple-horizon estimators were IID. In this setting, the probability of a monotonic relation is 0.83%, approximately $1/78^{\text{th}}$ of the true probability for the multiple-horizon estimators. Even with overlapping horizons, monotonicity drops sharply as ρ falls, i.e., from the 65.2% value to 36.7%, 19.7%, 10.8%, and 6.0% for $\rho = 0.953, 0.750, 0.500, 0.250,$ and $0.000,$ respectively. This result further highlights the importance of persistence in the predictive variable for generating these patterns.

One possible explanation for this finding is that the small sample bias increases with horizon (e.g., Stambaugh (1999), Goetzmann and Jorion (1993), and Kim and Nelson (1993)). Table 1A confirms that the small sample bias increases with the horizon, with the means of the 1- to 5-year coefficients equal to 0.048, 0.092, 0.132, 0.169, and 0.203, respectively. To investigate whether the monotonicity is due to this bias, Table 1B duplicates Table 1A under the assumption that $\sigma_{\varepsilon u} = 0$. For this value, the small sample bias is theoretically zero, and the estimates are unbiased in our simulations. Interestingly, the monotonicity falls only slightly, to 57.2%. Furthermore, Table 1B shows that the correlation matrix across the multiple-horizon estimators is virtually identical to that in Table 1A. Thus, the monotonicity is driven by the almost perfect correlation across the estimators and the increasing horizon.

σ_u and σ_ε do not matter, this is not true for either the persistence variable ρ (Boudoukh and Richardson (1994)) or the correlation $\sigma_{\varepsilon u}$ (e.g., Stambaugh (1999)). Thus, we also investigate different values for these parameters.

As described in Section II.A, much of the literature has argued for predictability by focusing on the increase in the coefficient estimates as a function of the horizon. Both theoretically and in simulation, we show that this result is expected under the null hypothesis of no predictability. An alternative measure of predictability also considered in the literature is the magnitude and pattern of R^2 s across horizons. While the R^2 is linked directly to the coefficient estimate, it is nonetheless a different statistic of the data. Table 2A reports the simulated 5x5 correlation matrix of the multiple horizon R^2 s as well as their means, medians, standard deviations, and monotonicity properties.

Similar to Table 1A, the R^2 s are all highly correlated across horizons. For example, the simulated correlations between the 1-year and 2- to 5-year horizon R^2 s are 0.948, 0.886, 0.823, and 0.762, respectively. This degree of correlation leads to R^2 s that are monotonic in the horizon 51.3% of the time under the null hypothesis—the exact pattern documented in the literature. This result is not due to the Stambaugh (1999) small sample bias, as both the degree of correlation and monotonicity also appear in the simulated data without the bias (see Table 2B, where the cross-equation correlation is zero). Also, analogous to the evidence for the multiple-horizon coefficient estimators, the degree of cross-correlation and monotonicity depends crucially on the level of persistence ρ of the predictive variable.

Figures 1A and 1B show the correlation coefficients between the 1- and the J -period betas and R^2 s. The correlations are plotted for different persistence parameters, and the figures illustrate both the monotonicity and near linearity one would expect and the dependence of this effect on the persistence parameter.

The theoretical calculations of Section II.B imply an even stronger condition than monotonicity. For ρ close to 1, the coefficients and R^2 s should increase one-for-one with the

horizon under the null hypothesis. Because this is the typical pattern found in US data, it seems worthwhile to investigate this implication through a simulation under the null hypothesis of no predictability. We compare the ratio of the 2- to 5-year coefficient and R^2 estimates to the 1-year estimates. Since there are numerical issues when using denominators close to zero, we run the analysis under the condition that the 1-year estimate have an absolute value greater than 0.01, or an R^2 greater than 0.5%. Approximately 87% and 61% of the simulations respectively satisfy these criteria.

Table 3A contains the results. As predicted by the theory, the mean ratios of the estimates are 1.93, 2.78, 3.56, and 4.28 for the 2-, 3-, 4-, and 5-year horizons, respectively. The R^2 s are equally dramatic, with corresponding ratios of 1.95, 2.87, 3.75, and 4.59.⁵ Note that these simulations are performed under the null hypothesis of no predictability. The β s are zero, but the other parameters are calibrated to match the joint distribution of stock returns and dividend yields in the data. How do these results compare with the estimated coefficients and R^2 s from the actual data? In the data, the ratios for the 2-, 3-, 4-, and 5-year horizons are 2.01, 3.03, 3.62, and 4.00 for the β estimates, and 1.86, 3.01, 3.44, and 3.73 for the R^2 s. The similarities are startling.

Figures 2A and 2B plot the ratios as a function of the persistent parameter ρ . For large ρ , both the coefficient estimates and R^2 s increase linearly with the horizon, with fairly steep slopes (albeit not quite one-for-one). As persistence drops off, the slope diminishes dramatically. For $\rho = 0$, the ratio plot is actually flat. Nevertheless, given the high persistence of the predictive variables used in practice, the more relevant ratios would be those

⁵ Similar to the earlier tables, Table 3b shows that these findings are not due to the Stambaugh bias and hold equally well for $\sigma_{zu} = 0$.

corresponding to steep slopes. These graphs show the mean of the ratio: however, understanding the full distribution allows us to examine whether the actual estimates fall within the empirical null distributions.

To better understand the statistical likelihood of the observed evidence in light of the distribution of the various relevant coefficients under the null hypothesis, Figures 3A and 3B show box plots of the distribution of the multiple horizon coefficient estimates and R^2 s conditional on the 1-year coefficient estimate and R^2 being close to the actual values (i.e., $\hat{\beta} = 0.117$ and $R^2 = 4.31\%$). The box plots shows the median, the 25th and 75th percentiles, and the more extreme 10th and 90th percentiles of the distribution. Several observations are in order. First, consistent with Figures 2A and 2B, the percentiles linearly increase at a fairly steep rate. Second, the actual values of the coefficients and R^2 s from the data (marked as diamonds in the graph) lie uniformly between the 25th and 75th percentiles. This explains why there is almost no hope of finding evidence against the null hypothesis: Given some amount of sampling error, the hypothesis of no predictability produces precisely the pattern one would expect in the coefficients under the alternative hypothesis. Because the sample sizes are relatively small, the presence of sampling error is almost guaranteed. Third, the plots show that what matters is the magnitude of the coefficient at short horizons. In the voluminous literature on stock return predictability in finance, no researcher has ever considered the short horizon evidence to be remarkable.

III. Evidence

The theory and corresponding simulation evidence in Section II suggests that it will be very difficult to distinguish between the null hypothesis of no predictability and alternative models

of time-varying expected returns that involve persistent autoregressive processes. The reason is that sampling error produces virtually identical patterns in both R^2 s and coefficients across horizons. However, this finding does not necessarily imply that joint tests will not distinguish the null from other alternatives. Recall that the null hypothesis implies highly correlated regression coefficient estimators, which induce the coefficient pattern. Even with unremarkable coefficient estimators, yet nonconforming coefficient patterns, one might find strong evidence against the null hypothesis of no predictability.

In this section, we look at a number of commonly used variables to test the predictability of stock returns. For stock returns, we use the value-weighted (VW) CRSP portfolio. For predictors, we use the log dividend yield and log earnings yield on the S&P500, the default spread between Baa and Aaa yields, the term spread between long-term government bond yields and T-bill yields, and the log book-to-market ratio.⁶ In addition, we use the log dividend yield on the CRSP VW index as well as various adjustments for repurchases and new equity issues.⁷ A final variable chosen is the aggregate equity share of new issues in the market as looked at by Baker and Wurgler (2000).

The regression analysis corresponds to equation (1), and covers horizons of 1 to 5 years over the period 1926–2004. For each set of multiple-horizon regressions, we calculate the coefficient, its analytical standard error (using results from Section II.B), its asymptotic p-value, and its simulated p-value under an AR(1) with matching parameters.⁸ In addition, we

⁶ See Goyal and Welch (2003). The data are available on Amit Goyal's website (<http://www.bus.emory.edu/AGoyal/>).

⁷ See Boudoukh, Michaely, Richardson, and Roberts (2005) for a detailed description of the various measures of payout yield. The data are available on Michael Robert's website (http://www.finance.wharton.upenn.edu/mrrobert/public_html/Research/Data/PayoutPaperDataTS.xls).

⁸ Because the equations involve overlapping observations across multiple horizons, small sample adjustments for coefficient estimators and standard errors (e.g., Amihud and Hurvich (2004) and Amihud, Hurvich, and Wang (2005)) are no longer strictly valid. As developing methods for our particular regression system lies outside the

conduct a joint Wald test across the equations, using both asymptotic and simulated p-values. These results are reported in Table 4.

Most of the series show the much-cited pattern of increasing coefficient estimates and corresponding R^2 s. For the dividend yield, the payout yield including total repurchases, and the payout yield including treasury stock-adjusted repurchases on the CRSP VW index, the log dividend and earnings yields on the S&P500, the default spread, the term spread, and the log book-to-market ratio, the increases in R^2 from the 1-year to the 5-year horizon are 4.98% to 17.24%, 7.39% to 21.04%, 6.40% to 19.37%, 4.31% to 16.11%, 3.31% to 10.65%, 0.03% to 3.31%, 3.30% to 14.43%, and 3.00% to 16.37%, respectively. However, the persistence levels of the associated variables are 93.3%, 86.4%, 90.8%, 95.3%, 76.2%, 79.2%, 61.5%, and 93.9%, respectively. It should not be surprising that many of the series, especially the dividend yield and payout yields, earnings yield, term spread, and book-to-market, have significant coefficients using asymptotic p-values across most of the horizons. Under the null hypothesis, the regressions at each horizon are virtually the same.

Table 5 is an alternative representation of the results in Table 4, i.e., the ratios of the coefficient estimates and R^2 s across horizons. For the series cited above (except for the default spread), the ratios for both quantities are similar to the simulated ratios under the null hypothesis of no predictability. In all cases, the ratios (and therefore the underlying coefficient estimates and R^2 s) increase with the horizon. Thus, the finding that some of the 1-year regressions are significant, and that the same variables produce virtually identical patterns at longer horizons, is actually evidence that the annual regression results were due to sampling

scope of this paper, we simply rely on Monte Carlo estimated p-values as a first approximation of both the correlation (e.g., Stambaugh (1999)) and long horizon (e.g., Valkanov (2003)) biases.

error.⁹ The joint tests confirm this phenomenon by generally producing higher p-values, e.g., 0.13, 0.20 and 0.20 for the dividend yield and two payout yield variables on the CRSP VW index, 0.09 and 0.39 for the dividend and earnings yield on the S&P500, 0.75 for the default spread, 0.31 for the term spread, and 0.03 for the book-to-market ratio, one of the few significant findings at the 5% level.

Several observations illustrate the nature of the joint tests. First, consider the regression results for the dividend yield and payout yields on the CRSP VW index. By almost any eyeball measure, the evidence for the payout yield appears to be stronger. All of the horizons produce greater coefficient estimates and lower p-values for payout yields than for dividend yields. For example, while six of the individual p-values are less than 0.02 for payout yields, none of the coefficients satisfy this criterion for dividend yields. Nevertheless, the p-value of the joint test for the payout yield is higher than that for the dividend yield by a significant margin.

Second, the individual coefficient estimates and corresponding R^2 s of the one significant variable under the joint cross-horizon test, i.e., book-to-market, do not look any more impressive than those for the various payout yield measures, earnings yield, or term premium. Yet the significance level of the joint test is much higher. Why? The pattern in the coefficients, while monotonic, is much less linear and one-to-one than implied by a series with a 94% persistence level. This result illustrates the power of the joint test to uncover seemingly innocuous differences across horizons. Note though that our implementation of the joint test imposes an AR(1) in X_t , and the low p-value may be partly attributable to a rejection of this assumption. Alternatively, we could have estimated the full autocorrelation structure of X_t and not imposed the AR(1), trading off efficiency for consistency.

⁹ This has even greater support once the researcher takes into account the data-snooping arguments of Foster,

Third, the simulated p-values in general show much less significance for both the individual and joint tests. This mirrors the small sample findings of Goetzmann and Jorion (1993), Kim and Nelson (1993), and Valkanov (2003). However, while book-to-market is no longer significant at the individual horizon level (with p-values increasing from the 5% level to the 20% level), the joint test remains significant at the 10% level. Again, this is because the pattern in the coefficient estimates is not consistent with that of the null hypothesis with a persistent predictive variable. As Tables 1A and 1B show, the correlation pattern across multiple-horizon estimators is robust to small sample considerations.

As a final comment, two variables, the net payout yield (i.e., payout yield plus net issuance) and equity share of new issuances, are strongly significant across horizons as evidenced by Wald statistics with p-values of 0.00 and 0.01, respectively. These results are consistent with the short-horizon findings of Boudoukh, Michaely, Richardson, and Roberts (2005) and Baker and Wurgler (2000), and show that this evidence continues to long horizons. Of some interest, while the coefficients and R^2 s are large across all horizons, the pattern is no longer monotonic. This latter finding provides even sharper evidence against the null since the series have AR(1) coefficients of 0.649 and 0.454, respectively. With this degree of autocorrelation and the overlapping horizons, one would have expected a pattern similar to the other predictive variables.

IV. Conclusion

Long-horizon stock return predictability is considered to be one of the more important pieces of evidence in the empirical asset pricing literature over the last couple of decades (see, e.g.,

Smith, and Whaley (1997).

the textbooks of Campbell, Lo, and MacKinlay (1997) and Cochrane (2001)). The evidence is set forth as a yardstick for theoretical asset pricing models and is slowly penetrating the practitioner community (for two recent examples, see Brennan and Xia (2005) and Asness (2003)). In addition, long-horizon predictability has also been documented in other markets, which is perhaps not surprising, given our analysis. The highly cited work of Fama and Bliss (1987) and Mark (1995) document results similar in spirit to the ones discussed in this paper for bond returns and exchange rates, respectively. Both papers involve highly persistent regressors and document nearly linearly increasing β s and R^2 s.

In this paper, we show that stronger long-horizon results, in the form of higher β s and increasing R^2 s, present little if any independent evidence over and above the short-horizon results for persistent regressors. Under the null hypothesis of no predictability, sampling variation can generate small levels of predictability at short horizons. This result is well known. Our research shows that higher levels of predictability at longer horizons are to be expected as well.

Appendix A

Under the null hypothesis of no predictability, $\beta_1 \cdots = \beta_J = \cdots = \beta_{J^*} = \cdots = 0$, we can rewrite regression equation system (1) for any J and J^* horizons as:

$$E[f_t(\cdot)] = E \begin{pmatrix} (R_{t,t+1} - \alpha_1 + \beta_1 X_t) \\ (R_{t,t+1} - \alpha_1 + \beta_1 X_t) X_t \\ \vdots \\ (R_{t,t+J} - J\alpha_1 + \beta_J X_t) X_t \\ \vdots \\ (R_{t,t+J^*} - J^* \alpha_1 + \beta_{J^*} X_t) X_t \\ \vdots \end{pmatrix} = 0.$$

Under the assumption of conditional homoskedasticity of the error terms above, one can apply the approach of Richardson and Smith (1991) (see also Boudoukh and Richardson (1994)) to analytically derive the asymptotic distribution of the estimators $\hat{\beta}' = (\cdots \hat{\beta}_J \cdots \hat{\beta}_{J^*} \cdots)$. In particular, applying results from Hansen (1982), the vector of regression coefficients $\hat{\theta} = (\hat{\alpha}_1 \hat{\beta})'$ has an asymptotic normal distribution with mean $(\alpha_1 \ 0)'$ and variance-covariance matrix $[D_0' S_0^{-1} D_0]^{-1}$, where $D_0 = E[\partial f_t / \partial \theta]$ and $S_0 = \sum_{l=-\infty}^{+\infty} E[f_t f_{t-l}]$. Under these assumptions, it is possible to calculate D_0 and S_0 analytically. Specifically,

$$D_0 = \begin{pmatrix} 1 & \mu_X & \cdots & 0 & \cdots & 0 & \vdots \\ \mu_X & \sigma_X^2 + \mu_X^2 & \cdots & 0 & \cdots & 0 & \vdots \\ \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ J\mu_X & 0 & \cdots & \sigma_X^2 + \mu_X^2 & \cdots & 0 & \vdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \ddots & \vdots \\ J^* \mu_X & 0 & \cdots & 0 & \cdots & \sigma_X^2 + \mu_X^2 & \vdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots \end{pmatrix},$$

and

$$S_0 = \begin{pmatrix} \sigma_R^2 & \sigma_R^2 \mu_X & \cdots & J\sigma_R^2 \mu_X & \cdots & J^* \sigma_R^2 \mu_X & \vdots \\ \sigma_R^2 \mu_X & \sigma_R^2 (\mu_X^2 + \sigma_X^2) & \cdots & \sigma_R^2 \left(J\mu_X^2 + \sigma_X^2 \left[1 + \sum_{l=1}^{J-1} \rho_l \right] \right) & \cdots & \sigma_R^2 \left(J^* \mu_X^2 + \sigma_X^2 \left[1 + \sum_{l=1}^{J^*-1} \rho_l \right] \right) & \vdots \\ \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ J\sigma_R^2 \mu_X & \sigma_R^2 \left(J\mu_X^2 + \sigma_X^2 \left[1 + \sum_{l=1}^{J-1} \rho_l \right] \right) & \cdots & \sigma_R^2 \left(J^2 \mu_X^2 + \sigma_X^2 \left[J + 2 \sum_{l=1}^{J-1} (J-l) \rho_l \right] \right) & \cdots & \sigma_R^2 \left(J J^* \mu_X^2 + \sigma_X^2 \left[J + \sum_{l=1}^{J-1} (J-l) [\rho_l + \rho_{J^*-J+l}] + \sum_{l=1}^{J^*-J} J \rho_l \right] \right) & \vdots \\ \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ J^* \sigma_R^2 \mu_X & \sigma_R^2 \left(J^* \mu_X^2 + \sigma_X^2 \left[1 + \sum_{l=1}^{J^*-1} \rho_l \right] \right) & \cdots & \sigma_R^2 \left(J J^* \mu_X^2 + \sigma_X^2 \left[J + \sum_{l=1}^{J-1} (J-l) [\rho_l + \rho_{J^*-J+l}] + \sum_{l=1}^{J^*-J} J \rho_l \right] \right) & \cdots & \sigma_R^2 \left(J^{*2} \mu_X^2 + \sigma_X^2 \left[J^* + 2 \sum_{l=1}^{J^*-1} (J^*-l) \rho_l \right] \right) & \vdots \\ \vdots & \cdots & \ddots & \cdots & \cdots & \cdots & \ddots \end{pmatrix}$$

where μ_X is the mean of X_t , σ_X^2 is the unconditional variance of X_t , ρ_l is the l^{th} order autocorrelation of X_t , σ_R^2 is the variance of single period returns R_t , and $J^* > J$. Using D_0 and S_0 above, and performing the relevant matrix calculations, one gets the desired result given in equation (2).

References

- Ait-Sahalia, Y., and M. Brandt, 2001, "Variable Selection for Portfolio Choice," *Journal of Finance*, 56 (4), 1297–1351.
- Amihud, Y., and C. Hurvich, 2004, "Predictive Regression: A Reduced-Bias Estimation Method," *Journal of Financial and Quantitative Analysis*, 39 (4), 813–841.
- Amihud, Y., C. Hurvich, and Y. Wang, 2005, "Hypothesis Testing in Predictive Regressions," NYU working paper.
- Asness, C., 2003, "Fight the Fed Model," *Journal of Portfolio Management*, 30 (1), 11–24.
- Baker, M. and J. Wurgler, 2000, "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Finance*, 55 (5), 2219–2257.
- Barberis, N., 2000, "Investing for the Long Run when Returns Are Predictable," *Journal of Finance*, 55 (1), 225–264
- Barberis, N., and R. Thaler, 2003, "A Survey of Behavioral Finance," *The Handbook of the Economics of Finance*, ed. North Holland: Amsterdam.
- Bossaerts, P., and P. Hillion, 1999, "Implementing Statistical Criteria to Implement Return Forecasting Models: What Do We Learn?" *Review of Financial Studies*, 12 (2), 405–428.
- Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts, 2005, "On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing," *Journal of Finance*, forthcoming.
- Boudoukh, J., and M. Richardson, 1994, "The Statistics of Long-Horizon Regressions Revisited," *Mathematical Finance*, 4 (2), 103–120.
- Brennan, M., and Y. Xia, 2005, "Persistence, Predictability, and Portfolio Planning," Wharton working paper.
- Campbell, J., 2000, "Asset Pricing at the Millennium," *Journal of Finance*, 55 (4), 1515–1567.
- Campbell, J., 2001, "Why Long Horizons? A Study of Power Against Persistent Alternatives," *Journal of Empirical Finance*, 8, 459–491.
- Campbell, J., 2003, "Consumption-Based Asset Pricing" in Chapter 13 of *The Handbook of the Economics of Finance*, ed. North Holland: Amsterdam, 803–881.
- Campbell, J., and J. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107 (2), 205–251.

Campbell, J., A. Lo, and C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

Campbell, J., and R. Shiller, 1988, “The Dividend–Price Ratio and Expectations of Future Dividends and Discount Factors,” *Review of Financial Studies*, 1 (3), 195–228.

Campbell, J., and L. Viceira, 1999, “Consumption and Portfolio Decisions When Expected Returns Are Time Varying,” *Quarterly Journal of Economics*, 114 (2), 433–495

Cochrane, J., 1999, “New Facts in Finance,” *Economic Perspectives*, 23 (3), 36–58.

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

Cremers, M., 2002, “Stock Return Predictability: A Bayesian Model Selection Perspective,” *Review of Financial Studies*, 15 (4), 1223–1249.

Fama, E., 1998, “Market Efficiency, Long-Term Returns, and Behavioral Finance,” *Journal of Financial Economics*, 49, 283–306.

Fama, E., and R. Bliss, 1987, “The Information in Long Maturity Forward Rates,” *American Economic Review*, 77, 680–92.

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

Ferson, W., and R. Korajczyk, 1995, “Do Arbitrage Pricing Models Explain the Predictability of Stock Returns?,” *Journal of Business*, 68 (3), 309–349

Foster, D., T. Smith, and R. Whaley, 1997, “Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R-Squared,” *Journal of Finance*, 52 (2), 591–607.

Goetzmann, W., and P. Jorion, 1993, “Testing The Predictive Power of Dividend Yields,” *Journal of Finance*, 48, 663–679.

Goyal, A., and Welch I., 2003, “Predicting the Equity Premium with Dividend Ratios,” *Management Science*, 49 (5), 639–654.

Hansen, L., 1982, “Large Sample Properties of Generalized Method of Moments,” *Econometrica*, 50 (4), 1029–1054.

Hansen, L., and R. Hodrick, 1980, “Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis,” *Journal of Political Economy*, 88 (5), 829–853.

Hodrick, R., 1992, “Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement,” *Review of Financial Studies*, 5(3), 357–386.

- Kim, M.J., and C.R. Nelson, 1993, "Predictable Returns – the Role of Small Sample Bias," *Journal of Finance*, 48 (2), 641–661.
- Kirby, C., 1997, "Measuring the Predictable Variation in Stock and Bond Returns," *Review of Financial Studies*, 10 (3), 579–630.
- Lettau, M., and S. Ludvigson, 2001, "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56 (3), 815–849.
- Lettau, M., and S. Ludvigson, 2005, "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76 (3), 583–626.
- Lewellen, J., 2004, "Predicting Returns with Financial Ratios," *Journal of Financial Economics*, 74 (2), 209–235.
- Mark, N. C., 1995, "Exchange Rates and Fundamentals, Evidence on Long Horizon Predictability," *American Economic Review*, 85 (1), 201–218.
- Menzly, L., T. Santos and P. Veronesi, 2004, "Understanding Predictability," *Journal of Political Economy*, 112, 1–47.
- Patelis, A.D., 1997, "Stock Return Predictability and the Role of Monetary Policy," *Journal of Finance*, 52 (5), 1951–1972
- Poterba, J., and L. Summers, 1988, "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 22, 27–60.
- Richardson, M., 1993, "Temporary Components of Stock Prices: A Skeptic's View," *Journal of Business and Economics Statistics*, 11 (2), 199–207.
- Richardson, M., and T. Smith, 1991, "Tests of Financial Models in the Presence of Overlapping Observations," *Review of Financial Studies*, 4 (2), 227–254.
- Richardson, M., and T. Smith, 1994, "A Unified Approach to Testing for Serial Correlation in Stock Returns," *Journal of Business*, 67 (3), 371–399.
- Richardson, M., and J. Stock, 1989, "Drawing Inferences from Statistics Based on Multi-Year Asset Returns," *Journal of Financial Economics*, 25, 323–348.
- Stambaugh, R.F., 1993, "Estimating Conditional Expectations When Volatility Fluctuates," NBER Technical Paper 140.
- Stambaugh, R.F., 1999, "Predictive Regressions," *Journal of Financial Economics*, 54 (3), 375–421.

Valkanov, R., 2003, “Long-Horizon Regressions: Theoretical Results and Applications,” *Journal of Financial Economics*, 68 (2), 201–232.

Panel A: $\sigma_{eu} = -0.683$

Coefficient estimates				Correlations			
Horizon	Mean	SD	Median	Horizon			
				2	3	4	5
1	0.048	0.068	0.037	0.966	0.925	0.883	0.841
2	0.092	0.128	0.074		0.979	0.946	0.908
3	0.132	0.182	0.109			0.985	0.957
4	0.169	0.232	0.142				0.988
5	0.203	0.277	0.175				
% monotonic		65.24					

Test statistics				Size		
	Mean	SD	Median	10%	5%	1%
Wald	7.733	119.281	5.536	20.75	13.15	5.42
P-value	0.396	0.294	0.354			

Panel B: $\sigma_{eu} = 0$

Coefficient estimates				Correlations			
Horizon	Mean	SD	Median	Horizon			
				2	3	4	5
1	0.000	0.063	0.000	0.960	0.913	0.867	0.823
2	0.000	0.121	0.001		0.977	0.940	0.900
3	0.000	0.175	0.001			0.984	0.954
4	0.000	0.227	0.002				0.987
5	0.000	0.276	0.001				
% monotonic		57.22					

Test statistics				Size		
	Mean	SD	Median	10%	5%	1%
Wald	7.199	59.730	5.203	18.92	12.06	5.10
P-value	0.423	0.300	0.392			

Table 1: Distribution of Coefficient Estimates and Test Statistics

Panel A reports the mean, standard deviation, and median of the coefficient estimates from the predictive regression (equation (1)), and the correlations between these estimates for horizons of 1 to 5 years across 100,000 simulations. “% monotonic” is the percentage of the simulations that produce coefficients that are monotonic in the horizon. Panel A also reports the mean, standard deviation, and median of the joint Wald test statistic (across horizons), the associated p-values, and the percentage of statistics that reject the null hypothesis of no predictability at the 10%, 5%, and 1% levels. There are 75 observations for each simulation, and simulations are performed under the null hypothesis of no predictability using the parameters $\rho = 0.953$, $\sigma_{\varepsilon} = 0.196$, $\sigma_u = 0.166$, $\sigma_{eu} = -0.683$. Panel B reports the same statistics for $\sigma_{eu} = 0$ (all other simulation parameters are the same as in Panel A).

Panel A: $\sigma_{\varepsilon u} = -0.683$

Horizon	Mean	SD	Median	Correlations			
				Horizon			
				2	3	4	5
1	1.791	2.343	0.885	0.948	0.886	0.823	0.762
2	3.386	4.280	1.746		0.968	0.916	0.858
3	4.845	5.948	2.564			0.976	0.932
4	6.182	7.412	3.326				0.981
5	7.415	8.708	4.089				
% monotonic		51.30					

Panel B: $\sigma_{\varepsilon u} = 0$

Horizon	Mean	SD	Median	Correlations			
				Horizon			
				2	3	4	5
1	1.345	1.861	0.619	0.927	0.845	0.767	0.695
2	2.524	3.399	1.202		0.957	0.892	0.821
3	3.612	4.769	1.743			0.969	0.914
4	4.624	5.992	2.278				0.976
5	5.570	7.094	2.790				
% monotonic		42.50					

Table 2: Distribution of R^2 's

Panel A reports the mean, standard deviation, and median of the R^2 's from the predictive regression (equation (1)) and the correlations between them for horizons of 1 to 5 years across 100,000 simulations. “% monotonic” is the percentage of the simulations that produce R^2 's that are monotonic in the horizon. There are 75 observations for each simulation, and simulations are performed under the null hypothesis of no predictability using the parameters $\rho = 0.953$, $\sigma_\varepsilon = 0.196$, $\sigma_u = 0.166$, $\sigma_{\varepsilon u} = -0.683$. Panel B reports the same statistics for $\sigma_{\varepsilon u} = 0$ (all other simulation parameters are the same as in Panel A).

Panel A: $\sigma_{eu} = -0.683$

Horizon	Coefficient estimate ratios				R ² ratios			
	Mean	SD	Median	# of sim.	Mean	SD	Median	# of sim.
2	1.930	0.854	1.916	87,166	1.953	0.826	1.868	61,467
3	2.784	1.842	2.731	87,166	2.872	1.752	2.597	61,467
4	3.564	2.963	3.461	87,166	3.753	2.792	3.209	61,467
5	4.276	4.183	4.119	87,166	4.595	3.891	3.739	61,467
% monotonic		69.85		% monotonic		59.40		

Panel B: $\sigma_{eu} = 0$

Horizon	Coefficient estimate ratios				R ² ratios			
	Mean	SD	Median	# of sim.	Mean	SD	Median	# of sim.
2	1.869	1.102	1.886	85,717	1.904	0.961	1.782	54,603
3	2.615	2.309	2.663	85,717	2.754	1.992	2.397	54,603
4	3.261	3.655	3.346	85,717	3.548	3.086	2.884	54,603
5	3.822	5.085	3.94	85,717	4.292	4.204	3.234	54,603
% monotonic		61.57		% monotonic		48.88		

Table 3: Distribution of Coefficient Estimate and R² Cross-Horizon Ratios

Panel A reports the mean, standard deviation, and median of the coefficient estimate and R² ratios (i.e., $\hat{\beta}_i/\hat{\beta}_1$ and R^2_i/R^2_1 , $i = 2, \dots, 5$) from the predictive regression (equation (1)) across the simulations out of the 100,000 for which $\hat{\beta}_1 > 0.01$ or $R^2_1 > 0.5\%$, respectively. “% monotonic” is the percentage of these simulations that produce coefficient estimates and R²s that are monotonic in the horizon. There are 75 observations for each simulation, and simulations are performed under the null hypothesis of no predictability using the parameters $\rho = 0.953$, $\sigma_\varepsilon = 0.196$, $\sigma_u = 0.166$, $\sigma_{eu} = -0.683$. Panel B reports the same statistics for $\sigma_{eu} = 0$ (all other simulation parameters are the same as in Panel A).

	Horizon					Wald
	1	2	3	4	5	
Log dividend yield, S&P500 ($\rho = 0.953$)						
$\hat{\beta}$	0.117	0.236	0.355	0.424	0.469	9.559
Std. err.	0.066	0.130	0.193	0.255	0.316	
Asym. p-value	0.037	0.034	0.033	0.048	0.069	0.089
Sim. p-value	0.144	0.129	0.116	0.137	0.166	0.190
R ²	4.315	8.019	13.009	14.856	16.107	
Log earnings yield, S&P500 ($\rho = 0.762$)						
$\hat{\beta}$	0.098	0.217	0.315	0.370	0.366	5.241
Std. err.	0.062	0.117	0.167	0.214	0.257	
Asym. p-value	0.058	0.032	0.030	0.042	0.077	0.387
Sim. p-value	0.127	0.078	0.068	0.081	0.128	0.417
R ²	3.308	7.396	10.907	12.315	10.647	
Default yield spread ($\rho = 0.792$)						
$\hat{\beta}$	0.429	3.299	4.967	7.209	9.113	2.673
Std. err.	2.782	5.268	7.582	9.742	11.767	
Asym. p-value	0.439	0.266	0.256	0.230	0.219	0.750
Sim. p-value	0.561	0.392	0.383	0.353	0.340	0.778
R ²	0.032	0.855	1.364	2.347	3.313	
Term yield spread ($\rho = 0.615$)						
$\hat{\beta}$	2.694	4.046	6.233	9.681	11.707	5.966
Std. err.	1.712	3.078	4.270	5.325	6.270	
Asym. p-value	0.058	0.094	0.072	0.035	0.031	0.310
Sim. p-value	0.076	0.111	0.085	0.041	0.035	0.329
R ²	3.300	3.394	5.669	11.175	14.432	
Log book-to-market ratio ($\rho = 0.939$)						
$\hat{\beta}$	0.075	0.174	0.272	0.341	0.366	12.050
Std. err.	0.050	0.099	0.147	0.193	0.239	
Asym. p-value	0.067	0.039	0.032	0.039	0.063	0.034
Sim. p-value	0.265	0.199	0.167	0.175	0.221	0.090
R ²	2.997	7.287	12.568	16.072	16.366	

Table 4: Coefficient Estimates and R²s from Predictive Regressions

The table reports results from the regression of 1- to 5-year CRSP value-weighted returns on various lagged predictor variables (equation (1)) for the period 1926–2004 (75 observations). $\hat{\beta}$ is the estimated coefficient, with associated asymptotic standard error (equation (3)), p-value under the null hypothesis of no predictability, and the asymptotic Wald test and p-value for the joint hypothesis of no predictability across horizons. The table also reports simulated p-values (100,000 simulations) for both the individual coefficients and the Wald test. There are 75 observations for each simulation, and simulations are performed under the null hypothesis of no predictability using parameters estimated from the data.

	Horizon					Wald
	1	2	3	4	5	
Log dividend yield, CRSP VW ($\rho = 0.933$)						
$\hat{\beta}$	0.136	0.273	0.404	0.474	0.526	8.405
Std. err.	0.071	0.140	0.207	0.273	0.337	
Asym. p-value	0.027	0.025	0.025	0.041	0.059	0.135
Sim. p-value	0.137	0.121	0.113	0.138	0.164	0.226
R ²	4.977	9.181	14.357	15.776	17.244	
Log payout yield, CRSP VW, cash flow ($\rho = 0.864$)						
$\hat{\beta}$	0.192	0.374	0.528	0.608	0.673	7.260
Std. err.	0.082	0.159	0.232	0.302	0.369	
Asym. p-value	0.010	0.009	0.011	0.022	0.034	0.202
Sim. p-value	0.065	0.054	0.053	0.071	0.087	0.249
R ²	7.386	12.820	18.257	19.415	21.036	
Log payout yield, CRSP VW, Treasury stock ($\rho = 0.908$)						
$\hat{\beta}$	0.166	0.322	0.460	0.535	0.598	7.257
Std. err.	0.076	0.149	0.219	0.288	0.354	
Asym. p-value	0.015	0.015	0.018	0.031	0.046	0.202
Sim. p-value	0.094	0.086	0.084	0.106	0.126	0.278
R ²	6.400	11.097	16.158	17.539	19.374	
Log net payout yield, CRSP VW, cash flow ($\rho = 0.649$)						
$\hat{\beta}$	0.737	1.376	1.613	1.629	1.611	24.068
Std. err.	0.168	0.304	0.425	0.533	0.631	
Asym. p-value	0.000	0.000	0.000	0.001	0.005	0.000
Sim. p-value	0.000	0.000	0.000	0.002	0.007	0.000
R ²	26.105	41.778	41.025	33.537	29.023	
Equity share of new issuances ($\rho = 0.454$)						
$\hat{\beta}$	-0.742	-1.296	-1.503	-1.581	-1.432	14.647
Std. err.	0.212	0.361	0.484	0.588	0.680	
Asym. p-value	0.000	0.000	0.001	0.004	0.018	0.012
Sim. p-value	0.001	0.000	0.001	0.004	0.018	0.011
R ²	17.079	24.537	23.496	20.410	15.755	

Table 4 Cont'd

		Horizon			
		2	3	4	5
ln (D/P) (S&P500)	$\hat{\beta}_i / \hat{\beta}_1$	2.012	3.031	3.618	4.005
	R^2_i / R^2_1	1.858	3.015	3.443	3.733
ln (E/P) (S&P500)	$\hat{\beta}_i / \hat{\beta}_1$	2.214	3.205	3.768	3.728
	R^2_i / R^2_1	2.236	3.297	3.723	3.219
Default spread	$\hat{\beta}_i / \hat{\beta}_1$	7.686	11.572	16.796	21.232
	R^2_i / R^2_1	26.939	42.980	73.974	104.411
Term spread	$\hat{\beta}_i / \hat{\beta}_1$	1.502	2.314	3.594	4.345
	R^2_i / R^2_1	1.028	1.718	3.386	4.373
ln (B/M)	$\hat{\beta}_i / \hat{\beta}_1$	2.309	3.615	4.522	4.855
	R^2_i / R^2_1	2.431	4.193	5.362	5.460
ln (D/P) (CRSP VW)	$\hat{\beta}_i / \hat{\beta}_1$	2.005	2.965	3.471	3.858
	R^2_i / R^2_1	1.845	2.884	3.169	3.464
ln (payout/P) (CRSP VW, CF)	$\hat{\beta}_i / \hat{\beta}_1$	1.945	2.745	3.161	3.498
	R^2_i / R^2_1	1.736	2.472	2.628	2.848
ln (payout/P) (CRSP VW, TS)	$\hat{\beta}_i / \hat{\beta}_1$	1.944	2.774	3.228	3.607
	R^2_i / R^2_1	1.734	2.525	2.740	3.027
ln (net payout/P) (CRSP VW, CF)	$\hat{\beta}_i / \hat{\beta}_1$	1.867	2.189	2.210	2.186
	R^2_i / R^2_1	1.600	1.572	1.285	1.112
Equity share of new issuances	$\hat{\beta}_i / \hat{\beta}_1$	1.746	2.026	2.131	1.930
	R^2_i / R^2_1	1.437	1.376	1.195	0.922

Table 5: Coefficient Estimate and R² Ratios from Predictive Regressions

The table reports the coefficient estimate ratios ($\hat{\beta}_i / \hat{\beta}_1$, $i = 2, \dots, 5$) and the R² ratios (R^2_i / R^2_1 , $i = 2, \dots, 5$) from the regression of 1- to 5-year CRSP value-weighted returns on various lagged predictor variables (equation (1)) for the period 1926–2004 (75 observations).

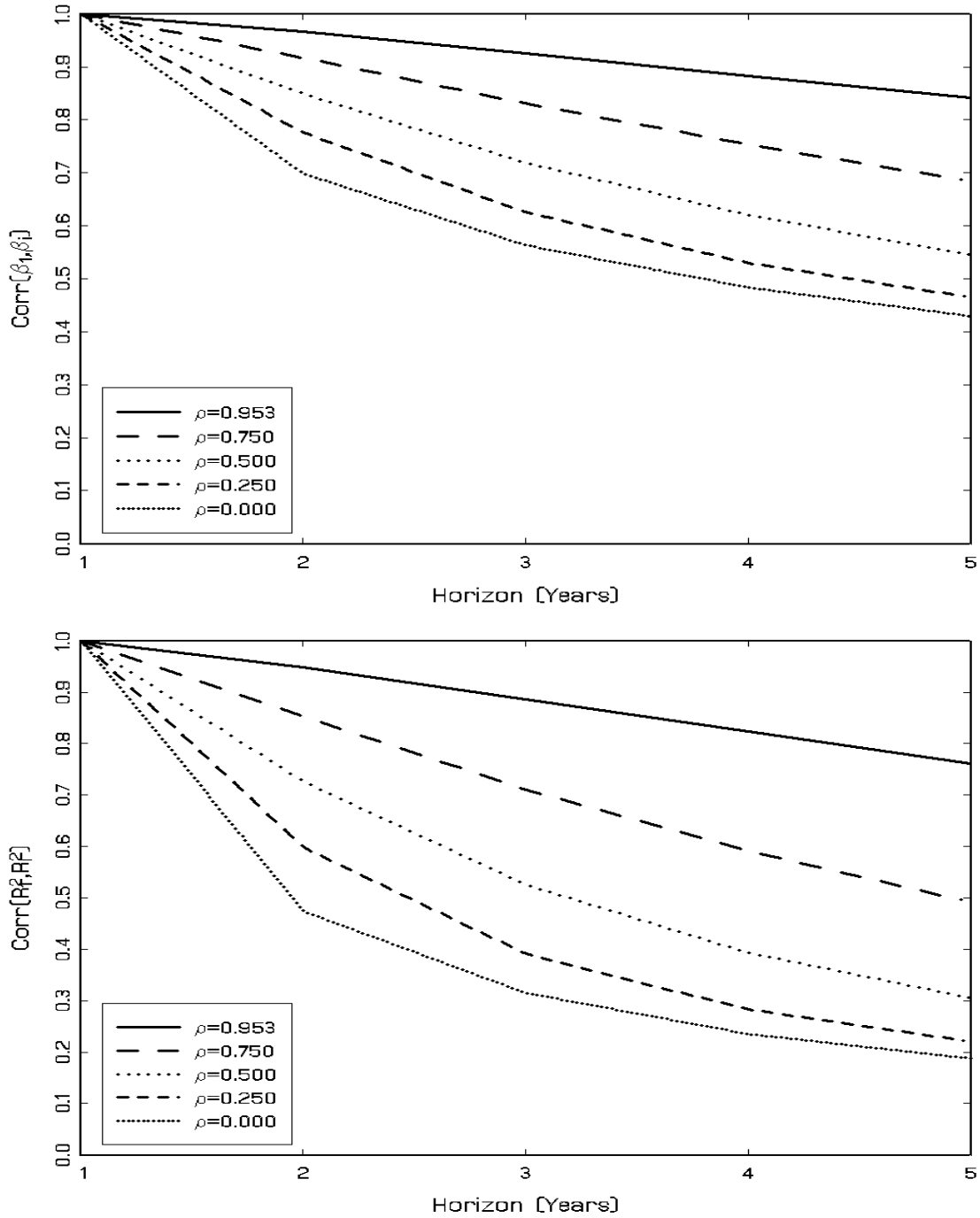


Figure 1: Cross-Horizon Correlations between Coefficient Estimates and R^2 s

The top panel plots the correlation between the coefficient estimate at the 1-year horizon and those at the 2- to 5-year horizons from the predictive regression (equation (1)) across 100,000 simulations for different values of ρ (the autocorrelation of the predictor variable). There are 75 observations for each simulation, and simulations are performed under the null hypothesis of no predictability using the parameters $\sigma_e = 0.196$, $\sigma_u = 0.166$, $\sigma_{eu} = -0.683$. The bottom panel plots the analogous correlations for the predictive regression R^2 s.

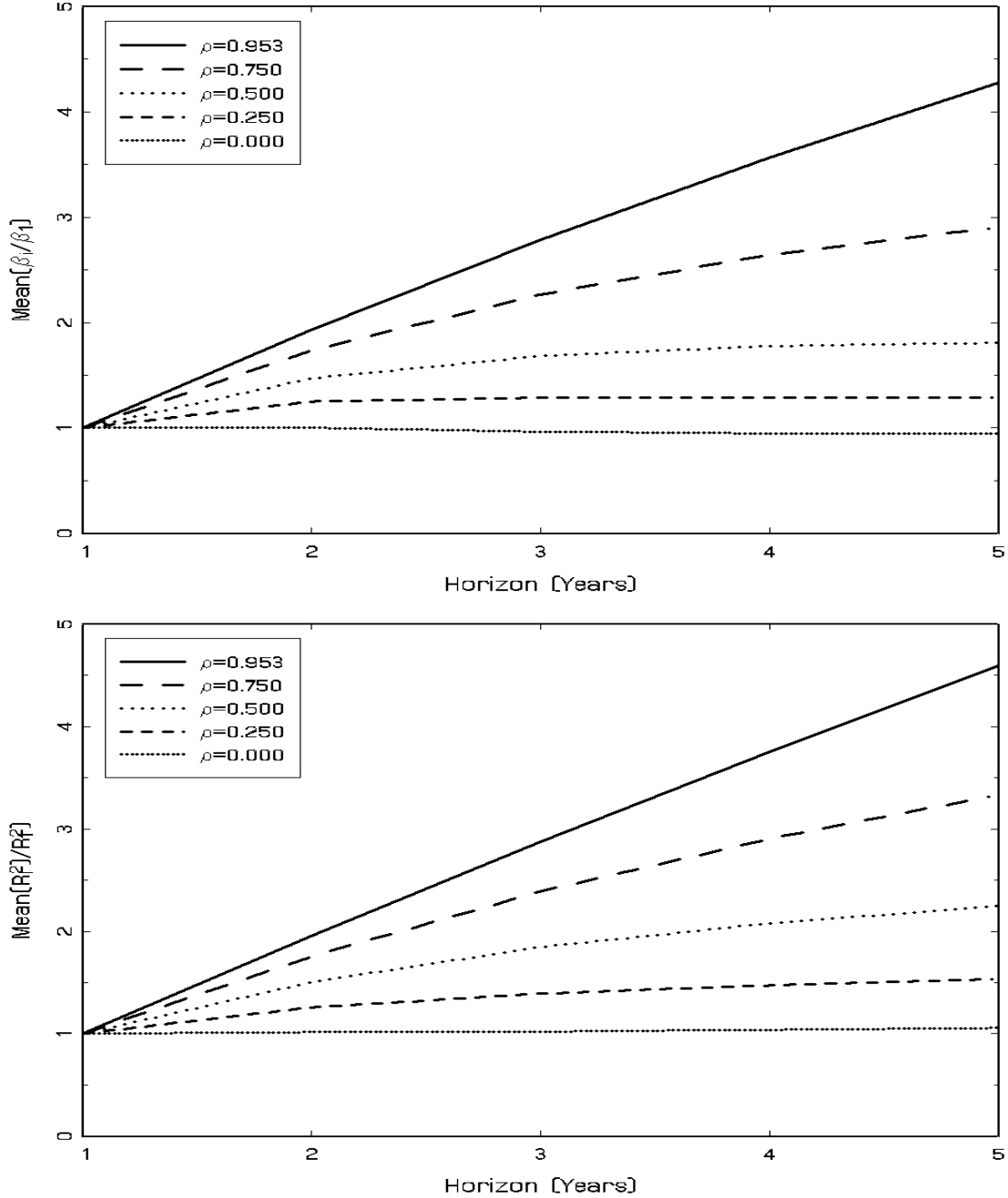


Figure 2: Mean Coefficient Estimate and R^2 Cross-Horizon Ratios

The top panel plots the mean coefficient estimate ratios (i.e., $\hat{\beta}_i/\hat{\beta}_1$, $i = 2, \dots, 5$) from the predictive regression (equation (1)) across the simulations out of the 100,000 for which $|\hat{\beta}_1| > 0.01$ for different values of ρ (the autocorrelation of the predictor variable). There are 75 observations for each simulation, and simulations are performed under the null hypothesis of no predictability using the parameters $\sigma_\varepsilon = 0.196$, $\sigma_u = 0.166$, $\sigma_{\varepsilon u} = -0.683$. The bottom panel plots the means of the analogous R^2 ratios (i.e., R^2_i/R^2_1 , $i = 2, \dots, 5$) for simulations with or $R^2_1 > 0.5\%$.

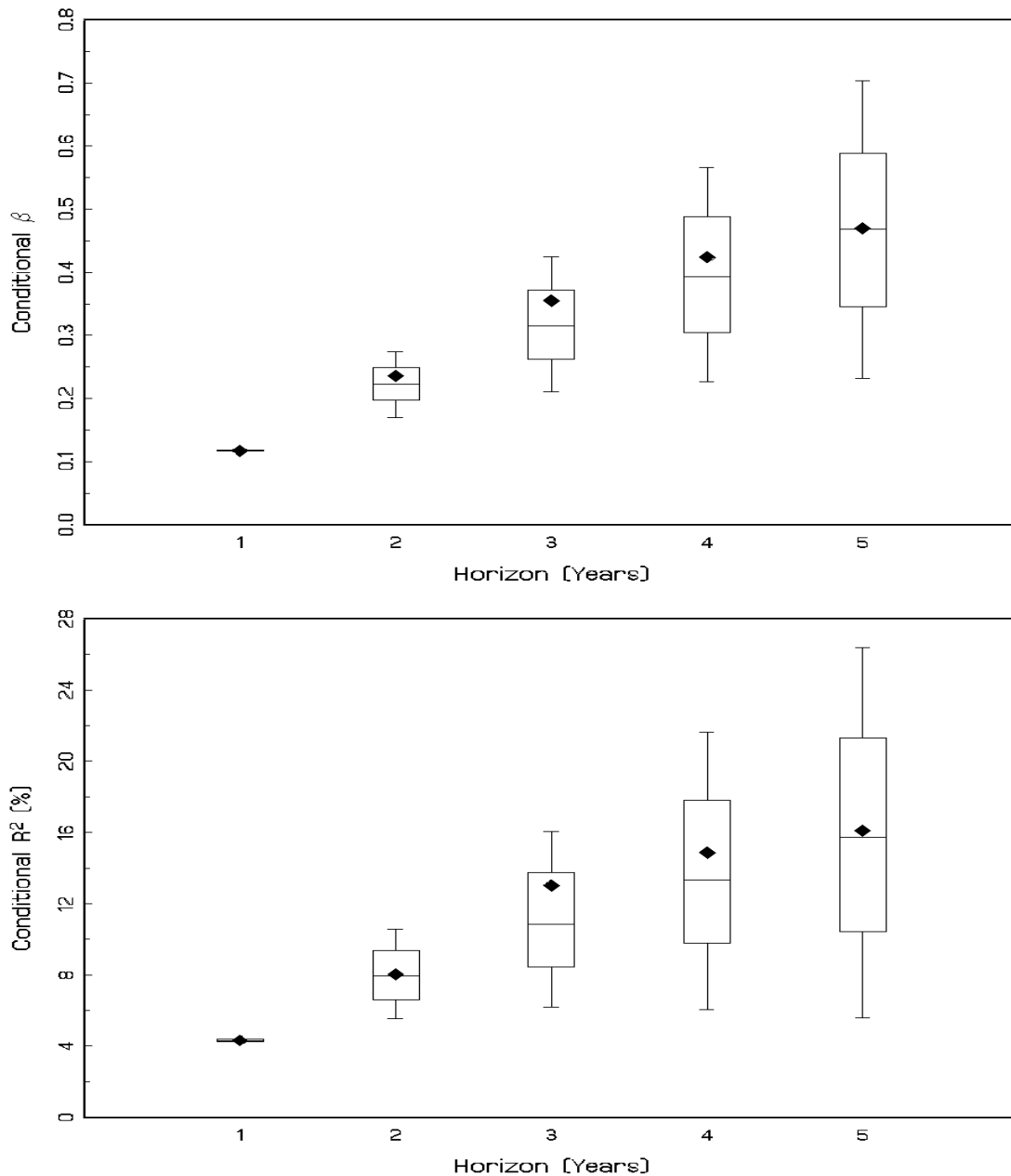


Figure 3: Conditional Distribution of Coefficient Estimates and R^2 s

The top panel provides a box plot of the simulated distributions of the coefficient estimates for horizons 2- to 5-years from the predictive regression (equation (1)) for the 971 out of 100,000 simulations for which $0.115 < \hat{\beta}_1 < 0.119$. The boxes show the median, 25th/75th percentiles, and 10th/90th percentiles. The diamonds mark the actual coefficient estimates from the first regression in Table 4. There are 75 observations for each simulation, and simulations are performed under the null hypothesis of no predictability using the parameters $\rho = 0.953$, $\sigma_\varepsilon = 0.196$, $\sigma_u = 0.166$, $\sigma_{\varepsilon u} = -0.683$. The bottom panel plots the analogous simulated distributions for the predictive regression R^2 s for the 899 simulations for which $4.215\% < R^2_1 < 4.414\%$ and the corresponding actual R^2 s.