

Valuation Ratios and Long-Horizon Stock Price Predictability

David E. Rapach
Department of Economics
Saint Louis University
3674 Lindell Boulevard
Saint Louis, MO 63108-3397
Phone: 314-977-3601
Fax: 314-977-1478
E-mail: rapachde@slu.edu

Mark E. Wohar*
Department of Economics
University of Nebraska at Omaha
CBA 512K
Omaha, NE 68182-0286
Phone: 402-554-3712
Fax: 402-554-2853
E-mail: mwohar@unomaha.edu

October 23, 2003

ABSTRACT

Using annual data for 1872-1997, this paper re-examines the predictability of real stock prices based on price-dividend and price-earnings ratios. In line with the extant literature, we find significant evidence of increased long-horizon predictability; that is, the hypothesis that the current value of a valuation ratio is uncorrelated with future stock price changes cannot be rejected at short horizons but can be rejected at longer horizons based on bootstrapped critical values constructed from linear representations of the data. While increased statistical power at long horizons in finite samples provides a possible explanation for the pattern of predictability in the data, we find via Monte Carlo simulations that the power to detect predictability in finite samples does not increase at long horizons in a linear framework. An alternative explanation for the pattern of predictability in the data is nonlinearities in the underlying data-generating process. We consider exponential smooth-transition autoregressive models of the price-dividend and price-earnings ratios and their ability to explain the pattern of stock price predictability in the data.

JEL classifications: C22, C32, C53, G12

Key words: Predictive regression; Monte Carlo simulation; Nonlinear mean-reversion

*Corresponding author. The authors thank two anonymous referees and the co-editor for very helpful comments. The usual disclaimer applies. The results reported in this paper were generated using GAUSS 3.6. The GAUSS programs are available at <http://pages.slu.edu/faculty/rapachde/>.

1. INTRODUCTION

A central research topic in financial economics is the predictability of stock prices. Predictability is typically assessed in the context of a predictive regression model, where real stock price growth or real equity returns over various horizons are regressed on a variable thought to potentially explain future movements in stock prices. While a number of econometric difficulties are inherent in predictive regressions (Mankiw and Shapiro, 1986; Stambaugh, 1986, 1999; Nelson and Kim, 1993; Kirby, 1997), the consensus appears to be that some variables can, in fact, predict future movements in stock prices (Campbell, 1999, 2000). Among these variables are valuation ratios based on measures of fundamental value, such as the price-dividend and price-earnings ratios. The oft-cited studies of Fama and French (1988) and Campbell and Shiller (1988) find that price-dividend and price-earnings ratios predict future real equity returns, and, more recently, Campbell and Shiller (1998) find that these ratios are useful in predicting future growth in real stock prices using data spanning the late nineteenth to the late twentieth centuries. An interesting pattern emerges in studies of stock price predictability based on valuation ratios: significant evidence of predictability is typically found at long, but not short, horizons; that is, the hypothesis that the current value of a valuation ratio is uncorrelated with future stock price changes cannot be rejected at short horizons but can be rejected at longer horizons. For example, Campbell and Shiller (1998) find that the price-dividend and price-earnings ratios have a very limited ability to predict real stock price growth over the next year, but a strong and significant ability to predict real stock price growth over the next ten years. This pattern has attracted a considerable amount of attention in the literature, and Campbell and Shiller (1998, p. 24) conclude that, despite the econometric difficulties inherent in long-horizon predictive regressions, “it is striking how well the evidence for [long-horizon] stock market predictability survives the various corrections that have been proposed.”

In this paper, we re-examine the ability of valuation ratios to predict real stock price growth at long horizons in light of the recent studies of Kilian (1999), Berkowitz and Giorgianni (2001), and Mark and Sul (2002). These studies are primarily concerned with the long-horizon predictability of nominal exchange

rates based on monetary fundamentals, originally detected by Mark (1995), but their insights are also applicable to long-horizon stock price predictability based on measures of fundamental equity value. We begin our empirical analysis by estimating predictive regression models for real stock price growth with the log-level of either the price-dividend or price-earnings ratio serving as the explanatory variable. Using the annual data spanning 1872-1997 from Campbell and Shiller (1998), and in line with the extant literature, we find significant evidence of real stock price predictability at long, but not short, horizons when inference is based on linear models.¹ A stable long-run valuation ratio provides an apparently plausible explanation for the long-horizon predictability of real stock prices: a stable long-run valuation ratio means that a measure of fundamental value, such as real dividends or real earnings, provides a long-run anchor for real stock prices, so that we might expect stock prices to eventually adjust to restore the long-run equilibrium between stock prices and fundamental value. While the existence of a stable long-run valuation ratio seems to provide a rationale for estimating long-horizon regressions, recent econometric research by Berkowitz and Giorgianni (2001) questions this justification in a linear framework. This research emphasizes that a linear framework characterized by a stable long-run valuation ratio logically implies real stock price predictability at *all* horizons or predictability at *no* horizon. Intuitively, in a linear framework, long-horizon forecasts are simple extrapolations of short-horizon forecasts. The pattern of stock price predictability in the data thus presents something of a puzzle in a linear framework.

One possible explanation for the pattern of stock price predictability in the data centers on statistical power. While the Berkowitz and Giorgianni (2001) argument that a linear framework implies predictability at all horizons or no horizon is logically correct, it may be the case in certain circumstances that the power to detect predictability in a linear framework is greater at long horizons. Using various asymptotic frameworks and Monte Carlo simulations for finite samples, several recent studies investigate whether there are power advantages at long horizons in predictive regression tests. The results are somewhat mixed.

¹ While real stock price predictability can be interpreted as evidence against the efficient market hypothesis, it can be reconciled with the efficient market hypothesis if we allow for time-varying expected real returns (Fama and French, 1989; Fama, 1991).

Berben and van Dijk (1998) find that local asymptotic power is not increasing in the horizon, while Mark and Sul (2002) identify local asymptotic power advantages at long horizons in certain regions of the admissible parameter space that are confirmed for finite samples in Monte Carlo experiments. Campbell (2001) also finds potential asymptotic power gains at long horizons and some power gains at long horizons for finite samples in Monte Carlo simulations. Ultimately, whether increased statistical power ensues at long horizons is an issue that needs to be analyzed for the sample sizes and data properties relevant to the particular application at hand, as in Kilian (1999). Using Monte Carlo simulations based on data from the post-Bretton Woods period, Kilian (1999) investigates whether there are power advantages at long horizons in tests of nominal exchange rate predictability based on monetary fundamentals.

In the spirit of Kilian (1999), we search for evidence of increased power at long horizons using Monte Carlo simulations based on the Campbell and Shiller (1998) real stock price, dividends, and earnings data. Interestingly—and similar to Kilian (1999)—the Monte Carlo simulations indicate that the power to detect real stock price predictability does not increase at long horizons; in fact, the simulations reveal that power decreases at long horizons in predictive regression tests based on either the price-dividend or price-earnings ratio. Our results suggest that increased statistical power at long horizons cannot explain the pattern of stock price predictability detected in the Campbell and Shiller (1998) data.

As argued by Kilian (1999) and Kilian and Taylor (2003) in the context of exchange rate predictability, the observed pattern of real stock price predictability in the Campbell and Shiller (1998) data could be construed as evidence for nonlinearities in the underlying data-generating process (DGP), as increased long-horizon predictability may arise naturally in a nonlinear framework. In light of this, we utilize a version of the parsimonious exponential smooth-transition autoregressive (ESTAR) model proposed by Kilian and Taylor (2003). This model allows for nonlinear mean-reversion in the relevant valuation ratio and has a straightforward economic interpretation. We find that the ESTAR specification fits the data reasonably well. We also re-assess the significance of the predictive regression tests using a modified bootstrap procedure, similar to the procedure developed by Kilian and Taylor (2003), based on a

nonlinear DGP. Using this modified bootstrap procedure, we find significant evidence of stock price predictability at long, but not short, horizons. Finally, we provide Monte Carlo evidence that these test results may plausibly be attributed to power gains that accrue at longer horizons when inference is based on a nonlinear framework. We conclude that, at least for the price-dividend ratio, a nonlinear framework provides a plausible explanation for the pattern of stock price predictability in the Campbell and Shiller (1998) data.

The rest of this paper is organized as follows. Section 2 presents real stock price predictive regression results using the long span of annual data from Campbell and Shiller (1998). In Section 3, we conduct Monte Carlo simulations in order to examine whether the power to detect predictability increases with the horizon in a linear framework for the Campbell and Shiller (1998) data. Section 4 considers modeling the price-dividend and price-earnings ratios using a nonlinear ESTAR specification and analyzes long-horizon stock price predictability in a nonlinear framework. Section 5 concludes.

2. PREDICTIVE REGRESSIONS

In this section, we estimate predictive regressions at both short and long horizons, as in Campbell and Shiller (1988, 1998). We use the annual S&P 500 nominal stock price, dividends, and earnings indexes from Campbell and Shiller (1998), which begin in 1871 and extend to 1997. We deflate the three nominal indexes using the consumer price index in order to obtain series for real stock prices, real dividends, and real earnings.² As in Campbell and Shiller (1998), we consider the ability of the price-dividend ratio (January real stock price divided by real dividends over the previous calendar year) and the price-earnings ratio (January real stock price divided by real earnings over the previous calendar year) to predict future real stock prices. The two valuation ratios are plotted in Figure 1. Campbell and Shiller (1998) find that these valuation ratios are useful for forecasting changes in real stock prices at long horizons. They present

² The S&P 500 data are available from Robert Shiller's home page at <http://www.econ.yale.edu/~shiller>. The complete documentation for the data sources is also provided there. Note that Campbell and Shiller (1998) use the producer price index to deflate the nominal stock price, dividends, and earnings series. Our results change little if we use the producer price index instead of the consumer price index to deflate the nominal variables.

scatterplots and R^2 measures that indicate a weak ability for the price-dividend ratio to forecast real stock price growth over the next year, but a strong ability for the price-dividend ratio to forecast real stock price growth over the next ten years. A similar result obtains for the price-earnings ratio. We re-examine these findings based on formal statistical tests of the no-predictability null.

We report predictive regression results for the price-dividend and price-earnings ratios in Table 1. More specifically, Table 1 reports estimation results for the predictive regression model,

$$\Delta p_{t+k}^k = \alpha_k + \beta_k z_t + \varepsilon_{t+k}^k, \quad (1)$$

where p_t is the real stock price in log-levels; $\Delta p_{t+k}^k = p_{t+k} - p_t$; $z_t = p_t - f_t$ is the relevant valuation ratio in log-levels, with f_t signifying the relevant fundamental (real dividends or real earnings) in log-levels. We denote the log-level of real dividends by d_t and the log-level of real earnings by e_t , so that the two valuation ratios in log-levels are $z_t = p_t - d_t$ and $z_t = p_t - e_t$. We consider values of k ranging from 1-10 years in equation (1). In a predictive regression model such as equation (1), the predictive ability of z_t is assessed through the t -statistic corresponding to $\hat{\beta}_k$, the OLS estimate of β_k . As is well known, when $k > 1$, the observations for the dependent variable in equation (1) are overlapping, and this induces serial correlation in the disturbance term, ε_{t+k}^k . Following much of the extant literature, our standard errors use the Newey and West (1987) adjustment based on the Bartlett kernel, which is robust to heteroskedasticity and serial correlation.³

A potential problem with estimating a predictive regression such as equation (1) is small-sample bias. Stambaugh (1986, 1999) considers the case where $k = 1$ in equation (1),

$$\Delta p_{t+1} = \alpha_1 + \beta_1 z_t + \varepsilon_{t+1}, \quad (2)$$

where ε_{t+1} is an independently and identically distributed disturbance term with mean zero. He shows that the OLS estimate of β_1 in equation (2), $\hat{\beta}_1$, is biased in finite samples when the DGP for Δp_{t+1} is governed

³ We select the lag truncation for the Newey and West (1987) adjustment using the Andrews (1991) automatic bandwidth selector with Andrews and Monohan (1992) prewhitening.

by equation (2) and z_t is generated by the stationary first-order autoregressive (AR(1)) process,

$$z_{t+1} = \phi + \rho z_t + \mu_{t+1}, \quad (3)$$

where μ_{t+1} is an independently and identically distributed disturbance term with mean zero and $-1 < \rho < 1$ ($0 < \rho < 1$ is the relevant range for our valuation ratios). In particular, Stambaugh (1986, 1999) shows that

$$E(\hat{\beta}_1 - \beta_1) \approx (\sigma_{\varepsilon\mu} / \sigma_\mu^2) E(\hat{\rho} - \rho), \quad (4)$$

where $\hat{\rho}$ is the OLS estimate of ρ in equation (3); $\sigma_{\varepsilon\mu}$ is the covariance between the disturbances terms, ε_t and μ_t ; σ_μ^2 is the variance of μ_t . It is well known that $\hat{\rho}$ is a biased estimator of ρ , with the bias given by (Hurwicz, 1950; Kendall, 1954; Shaman and Stine, 1988)

$$E(\hat{\rho} - \rho) \approx -(1 + 3\rho) / T, \quad (5)$$

where T is the effective sample size. From equation (5), we see that $\hat{\rho}$ can substantially underestimate ρ in finite samples, with the bias in $\hat{\rho}$ increasing as ρ approaches unity (that is, as z_t becomes more persistent). The bias in $\hat{\rho}$ is passed on to $\hat{\beta}_1$ via equation (4) when $\sigma_{\varepsilon\mu} \neq 0$ (and σ_μ^2 is not “too large”). In fact, $\sigma_{\varepsilon\mu}$ is likely to be positive and sizable, as a positive shock to the growth rate of real stock prices is also likely to generate an increase in a valuation ratio such as $p_t - d_t$ or $p_t - e_t$. Taken together, a persistent z_t series and highly correlated residuals across equations (2) and (3) can result in a substantial bias in finite samples for $\hat{\beta}_1$.⁴ As shown by, among others, Nelson and Kim (1993), these biases and the overlapping nature of the observations at horizons beyond the one-step-ahead horizon can have the effect of severely shifting the distribution of the t -statistic for $\hat{\beta}_k$, even when the t -statistic is computed using the Newey and West (1987) procedure, so that basing inferences on standard asymptotic results can lead to considerable size distortions when testing the null hypothesis of no predictability ($\beta_k = 0$).

⁴ Using equation (4) and values of $\sigma_{\varepsilon\mu}$, σ_μ^2 , and $E(\hat{\rho} - \rho)$ implied by the Campbell and Shiller (1998) data, the ratio of the bias to the least-squares point estimate of β_1 equals 0.47 for $z_t = p_t - d_t$ and 0.80 for $z_t = p_t - e_t$. These ratios are sizable and similar to those reported in Stambaugh (1999).

In order to make valid inferences for our predictive regression tests, we rely on a bootstrap procedure similar to the procedures in Nelson and Kim (1993), Mark (1995), Kothari and Shanken (1997), and Kilian (1999). As a first step, we use the original data to estimate a DGP for Δp_t under the null hypothesis of no predictability by estimating equation (2) via OLS with β_1 set to zero (so that p_t is generated by a random walk with drift under the null hypothesis). We also use the original data to estimate an AR process for z_t , with the lag order for the AR model selected using the AIC (considering a maximum lag order of four),⁵ and we adjust the OLS estimates of the AR parameters in order to correct for their small-sample biases using the theoretical results in Shaman and Stine (1988, Table 1). Armed with the parameter estimates, we can build up a pseudo-sample of observations for Δp_t and z_t by randomly drawing (with replacement) from the OLS residuals. Note that we draw the OLS residuals, $\hat{\varepsilon}_t$ and $\hat{\mu}_t$, in tandem, so that the pseudo-sample preserves the contemporaneous correlation in the disturbances present in the original data.⁶ We then estimate the predictive regression, equation (1), for each $k = 1, \dots, 10$ for the pseudo-sample, storing the Newey and West (1987) t -statistic for each k . We repeat this procedure 500 times in order to generate an empirical distribution of t -statistics for each k under the null hypothesis of no predictability. In order to test the null hypothesis that $\beta_k = 0$ against the one-sided alternative hypothesis that $\beta_k < 0$, we compute the p -value as the proportion of the bootstrapped t -statistics that are less than the t -statistic calculated for the original data.

Table 1 reports estimation results for equation (1) for the price-dividend and price-earnings ratios and horizons of 1-10 years. Campbell and Shiller (1998) focus on the 1-year and 10-year growth rates in real stock prices, and we generalize the Campbell and Shiller (1988) results by examining all horizons from 1-10

⁵ This allows for a greater degree of serial correlation in z_t than the AR(1) specification assumed by Stambaugh (1986, 1999) in equation (2). Kilian (1999) advises using the AIC to select the AR lag order for z_t . Our bootstrap procedure matches that of Mark (1995). This bootstrap procedure is closely related to the vector error correction bootstrap procedure in Kilian (1999); see equations (9)-(10") in Kilian (1999).

⁶ We set the initial values of the lagged variables to zero, and we include 100 transient start-up observations, which we ultimately discard, in order to randomize the initial values of the lagged variables. We follow this practice for all of the other bootstrap procedures and Monte Carlo simulations described below.

years. Our results confirm those in Campbell and Shiller (1998) in that real stock price growth predictability is detected at long—but not short—horizons. Using conventional significance levels, we find no evidence of predictability from the 1-year to 5-year horizons for the price-dividend ratio and from the 1-year to 7-year horizons for the price-earnings ratio, while we find evidence of predictability at horizons of 6-10 years for the price-dividend ratio and 8-10 years for the price-earnings ratio. As mentioned in the introduction, the pattern evident in the predictive regression results reported in Table 1 is a common finding in the extant literature on the predictability of real stock returns based on valuation ratios.

3. LONG-HORIZON PREDICTABILITY IN A LINEAR FRAMEWORK

As pointed out by Berkowitz and Giorgianni (2001), predictability at long, but not short, horizons presents a logical problem in the linear framework characterized by equations (2) and (3).⁷ While Berkowitz and Giorgianni (2001) are concerned with the finding of long-horizon exchange rate predictability based on monetary fundamentals, their argument also applies to real stock price predictability based on valuation ratios. The intuition behind the predictive regression, equation (1), is that if a stable long-run relationship exists between real stock prices, p_t , and a market fundamental, f_t , such a relationship can be exploited by using $z_t = p_t - f_t$ to predict future movements in stock prices. A stable long-run relationship exists between p_t and f_t when z_t is stationary, as in equation (3).⁸ Suppose that Δp_t and z_t are generated by equations (2) and (3), where we ignore constant terms in all equations without loss of generality. The predictive regression model at any horizon k , equation (1), can then be expressed in terms of the parameters of the DGP as

⁷ As in our bootstrap procedure, we can allow for a greater degree of serial correlation in z_t in equation (3) in what follows, as long as z_t remains stationary.

⁸ This means that Δp_t and Δf_t can be expressed as a linear vector-error correction (VEC) model, as in Kilian (1999) and Berkowitz and Giorgianni (2001). The unit root tests of Ng and Perron (2001) with good size and power indicate that $p_t \sim I(1)$ ($\Delta p_t \sim I(0)$) and $z_t \sim I(0)$ for $z_t = p_t - d_t$ and $z_t = p_t - e_t$ for the Campbell and Shiller (1998) data. The first-order autocorrelation equals 0.74 for $z_t = p_t - d_t$ and 0.66 for $z_t = p_t - e_t$.

$$\Delta p_{t+k}^k = \beta_1 \left(\frac{1-\rho^k}{1-\rho} \right) z_t + \beta_1 \sum_{j=1}^{k-1} \left(\frac{1-\rho^j}{1-\rho} \right) \mu_{t+(k-j)} + \sum_{j=1}^k \varepsilon_{t+j}. \quad (6)$$

It is clear from comparing equations (1) and (6) that

$$\beta_k = \beta_1 \left(\frac{1-\rho^k}{1-\rho} \right) \quad (7)$$

and

$$\varepsilon_{t+k}^k = \beta_1 \sum_{j=1}^{k-1} \left(\frac{1-\rho^j}{1-\rho} \right) \mu_{t+(k-j)} + \sum_{j=1}^k \varepsilon_{t+j}. \quad (8)$$

Equation (7) has an important implication for predictive regressions: if $\beta_1 = 0$, so that there is no predictability at the one-period-ahead horizon, then there is no predictability at any horizon, as $\beta_k = 0$ for all $k \geq 1$. Intuitively, multiple-step-ahead forecasts of a variable are simple extrapolations of the one-step-ahead forecast in a linear framework, so that lack of predictability at short horizons translates directly into lack of predictability at long horizons. Of course, equation (7) also implies that if $\beta_1 \neq 0$, then $\beta_k \neq 0$ for all $k \geq 1$.⁹ In summary, we should have predictability at *all* horizons or *no* horizon in a linear framework characterized by a stable long-run relationship between p_t and f_t as in equations (2) and (3).

Given the Berkowitz and Giorgianni (2001) result, how do we explain the predictive regression test results in Table 1? As discussed in the introduction, the pattern of predictability reported in Table 1 may be due to increased statistical power in finite samples as the horizon lengthens. If this is the case, increased power provides a rationale for running long-horizon predictive regressions, despite the Berkowitz and Giorgianni (2001) result. As noted in the introduction, Berben and van Dijk (1998), Campbell (2001), and Mark and Sul (2002) examine power at long horizons using asymptotic analyses, as well as Monte Carlo experiments for finite samples, and the results are somewhat mixed. The best

⁹ Equation (8) shows that the disturbance term in equation (1), ε_{t+k}^k , is serially correlated when $k > 1$, as noted in Section 2 above.

course of action appears to be analyzing power at different horizons via Monte Carlo experiments based on the sample sizes and data properties relevant to the application at hand.

We begin by checking the size properties of the long-horizon regression tests in Table 1 based on the best-fitting models under the null hypothesis of no predictability for the Campbell and Shiller (1998) data. The results are reported in columns (2) and (4) of Table 2 for predictive regression models based on the price-dividend and price-earnings ratios, respectively. In order to compute the size of our tests at the nominal 10% level, we conduct a Monte Carlo experiment with the data generated under the null hypothesis of no predictability. More specifically, we use the original data to estimate equation (2) with β_1 restricted to zero, as well as an AR model for z_t (with the lag order selected by the AIC). These estimated equations serve as the DGP under the null hypothesis of no predictability. By randomly re-sampling the residuals from these estimated equations (in tandem), we can build up a pseudo-sample of data for Δp_t and z_t that matches the original sample size. We then estimate equation (1) for the pseudo-sample for $k = 1, \dots, 10$ and calculate the t -statistic corresponding to $\hat{\beta}_k$ for each k . Next, we use the same bootstrap procedure described above in Section 2 to generate a p -value for each of the t -statistics. We repeat this process 500 times, giving us 500 p -values for each of the t -statistics corresponding to $k = 1, \dots, 10$. The size reported in columns (2) and (4) of Table 2 is the proportion of the t -statistics at each horizon that have a p -value less than 0.10. From columns (2) and (4) of Table 2, the predictive regression tests are very close to being correctly sized, so size distortions are not an issue for our method of inference.

In order to investigate power, we conduct another set of Monte Carlo simulations based on the best-fitting models under the alternative hypothesis of predictability for the Campbell and Shiller (1998) data. The results are reported in columns (3) and (6) of Table 2 for the price-dividend and price-earnings ratios, respectively. The first step is to estimate equation (2) using the original data without restricting β_1 to zero, so that the DGP is characterized by the predictability of Δp_t . We again use the original data

to estimate an AR model for z_t . These estimated equations serve as the DGP. Note that the DGP thus assumes that $\beta_1 = -0.04$ ($\beta_1 = -0.01$) for $z_t = p_t - d_t$ ($z_t = p_t - e_t$) in equation (2), corresponding to the OLS estimate of β_1 in Table 1. By randomly re-sampling the residuals (in tandem) and using the estimated equations, we can build up a pseudo-sample of data for Δp_t and z_t that matches the original sample size. We then estimate equation (1) for the pseudo-sample for $k = 1, \dots, 10$, calculate the t -statistic corresponding to $\hat{\beta}_k$ for each k , and compute the p -value corresponding to each t -statistic using the bootstrap procedure described in Section 2 above. We repeat this process 500 times, leaving us with 500 p -values for each of the t -statistics corresponding to $k = 1, \dots, 10$. The power reported in columns (3) and (6) of Table 2 is the proportion of the p -values that are less than 0.10 for each k .

From column (3) of Table 2, we see that the power to detect predictability using the price-dividend ratio does not increase with the horizon. In fact, power decreases at long horizons, with the maximum power level of 0.38 achieved at the 1-year horizon. Looking at column (5) of Table 2, we see a similar pattern for the price-earnings ratio: power does not increase at long horizons, and the maximum power level of 0.15 is achieved at the 1-year horizon. Comparing columns (3) and (5) of Table 2, we see that the power of the predictive regression tests based on the price-earnings ratio is considerably lower than the predictive regression tests based on the price-dividends ratio, with the power level close to the nominal size for the tests based on the price-earnings ratio. Overall, the Monte Carlo simulations indicate that there is no evidence of increased power at long horizons that could explain the pattern of predictability reported in Table 1.¹⁰

It is interesting to observe that there is a sense in which the OLS estimates of $\hat{\beta}_k$ in Table 1 are consistent with equation (7) and thus potentially consistent with a linear framework. While we cannot reject the null hypothesis that $\beta_1 = 0$ in equation (2) for the price-dividend ratio in Table 1, we also

¹⁰ Kilian (1999) obtains similar Monte Carlo simulation results for tests of nominal exchange rate predictability based on monetary fundamentals using quarterly data from the post-Bretton Woods period.

would be unable to reject the null hypothesis that, say, $\beta_1 = -0.10$. Plugging this hypothesized value of β_1 and a bias-adjusted estimated ρ of 0.77 into equation (7), we obtain a series of values for β_k , $k = 2, \dots, 10$, implied by a linear framework where $\beta_1 = -0.10$.¹¹ If we compare this series of β_k values to the series of OLS estimates in column (2) of Table 1, we could not reject the null hypothesis that each value of $\hat{\beta}_k$ in the series of OLS estimates equals the corresponding element of the series of β_k values implied by $\beta_1 = -0.10$ in equation (7). From this perspective, the estimation results in Table 1 could be construed as being consistent with a linear framework. However, the power results reported in column (4) of Table 2 show that it is still difficult to explain the pattern of rejections in Table 1 for the price-dividend ratio in a linear framework where $\beta_1 = -0.10$. The power results in column (4) of Table 2 are based on the linear DGP defined by equations (2) and (3) with β_1 taking on an assumed value of -0.10 . Similar to the case where $\beta_1 = -0.04$, the power of the predictive regression test does not increase with the horizon; in fact, it decreases monotonically with the horizon. These simulation results indicate that it is very unlikely that we would obtain the pattern of rejections observed in Table 1 for the price-divided ratio if the data were generated by a linear framework where $\beta_1 = -0.10$. From column (7) of Table 2, we also see that power does not increase with the horizon for the price-earnings ratio when we assume that the DGP is characterized by a linear framework where $\beta_1 = -0.10$. These additional Monte Carlo simulations reinforce the notion that it is difficult to account for the pattern of predictability observed in Table 1 in a linear framework characterized by equations (2) and (3).¹²

4. AN ALTERNATIVE NONLINEAR FRAMEWORK

In the previous section, we found that the evidence of real stock price predictability at long, but not short, horizons based on price-dividend and price-earnings ratios is difficult to explain in a linear framework.

¹¹ The implied values of β_k are $-0.18, -0.24, -0.28, -0.31, -0.34, -0.36, -0.38, -0.39, -0.40$ for $k = 2, \dots, 10$.

¹² We also conducted Monte Carlo simulations where $\beta_1 = -0.20$, and we continued to fail to find evidence that power increases with the horizon.

However, Berkowitz and Giorgianni (2001), Kilian (1999), and Kilian and Taylor (2003) observe that such a pattern of predictability is possible in a nonlinear world. In fact, in the case of long-horizon exchange rate predictability based on monetary fundamentals, Kilian (1999, p. 507) interprets the evidence of predictability at long, but not short, horizons as indirect evidence of a nonlinear DGP:

Perhaps the most interesting finding to emerge from this study is that the linear VEC model framework underlying the existing long-horizon regression tests is likely to be misspecified. In particular, the observed pattern of p -values across forecast horizons in the empirical study is inconsistent with the size and power results for the linear VEC model. This fact is suggestive of a non-linear DGP.

In this section, we undertake a preliminary investigation of real stock price predictability in a nonlinear framework along the lines of Kilian and Taylor (2003). We begin by postulating a DGP for $z_t = p_t - d_t$ or $z_t = p_t - e_t$ that allows for nonlinear mean-reversion in these variables. Kilian and Taylor (2003) argue that nonlinear mean-reversion better describes asset price movements in a world of noise trading and risky arbitrage modeled along the lines of De Long, et al. (1990a,b).¹³ In models with noise traders and arbitrageurs, noise traders' demand for assets is affected by beliefs that are not fully justified by news about fundamentals; for example, noise traders may follow the advice of technical analysts. In contrast, arbitrageurs form fully rational expectations about the returns to holding an asset and can potentially profit from the mistaken beliefs of noise traders. However, arbitrage is risky in these models. The problem is that the mistaken beliefs of noise traders may cause asset prices to deviate from their underlying fundamentals for considerable periods of time, even though they will ultimately return to a level in line with the fundamentals. An arbitrageur may have to borrow to execute trades or be compared to other financial advisors, and if the mispricing persists, the arbitrageur can suffer serious losses or fare poorly in comparisons to other advisors. These frictions prevent the arbitrageurs from quickly eliminating deviations in asset prices from their underlying fundamentals. Kilian and Taylor (2003) hypothesize that

¹³ See Shleifer (2000) for an overview of models with investor sentiment and risky arbitrage in behavioral finance.

the risk to arbitrage decreases as the asset becomes increasingly overvalued or undervalued, leading to disproportionately quicker adjustment toward the equilibrium. Thus, arbitrageurs will exert strong pressure on the asset's price to move in line with the underlying fundamentals when deviations from the fundamentals are large—as they are more confident that their beliefs will be justified—while smaller deviations are likely to be more persistent.

Kilian and Taylor (2003) propose a parsimonious ESTAR model specification for nominal exchange rate deviations from purchasing power parity fundamentals that incorporates the notion of risky arbitrage. We consider a similar ESTAR specification for the price-dividend or price-earnings ratio. Our version of the Kilian and Taylor (2003) model takes the form,

$$z_t - \mu_z = \{\exp[\gamma(z_{t-1} - \mu_z)^2]\}(z_{t-1} - \mu_z) + u_t, \quad (9)$$

where μ_z is the mean of z_t and u_t is a white-noise disturbance term ($u_t \sim iid(0, \sigma_2)$).¹⁴ The transition function for the ESTAR model, equation (9), is defined by $\exp[\gamma(z_{t-1} - \mu_z)^2]$, and it has a straightforward economic interpretation: assuming that $\gamma < 0$ (as is the case for our data), mean-reversion will be stronger the larger (in absolute value) the deviation of p_t from f_t . Essentially, the transition function allows the value of the AR coefficient to decrease as the size of the deviation increases, so that mean-reversion is faster for larger deviations. Note that as the deviation approaches zero, the AR process in equation (9) approaches a unit root process. (The value of the transition function approaches zero as the deviation approaches $\pm \infty$.) Kilian and Taylor (2003) find that a parsimonious ESTAR model fits exchange rate deviations from purchasing power parity fundamentals well over the post-Bretton Woods period. We estimate equation (9) for $z_t = p_t - d_t$ and $z_t = p_t - e_t$ using the

¹⁴ It is not our intent to undertake an extensive analysis of how well different classes of nonlinear models fit the Campbell and Shiller (1998) data. Instead, we investigate a parsimonious nonlinear model with a straightforward economic interpretation and examine whether it can help us understand the predictive regression test results reported in Table 1.

Campbell and Shiller (1998) data.¹⁵

We estimate equation (9) for each z_t using nonlinear least squares (NLLS), and the results are reported in Table 3. It is important to note that care must be taken when assessing the significance of the NLLS estimate of γ , $\hat{\gamma}$, as z_t is nonstationary under the null hypothesis that $\gamma=0$. Following Taylor and Peel (2000) and Kilian and Taylor (2003), we use a bootstrap procedure to calculate a p -value for the NLLS t -statistic corresponding to $\hat{\gamma}$. For each of 500 bootstrap iterations, we generate a pseudo-sample by re-sampling the residuals from a random walk with drift model for z_t . We estimate equation (9) for each of the 500 pseudo-samples, and we store the t -statistics in order to form an empirical distribution of t -statistics. In order to compute a p -value for a test of the null hypothesis that $\gamma=0$ against the alternative hypothesis that $\gamma<0$, we calculate the proportion of the simulated t -statistics that are less than the t -statistic corresponding to the original data. From column (2) of Table 3, we see that the NLLS estimate of γ is negative for each z_t . From column (3), we see that $\hat{\gamma}$ is significant at the 5% level for $z_t = p_t - d_t$ and at the 1% level for $z_t = p_t - e_t$.

We can glean more insight into the nonlinear nature of $z_t = p_t - d_t$ and $z_t = p_t - e_t$ by examining the estimated conditional expectation function, $E(z_t - \hat{\mu}_z | z_{t-1} - \hat{\mu}_z)$, which is plotted for each valuation ratio in Figure 2. As noted above, the transition function, $\exp[\hat{\gamma}(z_{t-1} - \hat{\mu}_z)^2]$, can be interpreted as the value of the AR coefficient, which corresponds to the slope of the conditional expectation function. From Figure 2, we see that the slope of the conditional expectation function decreases as $z_{t-1} - \hat{\mu}_z$ increases in absolute value, so that mean-reversion is faster for larger deviations.

In order to more directly compare a linear to a nonlinear specification for z_t , we also tested the null hypothesis of a linear AR model specification against the alternative hypothesis of an ESTAR specification using the Lagrange-multiplier test of Granger and Teräsvirta (1993) and Teräsvirta (1994).

¹⁵ We include a test for any remaining autocorrelation in the residuals of a nonlinear regression model due to Eitrheim and Teräsvirta (1996). Supporting a lag order of one in equation (9) for each z_t , the null hypothesis of no autocorrelation in the residual term in equation (9) cannot be rejected (see column (7) of Table 3).

Given the parsimonious specification in equation (9), this entails regressing z_t on a constant, z_{t-1} , z_{t-1}^2 , and z_{t-1}^3 and testing the joint significance of the coefficients on z_{t-1}^2 , and z_{t-1}^3 . For the price-dividend ratio, we can reject the null hypothesis of linearity at the 5% level using either the F -statistic or the χ^2 -statistic form of the test (p -value of 0.02 for each statistic). The Lagrange-multiplier test thus provides support for an ESTAR specification over a linear AR specification for the price-dividend ratio. For the price-earnings ratio, neither statistic is significant at the 10% level (p -value of 0.23 for each statistic), so that the test does not support an ESTAR over a linear AR specification for the price-earnings ratio at standard significance levels.

If the underlying DGP is nonlinear, the p -values obtained for the predictive regression tests in Table 1 under the assumption of a linear DGP become invalid. In light of this, we again estimate the predictive regression model, equation (1), for $z_t = p_t - d_t$ and $z_t = p_t - e_t$ and $k=1, \dots, 10$, this time accounting for a nonlinear adjustment to the fundamentals using the modified bootstrap methodology suggested by Kilian and Taylor (2003). We assume the following DGP under the null hypothesis of no predictability in order to capture the nonlinear dynamics in z_t ,

$$\Delta p_t - \mu_{\Delta p} = u_{1,t}, \quad (10)$$

$$z_t - \mu_z = \{\exp[\gamma(z_{t-1} - \mu_z)^2]\}(z_{t-1} - \mu_z) + u_{2,t}, \quad (11)$$

where the disturbance vector, $(u_{1,t}, u_{2,t})'$, is independently and identically distributed. This process can be estimated by NLLS, and we can re-sample the residuals (in tandem) in order to generate a pseudo-sample of observations for Δp_t and z_t matching the original sample size. We compute and store the t -statistic corresponding to $\hat{\beta}_k$ for $k=1, \dots, 10$ in equation (1) for the pseudo-sample. We repeat this process 500 times in order to generate an empirical distribution of t -statistics for each k and we calculate a p -value for each k as the proportion of the simulated t -statistics that are less than the t -statistic calculated using the original data. This allows us to assess the significance of $\hat{\beta}_k$ in equation (1) under

the assumption of a nonlinear DGP.¹⁶

Estimation results for predictive regression models based on $z_t = p_t - d_t$ and $z_t = p_t - e_t$ are reported in Table 4. Note that the $\hat{\beta}_k$ estimates and their associated t -statistic are computed in the same manner as Table 1; the only difference in Table 4 is that we use the modified bootstrap procedure, which assumes that z_t follows an ESTAR process, to generate p -values and assess statistical significance. From column (3) of Table 4, we see significant evidence of real stock price predictability at horizons of 5-10 years based on the price-dividend ratio, and from column (5) of Table 4, we see significant evidence of predictability at horizons of 6-10 years based on the price-earnings ratio. Overall, we still find predictability at long, but not short, horizons when we assume a nonlinear DGP for either valuation ratio.

We now proceed to show that this evidence of increased long-horizon predictability can be attributed to power gains at longer horizons, as opposed to size distortions, in a nonlinear framework, at least for the price-dividend ratio. As a first step, we ensure that the modified bootstrap procedure that we use to make inferences in Table 4 has good size properties. In order to investigate this, we assume that the DGP for Δp_t and z_t is given by equations (10) and (11). Armed with estimates of the parameters in equations (10) and (11), we re-sample from the fitted residuals (in tandem) in order to build up a pseudo-sample of observations for Δp_t and z_t . We then estimate equation (1) for $k=1, \dots, 10$ and use the modified bootstrap procedure to calculate a p -value corresponding to $\hat{\beta}_k$ for each k . We repeat this process 500 times, and measure the empirical size at the nominal 10% level as the proportion of p -values that are less than 0.10. From columns (2) and (4) of Table 5, we see that the modified bootstrap procedure works quite well, as the empirical size of each test statistic is very close to the nominal size of 0.10.

We conduct a final set of Monte Carlo simulations in order to investigate power in a nonlinear framework. As pointed out by Kilian and Taylor (2003), the power of the test will in general depend on

¹⁶ See the Appendix in Kilian and Taylor (2003) for a detailed description of the modified bootstrap methodology.

the specific form of the alternative model, and it will be difficult to identify the actual underlying nonlinear process at work in the data. We follow Kilian and Taylor (2003) and assume a linear process for the fundamental. More specifically, we use a version of Kilian and Taylor's (2001) DGP 3. In order to arrive at a relatively parsimonious model for the dividends or earnings process, we select the number of lags of each variable to include in the DGP for the fundamental using a general-to-specific procedure, where the general specification included four lags of Δd_t or Δe_t and four lags of Δp_t . Our assumed process for d_t is

$$\Delta d_t = \gamma_0 + \gamma_1 \Delta d_{t-1} + \gamma_2 \Delta p_{t-1} + \gamma_3 \Delta p_{t-2} + \gamma_4 \Delta p_{t-3} + u_{1t}, \quad (12)$$

while our assumed process for e_t is

$$\Delta e_t = \gamma_0 + \gamma_1 \Delta e_{t-4} + \gamma_2 \Delta p_{t-1} + u_{1t}. \quad (13)$$

The DGP under predictability is thus given by equations (12) and (11) for $z_t = p_t - d_t$ and equations (13) and (11) for $z_t = p_t - e_t$. Note that equations (12) and (11) or equations (13) and (11) are sufficient to determine p_t , as we can back out p_t using $p_t = z_t + d_t$ or $p_t = z_t + e_t$. After estimating equations (12) and (11) or equations (13) and (11), we build up a pseudo-sample of data for Δp_t and z_t by re-sampling from the fitted residuals (in tandem). We then estimate equation (1) for the pseudo-sample for $k=1, \dots, 10$ and conduct tests at the 10% significance level. We repeat this process 500 times, and columns (3) and (5) of Table 5 report the power of the tests.

An interesting result emerges in column (3) of Table 5 for predictive regression tests based on the price-dividend ratio. Unlike the situation in column (3) of Table 2, where we assume a linear DGP, power does increase at long horizons for tests based on the price-dividend ratio when we assume a nonlinear DGP. Indeed, the power figures reported in column (3) of Table (5) follow a pattern that closely matches the pattern of p -values in column (3) of Table 4: the power figures in column (3) of Table 5 increase monotonically as the horizon increases from 1-5 years, which can explain the nearly monotonic decrease in the corresponding p -values in column (3) of Table 4. While we could not account

for the pattern of predictability in the Campbell and Shiller (1998) data using a linear framework, we apparently can account for the pattern of predictability when we assume that the price-dividend ratio is generated by an ESTAR process. Of course, one could consider a host of alternative nonlinear DGPs for the price-dividend ratio, but this is beyond the scope of the present paper. Nevertheless, we find that a relatively simple nonlinear DGP, based on a parsimonious ESTAR specification for the price-dividend ratio, is able to explain the pattern of stock price predictability in the Campbell and Shiller (1998) data, and this is at least suggestive that nonlinearities are important in understanding long-horizon stock price predictability.

Turning to column (5) in Table 5, we see that power does increase slightly as the horizon increases for predictive regression tests based on the price-earnings ratio. However, the increase in power is much less noticeable than in column (3) of Table 5, and it does not appear marked enough to provide an explanation for the pattern of p -values in column (5) of Table 4. Overall, the results in Table 5 indicate that a nonlinear DGP based on an ESTAR specification for the price-dividend ratio provides a plausible explanation for the pattern of stock price predictability in the data, while a nonlinear DGP based on an ESTAR specification for the price-earnings ratio is less successful.

5. CONCLUSION

In this paper, we re-examine predictive regression models of real stock prices based on valuation ratios. Using the annual data for 1872-1997 from Campbell and Shiller (1998), and in line with the extant literature, we find that the price-dividend and price-earnings ratios have significant ability to predict real stock price growth at long, but not short, horizons. A possible explanation for the pattern of stock price predictability in the data is increased statistical power to detect predictability at long horizons. We explore this potential explanation using Monte Carlo simulations. Our simulation results demonstrate that power does not increase at long horizons (in fact, it decreases), so that the pattern of stock price predictability in the data is difficult to explain in a linear framework.

Following Kilian (1999) and Kilian and Taylor (2003), we argue that the observed pattern of real stock price predictability could be construed as evidence for nonlinearities in the underlying DGP, as a nonlinear framework does not preclude predictability at long, but not short, horizons. In light of this, we employ a version of the parsimonious ESTAR model proposed by Kilian and Taylor (2003). This model allows for nonlinear mean-reversion in the price-dividends and price-earnings ratios, and we find that this nonlinear model fits the data reasonably well. We also re-assess the significance of the predictive regression tests by computing p -values from a modified bootstrap procedure based on a nonlinear DGP, and we continue to find significant evidence of stock price predictability at long, but not short, horizons. Finally, we show that the evidence of increased long-horizon predictability can plausibly be attributed to power gains arising from nonlinearities in the DGP for the price-dividend ratio. Thus, similar to Kilian and Taylor (2003) in the context of exchange rate predictability, we show that nonlinear behavior in the price-dividend ratio provides a plausible explanation of the pattern of predictability found in annual data for the price-dividend ratio and stock price growth. There is less evidence that an ESTAR specification for the price-earnings ratio can explain long-horizon stock price predictability.

Our investigation of ESTAR models shows that this nonlinear specification provides a plausible explanation for long-horizon stock price predictability, at least for predictive regression tests based on the price-dividend ratio. In related work, Gallagher and Taylor (2001) recently find evidence of ESTAR dynamics in the price-dividend ratio using quarterly data from 1926-1997.¹⁷ The results in the present paper indicate that further analysis of nonlinear model specifications for valuation ratios is warranted and may help researchers to better understand long-horizon stock price predictability, especially given the difficulty in accounting for long-horizon predictability in a linear framework.

¹⁷ Also see the nonlinear models in Qi (1999) and McMillan (2001). Gallagher and Taylor (2001) do not consider predictive regression tests. It would be interesting in future research to analyze stock price predictability in a nonlinear framework for the Gallagher and Taylor (2001) data.

REFERENCES

- Andrews DWK. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**: 817-858.
- Andrews DWK, Monohan JC. 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* **60**: 953-966.
- Berben R-P, van Dijk D. 1998. Does the absence of cointegration explain the typical findings in long horizon regressions? Report 9814, Econometrics Institute, Erasmus University.
- Berkowitz J, Giorgianni L. 2001. Long-horizon exchange rate predictability? *Review of Economics and Statistics* **83**: 81-91.
- Campbell JY. 1999. Asset prices, consumption, and the business cycle. In *Handbook of Macroeconomics, Volume 1*, Taylor J, Woodford M (eds.). Amsterdam: North-Holland.
- Campbell JY. 2000. Asset pricing at the millennium. *Journal of Finance* **55**: 1515-1567.
- Campbell JY. 2001. Why long horizons? A study of power against persistent alternatives. *Journal of Empirical Finance* **8**: 459-491.
- Campbell JY, Shiller RJ. 1988. Stock prices, earnings, and expected dividends. *Journal of Finance* **43**: 661-676.
- Campbell JY, Shiller RJ. 1998. Valuation ratios and the long-run stock market outlook. *Journal of Portfolio Management* Winter: 11-26.
- De Long JB, Shleifer A, Summers LH, Waldmann, RJ. 1990a. Noise trader risk in financial markets. *Journal of Political Economy* **98**: 703-738.
- De Long JB, Shleifer A, Summers LH, Waldmann, RJ. 1990b. Positive investment feedback strategies and destabilizing rational speculation. *Journal of Finance* **45**: 379-395.
- Eitrheim O, Teräsvirta T. 1996. Testing the adequacy of smooth transition autoregressive models. *Journal of Econometrics* **74**: 59-76.
- Fama EF. 1991. Efficient capital markets: II. *Journal of Finance* **46**: 1575-1617.
- Fama EF, French KR. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* **22**: 3-25.
- Fama, EF, French KR. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* **25**: 23-49.
- Gallagher LA, Taylor MP. 2001. Risky arbitrage, limits of arbitrage, and nonlinear adjustment in the dividend-price ratio. *Economic Inquiry* **39**: 524-536.

Granger CWJ, Teräsvirta T. 1993. *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

Hurwicz L. 1950. Least squares bias in time series. In *Statistical Inference in Dynamic Economic Models*, Koopmans, T (ed). Wiley: New York; 365-383.

Kendall M. 1954. Note on bias in the estimation of autocorrelation. *Biometrika* **41**: 403-404.

Kilian L. 1999. Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions? *Journal of Applied Econometrics* **14**: 491-510.

Kilian, L, Taylor MP. 2003. Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics* **60**: 85-107.

Kirby C. 1997. Measuring the predictable variation in stock and bond returns. *Review of Financial Studies* **10**: 579-630.

Kothari SP, Shanken J. 1997. Book-to-market, dividend yield, and expected market returns: a time series analysis. *Journal of Financial Economics* **44**: 169-203.

Mankiw NG, Shapiro, MD. 1986. Do we reject too often? *Economic Letters* **20**: 134-145.

Mark NC, 1995. Exchange rates and fundamentals: evidence on long-horizon predictability. *American Economic Review* **85**: 201-218.

Mark NC, Sul D. 2002. Asymptotic power advantages of long-horizon regressions. Manuscript, Department of Economics, Ohio State University.

McMillan DG. 2001. Nonlinear predictability of stock market returns: evidence from nonparametric and threshold models. *International Review of Economics and Finance* **10**: 353-368.

Nelson CR, Kim MJ. 1993. Predictable stock returns: the role of small sample bias. *Journal of Finance* **48**: 641-661.

Newey W, West KJ. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **10**: 129-162.

Ng S, Perron P. 2001. Lag-length selection and the construction of unit root tests with good size and power. *Econometrica* **69**: 1519-1554.

Qi M. 1999. Nonlinear predictability of stock returns using financial and economic variables. *Journal of Business and Economic Statistics* **17**: 419-429.

Shaman P, Stine RA. 1988. The bias of autoregressive coefficient estimators. *Journal of the American Statistical Association* **83**: 842-848.

Shleifer A. 2000. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford: Oxford University Press.

Stambaugh RF. 1986. Biases in regressions with lagged stochastic regressors. Working Paper No. 156, Graduate School of Business, University of Chicago.

Stambaugh RF. 1999. Predictive regressions. *Journal of Financial Economics* **54**: 375-421.

Taylor MP, Peel DA. 2000. Nonlinear adjustment, long-run equilibrium and exchange rate fundamentals. *Journal of International Money and Finance* **19**: 33-53.

Teräsvirta T. 1994. Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* **89**: 208-218.

Table 1. Estimation results for the predictive regression model, $\Delta p_{t+k}^k = \alpha_k + \beta_k z_t + \varepsilon_{t+k}^k$, under the assumption of a linear data-generating process.

(1)	(2)	(3)	(4)	(5)
	$z_t = p_t - d_t$ (1872-1997)		$z_t = p_t - e_t$ (1872-1997)	
Horizon (k)	$\hat{\beta}_k$	t -statistic	$\hat{\beta}_k$	t -statistic
1 year	-0.04	-0.83 [0.40]	-0.01	-0.24 [0.49]
2 years	-0.15	-1.58 [0.18]	-0.08	-0.79 [0.30]
3 years	-0.19	-1.48 [0.28]	-0.10	-0.78 [0.31]
4 years	-0.32	-2.04 [0.19]	-0.15	-0.99 [0.28]
5 years	-0.43	-2.90 [0.11]	-0.23	-1.77 [0.16]
6 years	-0.47	-3.19 [0.07]	-0.31	-2.13 [0.11]
7 years	-0.52	-3.70 [0.07]	-0.38	-2.24 [0.12]
8 years	-0.64	-5.39 [0.02]	-0.51	-2.91 [0.05]
9 years	-0.68	-4.93 [0.04]	-0.63	-3.93 [0.03]
10 years	-0.70	-4.20 [0.06]	-0.73	-4.29 [0.02]

Notes: bootstrapped p -values given in brackets; bold entries indicate significance at the 10% level according to the bootstrapped p -values.

Table 2. Monte Carlo simulation results for the predictive regression tests under the assumption of a linear data-generating process.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$z_t = p_t - d_t$			$z_t = p_t - e_t$		
	Power			Power		
Horizon (k)	Size	$\beta_1 = -0.04$	$\beta_1 = -0.10$	Size	$\beta_1 = -0.01$	$\beta_1 = -0.10$
1 year	0.09	0.38	0.86	0.11	0.15	0.81
2 years	0.10	0.34	0.83	0.11	0.14	0.75
3 years	0.09	0.35	0.79	0.12	0.13	0.66
4 years	0.11	0.32	0.78	0.11	0.13	0.60
5 years	0.10	0.30	0.76	0.10	0.13	0.52
6 years	0.11	0.29	0.74	0.10	0.12	0.46
7 years	0.10	0.30	0.72	0.10	0.12	0.43
8 years	0.10	0.29	0.71	0.09	0.11	0.38
9 years	0.10	0.30	0.69	0.09	0.12	0.34
10 years	0.10	0.29	0.69	0.11	0.12	0.33

Notes: reported size and power are based on 500 Monte Carlo replications with 500 bootstrapped replications per Monte Carlo replication.

Table 3. Nonlinear least squares estimation results for the parsimonious ESTAR model, $z_t - \mu_z = \{\exp[\gamma(z_{t-1} - \mu_z)^2]\}(z_{t-1} - \mu_z) + u_t$.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Valuation ratio	$\hat{\gamma}$	t -statistic	R^2	s^a	DW	AR(1) ^b
$z_t = p_t - d_t$ (1872-1997)	-0.93	-2.92 [0.02]	0.58	0.20	2.00	0.31 [0.57]
$z_t = p_t - e_t$ (1872-1997)	-1.23	-3.42 [0.00]	0.42	0.24	2.11	0.65 [0.42]

Notes: bootstrapped p -values given in parentheses; bold entries indicate significance at the 10% level according to the bootstrapped p -values.

^aRegression standard error.

^bOne-sided (upper-tail) test of the null hypothesis that there is no first-order autocorrelation in the regression disturbance term.

Table 4. Estimation results for the predictive regression model, $\Delta p_{t+k}^k = \alpha_k + \beta_k z_t + \varepsilon_{t+k}$, under the assumption of a nonlinear data-generating process.

(1)	(2)	(3)	(4)	(5)
	$z_t = p_t - d_t$ (1872-1997)		$z_t = p_t - e_t$ (1872-1997)	
Horizon (k)	$\hat{\beta}_k$	t -statistic	$\hat{\beta}_k$	t -statistic
1 year	-0.04	-0.83 [0.32]	-0.01	-0.24 [0.46]
2 years	-0.15	-1.58 [0.18]	-0.08	-0.79 [0.31]
3 years	-0.19	-1.48 [0.22]	-0.10	-0.78 [0.34]
4 years	-0.32	-2.04 [0.13]	-0.15	-0.99 [0.27]
5 years	-0.43	-2.90 [0.05]	-0.23	-1.77 [0.13]
6 years	-0.47	-3.19 [0.04]	-0.31	-2.13 [0.09]
7 years	-0.52	-3.70 [0.04]	-0.38	-2.24 [0.09]
8 years	-0.64	-5.39 [0.00]	-0.51	-2.91 [0.04]
9 years	-0.68	-4.94 [0.02]	-0.63	-3.93 [0.01]
10 years	-0.70	-4.20 [0.05]	-0.73	-4.29 [0.01]

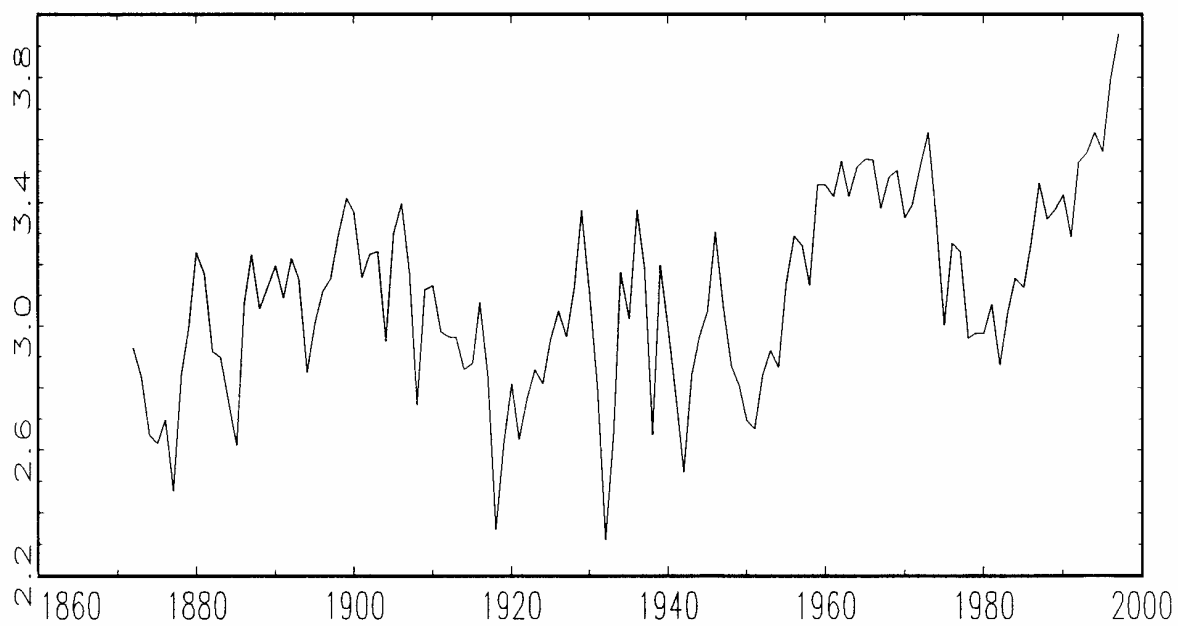
Note: bootstrapped p -values given in brackets; bold entries indicate significance at the 10% level according to the bootstrapped p -values

Table 5. Monte Carlo simulation results for the predictive regression tests under the assumption of a nonlinear data-generating process.

(1)	(2)	(3)	(4)	(5)
	$z_t = p_t - d_t$		$z_t = p_t - e_t$	
Horizon (k)	Size	Power	Size	Power
1 year	0.10	0.49	0.09	0.89
2 years	0.10	0.70	0.11	0.95
3 years	0.09	0.85	0.10	0.97
4 years	0.08	0.92	0.10	0.99
5 years	0.09	0.95	0.09	0.99
6 years	0.09	0.95	0.09	0.98
7 years	0.09	0.95	0.10	0.96
8 years	0.09	0.94	0.10	0.95
9 years	0.08	0.92	0.10	0.93
10 years	0.09	0.90	0.09	0.90

Notes: reported size and power are based on 500 Monte Carlo replications with 500 bootstrapped replications per Monte Carlo replication.

A. Price-dividend ratio



B. Price-earnings ratio

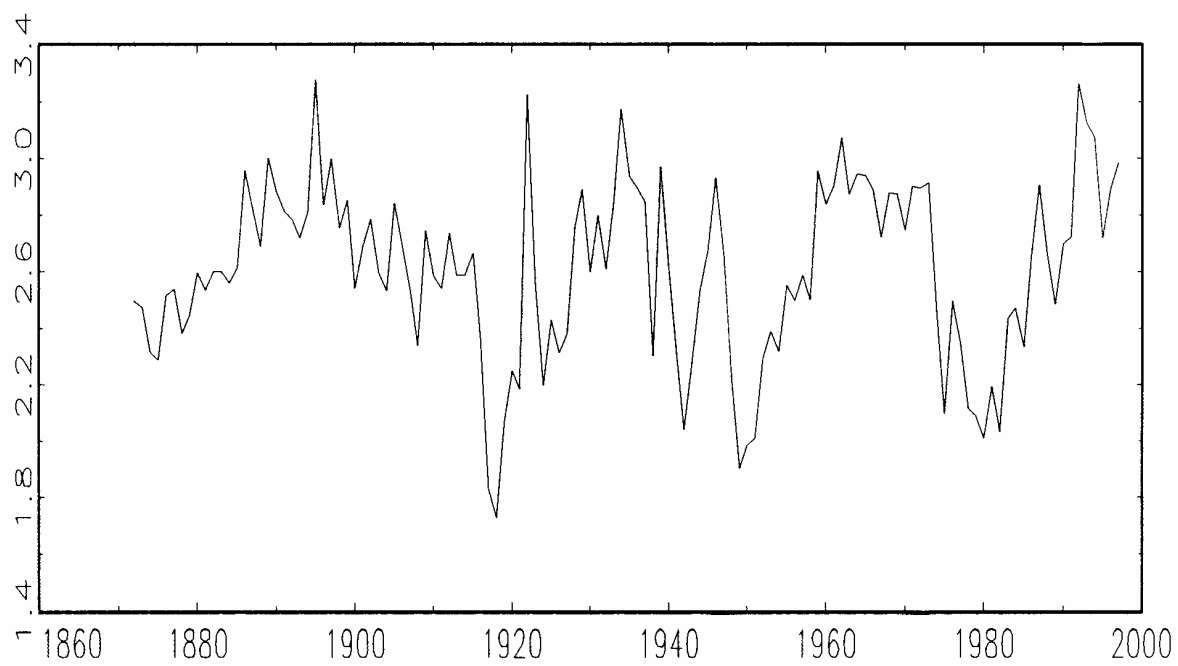


Figure 1. Valuation ratios (log-levels)

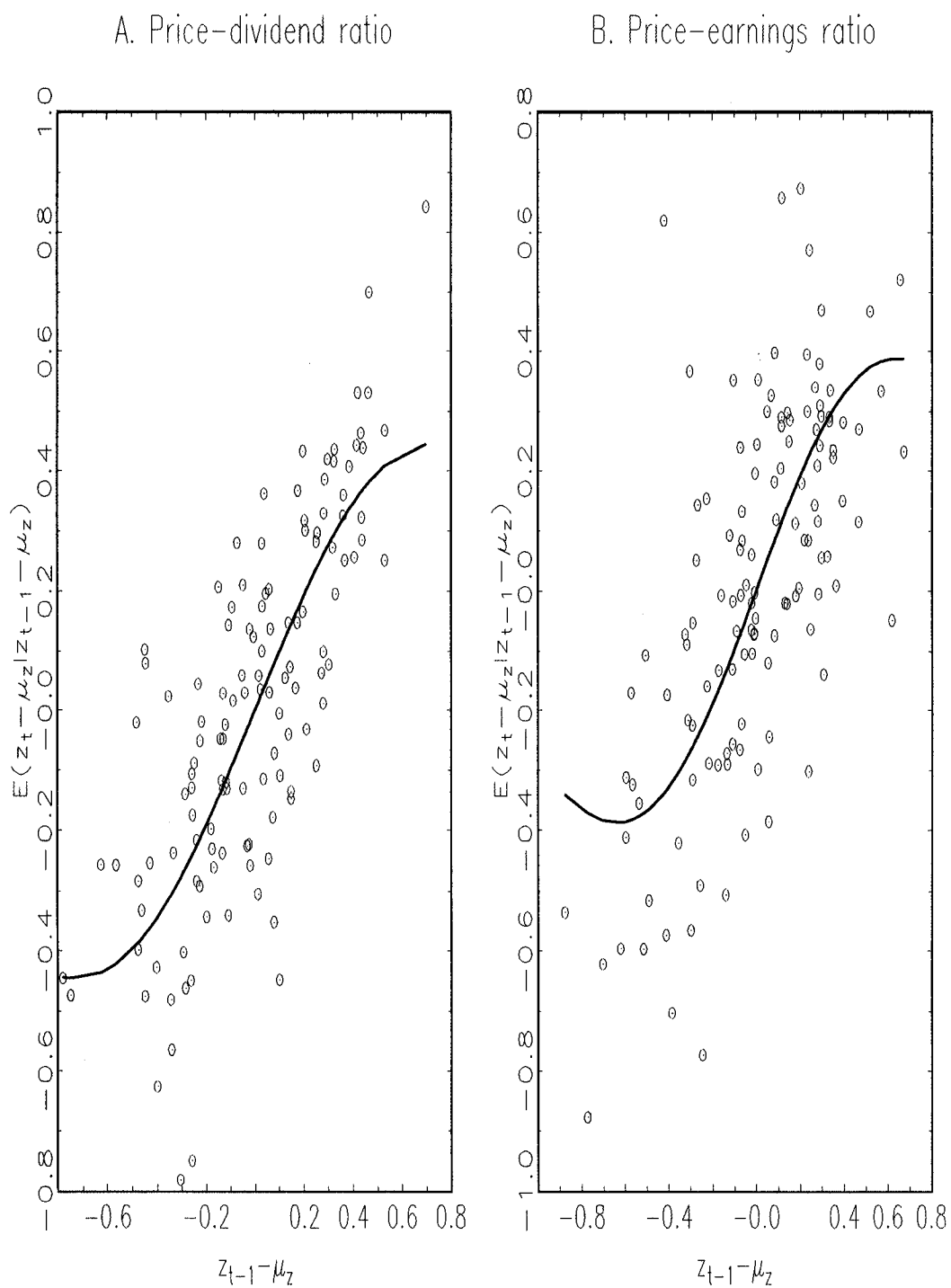


Figure 2. Conditional expectation of $z_t - \mu_z$ given $z_{t-1} - \mu_z$ for the estimated ESTAR model and scatterplot of the data for the price-dividend ratio (Panel A) or price-earnings ratio (Panel B)