JETEM

Journal of Economic Theory and Econometrics

The Journal of the Korean Econometric Society

(계 량 경 제 학 보)

March 2008

Volume 19. Issue 1. pp.17-48

Long-horizon stock return predictability test with a nonlinear nonparametric bootstrap method

Ho-Jin Lee*

Abstract A nonparametric bootstrap procedure with an LSTAR modeling of the valuation ratio is applied to the continuously compounded real stock return and the log of the price-dividend process. The empirical distribution of the test statistics shows that the evidence for a stock return predictability weakens when we take care of nonlinearity dynamics in the regressor. We split the sample into two regimes and implement the long-horizon predictability tests. Results show that the stock return is predictable in the stationary regime, while the test statistic under the null of unpredictability is insignificant in the non-stationary regime.

Keywords Structural Change, Price-Dividend Ratio, Long-Horizon Regression Model, Bootstrapping, LSTAR Model(Logistic Smooth Transition Autoregressive Model)

JEL Classification C22, G14

* Department of Business Administration, Myongji University, Namgajwa-dong 50-3, Seodaemun-gu, Seoul 120-728, Korea, Tel.: +82-2-300-0754, E-mail address: hlee07@mju.ac.kr

Received March 26, 2007, Revised August 23, 2007, Accepted March 13, 2008

1. Introduction

The spurious regression fit due to a persistent regressor suggests that the significant stock return predictability might be a product of size distortion rather than the predictability of financial fundamentals. When we test the statistical significance of prediction regression, we confronted the same kind of econometric problems that earlier researchers in this area had to deal with. Two notable issues are (1) The intercept and the slope coefficients of the regression are biased in finite samples due to the highly autocorrelated forecasting variable. (2) The overlapping forecast errors have an autocorrelation of at least $k-1$ degrees when we extend the forecasting horizons to k, and as a result of this, the construction of the functions of forecast errors such as the standard errors of regression coefficients would be complicated. Also, as pointed out in Elliott and Stock (1994), even if the forecasting variable is stationary, first order asymptotics is a poor approximation in finite samples because the first order asymptotic distribution that will be used for an inferential purpose shows a discontinuity between a stationary and a nonstationary process. As a way to circumvent this problem, the local-to-unity framework has been used to derive the asymptotic null distribution of t-statistic, where the explanatory variable is assumed to follow an autoregression with a root near to unity. Torous, Valkanov, and Yan (2004) find an evidence for predictability at short horizons but not at long horizons. Lanne (2002), on the other hand, finds no evidence that the stock returns are predictable using a highly persistent forecasting variable. Campbell and Yogo (2006), using the robust test to a persistent problem, find that some financial variables predict returns at all frequencies in the full sample. In other studies, the simulation method in order to get the empirical distribution of test statistic has been explored. To account for bias and size distortions, Nelson and Kim (1993) model the variables as a VAR under the null of no predictable stock return, and then generate the artificial histories of them using the estimated VAR and the randomized sequences of historical residuals. Repeated regressions for the artificial histories give us the empirical sampling distribution of test statistic. In his pathfinding study, Mark (1995), using parametric and nonparametric bootstrap long-horizon regression tests, builds bootstrap distributions upon estimated values of the restricted vector autoregression (VAR) that embodies the null hypothesis that the exchange rate is unpredictable. Recently, Berkowitz and Giorgianni (2001) and Berben and van Dijk (1998) show by means of a Monte Carlo simulation that the empirical critical values are sensitive to the null hypothesis. They show that, under the null of no cointegration, the asymptotic distributions of coefficient estimates are increasing in the estimation horizon despite being zero in population. Kilian (1999), taking the error correction relationship of exchange rate and monetary fundamental into account, modifies Mark's bootstrap DGP and invalidates the DGP of linear error correction valuation ratio in order to explain the long-horizon predictability. In a recent study by Rapach and Wohar (2005), following Kilian and Taylor (2001), they include the ESTAR procedure for valuation ratio in the data generating procedure.

In this paper, we will take advantage of the nonlinear time series setup for analyzing stock return predictability by including the nonlinear modeling into the nonparametric bootstrap data generating process. Lee (2006) investigates the nonlinear modeling of the price-dividend ratio in detail. In order to test the linearity null hypothesis and select the transition variable, the author computes the *LM* statistics and the associated *p*-values. Lee (2006) performs the following linearity tests allowing up to 4 AR lags for the transition variable. For each *LM* statistic, he computes the *LM*-type tests of Luukkonen *et al*. (1988) against the LSTAR alternative and the tests of Granger and Teräsvirta (1993) and Escribano and Jorda (1999) against the ESTAR, using least squares based (LS), and heteroskedasticity and autocorrelaion consistent versions of the tests outlined in Wooldridge (1991). The tests show that linearity is rejected in favor of the STAR-type nonlinearity and the appropriate transition variable has been selected as the lag order 3. Lee (2006) also exercises the sequential procedure suggested by Teräsvirta (1994) for the selection of the transition function and the test results select a logistic function as the transition function. For the detailed results, see Lee (2006).

By using the nonlinearity of regressor, we appreciate the performance of standard models in providing forecastability and evaluate whether the nonlinear regressor models can outperform a random walk in various forecasting horizons. The remainder of paper is organized as follows. Section 2 contains the long-horizon predictability tests using the linear valuation model. In order to correct a size distortion, a nonparametric bootstrap method will be used to construct *p*-values. Section 3 analyzes the long-horizon regressions using the modified bootstrap method which embodies the nonlinear modeling of valuation ratio. Different types of DGPs in the nonlinear modeling of valuation ratio will be attempted. Section 4 will investigate the predictability issues in each regime by simulation methods. The final section briefly concludes.

2. Long-horizon predictability test using linear valuation model

To overcome difficulties related to test the statistical significance, Mark (1995) suggested using the nonparametric bootstrapping which is a way of getting the empirical distribution of test statistic by resampling the actual errors estimated by the model rather than pseudo-random errors from a particular distribution. With the nonparametric bootstrap, the hypothesis that the log exchange rate is unpredictable from the monetary fundamental can be rejected at the 5% level for the Deutsche mark and the Swiss franc. Also, the hypothesis that the prediction regression and the random walk generate the same out-of-sample predictability can be rejected for the same currencies. If we apply Mark's bootstrap procedure to the long-horizon prediction regression of stock returns, then given the null hypothesis of no predictability, Mark's DGP can be specified as follows:

$$r_t = a_0 + \varepsilon_{1,t}$$
$$z_t = b_0 + \sum_{i=1}^{I} b_i z_{t-i} + \varepsilon_{2,t} \qquad (1)$$

where r_t is the continuously compounded real stock return, that is,

$r^k_{t+k} = p_{t+k} - p_t$, with p_t representing the log of the real stock price of January of each year, and $z_t = p_t - d_t$, with d_t representing the log of real dividends over the previous calender year.

This error-correction representation is motivated by the assumption that the log of stock price cannot deviate far from the dividend process over long time horizons. Such predictability contrasts with the conventional view that stock price movements can be best described as a random walk with drift. As a result, the long-horizon regression tests for stock return can be thought of as the test of whether the inclusion of log of price-dividend ratio can beat the random walk forecast. Also, Mark's approach conditions on a cointegration between the p_t and d_t. Thus, it assumes *a priori* the existence of a long-run relationship between the p_t and d_t. Berkowitz and Giorgianni (2001) suggest a two-step procedure instead of assuming a cointegration *a priori*. For example, by calculating the Horvath-Watson statistics, we can test the no-cointegration null. If we reject the null hypothesis, then we can test the significance of test statistic using the appropriate DGP. If we conduct the linearity test and find that the process is stationary in one regime and nonstationary in the other we can apply the nonlinear nonparametric bootstrap. Here, we are not necessarily assuming that two processes, p_t and d_t are linearly cointegrated. Instead, we can conduct two separate nonparametric bootstrap procedures, one with the cointegration assumption and the other without the cointegration assumption. In practice, Berkowitz and Giorgianni (2001) show that the difference in the stationary status of $p_t - d_t$ results in very different empirical critical values. Here, we have the benefit of using the nonlinear nonparametric bootstrap method over the null which assumes the cointegration in the linear nonparametric bootstrap.

Table 1 reports the predictive regression estimation results for various horizons of the form,

$$r^k_{t+k} = a_k + b_k z_t + \varepsilon^k_{t+k} \qquad (2)$$

where r_t is the continuously compounded real stock return.

[Insert Table 1 here]

When p_t is above its fundamental value, it is expected to fall over time, which means that the coefficient of forecasting variable z_t should be negative and increase with the forecasting horizon k. We consider k=1,2,...,10 years as the forecasting horizons. The data used in the long-horizon predictability tests are from Campbell and Shiller (1998, 2002). This data covers the period from 1871 to 2001. We deflate the price and dividend series using the consumer price index in order to get the real valued series. We employ the valuation ratio of price-dividend ratio (January real stock price divided by real dividends over the previous calendar year) as a forecasting variable for the continuously compounded real stock returns. We estimate each equation in the DGP using OLS. In estimating the valuation ratio equation, we allow up to 5 lags and the minimum AIC is achieved when we set lags as 3. After correcting for bias using Shaman and Stine (1988), the equation for valuation ratio, and the equation for stock return under the null of no predictability are estimated as,

$$
\begin{aligned}
r_t &= 0.0237 + \widehat{\varepsilon_{1,t}} \\
z_t &= 0.1941 + 0.8627 z_{t-1} - 0.1687 z_{t-2} + 0.2883 z_{t-3} + \widehat{\varepsilon_{2,t}} \\
\widehat{s^2} &= 0.0312
\end{aligned} \qquad (3)
$$

This serves as the data generating process in calculating the bootstrap distribution. However, when $k>1$, the dependent variables are overlapping, and this induces $(k-1)^{th}$−order serial correlation into the disturbance term, ε^k_{t+k} under the null. When estimating the spectral density matrix of $(\varepsilon^k_{t+k}, \varepsilon^k_{t+k} z_t)$ at frequency zero in order to construct the heteroskedasticity autocorrelation consistent covariance matrix, we take the version of Andrews (1991) in order to determine the truncation lag for the Bartlett window along with the arbitrary truncation lag of 20. Also, as an approximation to the parametric model of Andrews' procedure, we employ a univariate AR(1) model. Let ρ_1 and ρ_2 be the estimated autocorrelation coefficients of ε^k_{t+k} and $\varepsilon^k_{t+k} z_t$ respectively. Also, let s^2_1 and s^2_2 be the estimated disturbance variance from the *AR(1)* processes. The truncation lag, *a*, for the Bartlett window in Newey and West's estimator given by Andrews' rule is $a=1.1447[a_1 T]^{1/3}$, where

$$
a_1 = \frac{\sum_{j=1}^{2} \frac{4\rho^2 s_j^4}{(1-\rho_j)^6 (1+\rho_j)^2}}{\sum_{j=1}^{2} \frac{s_j^4}{(1-\rho)^4}} \qquad (4)
$$

Though we take the serial correlation induced into the disturbance terms by the inconsistency between the sampling period and forecasting horizons into account, basing inference on the asymptotic distribution in this case provides a poor approximation to the exact distribution of test statistic. So, the statistical inference in this context requires caution to be exercised. The more autocorrelated the valuation ratio is as in our case of price−dividend ratio, the more underestimated autoregressive coefficient we have. In turn, this bias will be transmitted to the bias in the coefficient of forecasting variable. Also pointed out in various studies, including Kilian (1999), Kothari and Shanken (1997), Mark (1995), and Nelson and Kim (1993), the asymptotic tests have the wrong size when we have the small sample bias. In order to cure this small sample bias and size distortion, Mark (1995) draw an inference from the parametric and nonparametric bootstrap distributions generated under the null hypothesis of unpredictability. We follow two steps described above to build the bootstrap distributions of test statistic for the unpredictability of stock returns. We estimated the VAR under the null hypothesis that the stock return is unpredictable. With these estimates from the VAR, first, resample the fitted residuals with replacement and build the pseudo−sample. Discard the first 1,000 observations of the pseudo−sample to get the adjusted sample of original sample size for each forecasting horizon. It is worth noting that we adjust the AR coefficients of valuation ratio in the DGP using Shaman and Stine (1988) results. Second, we estimate the DGP and calculate the t−statistic under the null to get the bootstrapped distribution of test statistic under the null hypothesis. Calculate the p−value as the ratio of bootstrapped t−statistic less than or equal to the t−statistic from the long−horizon regression with the original sample. Table 1 reports the bootstrap p−values based on 5,000 replications. With the conventional significance levels, we find no evidence of predictability from any forecasting horizons. In addition to this finding, we can confirm that the power increase in the long horizon is inconsistent with the DGP of linear price−dividend ratio. This proves the fact that in a world of linear mean reversion there is no rationale for conducting the long−horizon regression tests. The problem is that under linearity the k−step ahead forecasts are obtained by the linear extrapolation from 1−step ahead forecasts. Thus, by construction there can be no gain in power at longer horizons (see Berben and van Dijk (1998); Kilian (1999); Berkowitz and Giorgianni (2001); Kilian and Taylor (2001)). However, the pattern present in the predictive regression results of increasing coefficient and increasing t−statistic is a common finding based on the valuation ratio as the p−values decrease from 0.591 with k=1 to 0.146 with k=10. The exact significance level for 10−year forecasting horizon is 0.37. Size distortion is still a matter of concern because the effective size is about 0.22−0.26 under the nominal size of 0.10, which means that we still have a tendency to reject the null hypothesis too often. Even though we expect the power to increase as we lengthen the forecasting horizon, it actually decrease slightly from 0.66 to 0.56. This proves the fact that in a world of linear mean reversion there is no rationale for conducting long−horizon regression tests. Various econometric problems such as the small sample bias, non−exogenous, predetermined forecasting variable, persistency of regressor and slow convergence to the asymptotic distribution can be solved by estimating the empirical distribution of test statistic generated by the nonparametric bootstrap method under the null hypothesis of stock return unpredictability.

3. Long-horizon predictability tests using nonlinear valuation model

Instead of assuming that there is a linear long-run stable relationship between the log of the real stock prices and the log of the real dividends, Lee (2006) investigates the nonlinear modeling of the regressor and the implications of the nonlinear modeling for the stock return predictability. Among the class of common regime switching models, he favors the STAR (Smooth Transition Autoregressive) model as the nonlinear alternative. In order to test the null of linearity against the LSTAR and ESTAR alternatives, Lee (2006) uses LM_1, LM_3, and LM_3^e test statistics of Luukkonen, Saikkonen, and Teräsvirta (1988) and the LM_2 test statistic of Granger and Teräsvirta (1993) and the LM_4 test statistic of Escribano and Jorda (1999). Test results show that the evidence of nonlinearity in the log of the real price-dividend process is quite strong. He also exercises the tests of selection between LSTAR and ESTAR, and a logistic function is selected as the transition function.

In this paper, we use the results of Lee (2006) in setting up the DGP for the nonlinear nonparametric bootstrap procedure to derive the empirical distribution of the test statistics.

3.1. Logistic smooth transition autoregressive specification

of the price-dividend ratio

To simplify the nonlinear nonparametric simulation, we estimate the parsimonious LSTAR model for the price-dividend ratio. In the context of the estimation of nonlinear autoregressive models by Granger and Teräsvirta (1993) and Teräsvirta (1994), Akaike information criteria(AIC) and Hannan-Quinn criteria(HQ) both select third-order serial correlation in the data, suggesting a nonlinear AR(3) model.

$$z_t = (\phi_{1,0} + \phi_{1,1} z_{t-1} + \phi_{1,2} z_{t-2} + \phi_{1,3} z_{t-3})(1 - G(z_{t-2};\gamma, c)) + (\phi_{2,0} + \phi_{2,1} z_{t-1} + \phi_{2,2} z_{t-2} + \phi_{2,3} z_{t-3}) G(z_{t-2};\gamma, c) + \varepsilon_t \quad (5)$$

The estimation of parameters in the STAR model in equation (5) can be done by applying the nonlinear least squares (NLLS), i.e., the parameters $\Theta = (\phi_1', \phi_2', \gamma, c)'$ can be estimated as

$$\hat{\Theta} = \arg\min_{\Theta} Q_T(\Theta) = \arg\min_{\Theta} \sum_{t=1}^{T} (z_t - F(x_t;\Theta))^2 \quad (6)$$

where $F(x_t;\Theta)$ is,

$$F(x_t;\Theta) = \phi_1' x_t (1 - G(s_t;\gamma, c)) + \phi_2' x_t G(s_t;\gamma, c) \quad (7)$$

Under the additional assumption that the errors ε_t are normally distributed, the NLLS is equivalent to the maximum likelihood. Otherwise, the NLLS estimates can be interpreted as the quasi maximum likelihood estimates. It has been proved that under certain regularity conditions, the NLLS estimates are consistent and asymptotically normal. The estimation can be performed using any conventional nonlinear optimization procedure. In order to alleviate the computational burden, we use the fact that the STAR model is linear in the autoregressive parameters ϕ_1 and ϕ_2 for fixed values of γ and c. Thus, conditional upon γ and c, the estimates of $\phi=(\phi_1',\phi_2')'$ can be obtained by the ordinary least squares as

$$\phi(\hat{\gamma}, c) = (\sum_{t=1}^{T} x_t(\gamma, c) x_t(\gamma, c)')^{-1} (\sum_{t=1}^{T} x_t(\gamma, c) y_t) \qquad (8)$$

where $x_t(\gamma,c)=(x_t'(1-G(s_t;\gamma,c)),x_t'G(s_t;\gamma,c))'$ and $\phi(\gamma,c)$ indicate that the estimate of ϕ is conditional upon γ and c. A convenient method to obtain sensible starting values for the nonlinear optimization algorithm is to perform a two-dimensional grid search over γ and c and select the parameter estimates which make the smallest estimate for the residual variance $\widehat{\sigma^2}(\gamma,c)$. The estimation results of the log price-dividend ratio as the nonlinear specification are shown in Table 3.

[Insert Table 3 here]

The residual standard deviation of LSTAR model is about 8.6% smaller than that of fitted linear model shown in Table 4.

[Insert Table 4 here]

From the AIC and HQ criteria, an increase in the number of parameters is compensated by this residual standard deviation decrease. For the linear model, AIC = −3.209 and HQ = −3.172 while for the LSTAR model, AIC = −3.292 and HQ = −3.200. Figure 1 shows the transition function from the estimated LSTAR model.

[Insert Figure 1 here]

As it shows, the change in the logistic transition function takes place for the values of z_{t-2} around 3.2389, and historical data on log price dividend ratio are evenly distributed throughout each regime.

In Figure 2, the value of transition function from the estimated LSTAR model is drawn during the sample time period.

[Insert Figure 2 here]

One thing to note from this graph is that the log price−dividend ratio switches regimes between the two states, hence generates a nonlinearity though the transition function, $G(s_t;\gamma,c)$ takes the value of 1 persistently since the end of 1950s.

3.2. New approach to nonparametric bootstrapping with nonlinear DGP

Kilian (1999), and Kilian and Taylor (2001) attempt a new empirical methodology for assessing the degree of long horizon predictability of stock return using the nonparametric bootstrap procedure with the nonlinear DGP. Their proposed procedure is represented in the DGP under the null hypothesis of no predictability which also captures the nonlinear dynamics in the valuation ratio. Kilian invalidates the linear VEC model framework underlying the existing long-horizon regression tests. His argument is based on the finding that the observed pattern of p−values across the forecast horizons in the empirical study is inconsistent with the size and power results for the linear VEC model. We also confirmed this in section Ⅱ. While he acquires almost constant or even slightly declining pattern of power results from a Monte Carlo simulation, the long horizon prediction regression expects the increasing power as we extend the forecasting horizon. He concludes that this fact is suggestive of a non−linear DGP, which makes the forecast accuracy tests based on the linear mean reversion toward the economic fundamentals unreliable. Following Kilian, in this section, we exploit the nonlinear characteristic of the persistent forecasting variable to perform a nonparametric and nonlinear bootstrapping. By way of this simulation we will derive the empirical distribution of the test statistic. We use the result from the previous subsection by postulating the DGP for $z_t = p_t - d_t$ that allows the smooth transition nonlinearity in these variables. Kilian and Taylor argue that the nonlinear mean−reversion better describes the asset price movements in a world of noise trading and risky arbitrage modeled along the lines of De Long, Shleifer, Summers and Waldmann (1990a, 1990b). With this economic intuition, Kilian and Taylor (2001) and Rapach and Wohar (2005) propose the parsimonious ESTAR model specification for the exchange rate deviations from the purchasing power parity and for the valuation ratio in the stock return forecasting model respectively. However, our nonlinear modeling process reveals that the LSTAR model specification is more appropriate for the specification of valuation ratio. If the underlying DGP is nonlinear, and can be described as an LSTAR model, a nonlinear DGP invalidates the p−values obtained for the predictive regression tests under the assumption of linear DGP. Therefore, the p−values reported in Table 1 are no longer valid if we assume a nonlinear DGP. The bootstrap algorithm for the long horizon regression test with the nonlinear valuation ratio specification can be described as follows.

(i) Estimate the following long horizon regression with the actual data set.

$$r^k_{t+k} = a_k + b_k z_t + \varepsilon^k_{t+k} \qquad (9)$$

where r_t is the continuously compounded real stock return. For each k of forecasting horizon, construct the test statistic.

(ii) We assume the following DGP under the null hypothesis of no predictability in order to capture the nonlinear dynamics in z_t,

$$r_t = a_0 + \varepsilon_{1,t}$$

$$z_t = (\phi_{1,0} + \phi_{1,1} z_{t-1} + \phi_{1,2} z_{t-2} + \phi_{1,3} z_{t-3})(1 - G(z_{t-2};\gamma, c))$$
$$+ (\phi_{2,0} + \phi_{2,1} z_{t-1} + \phi_{2,2} z_{t-2} + \phi_{2,3} z_{t-3}) G(z_{t-2};\gamma, c) + \varepsilon_{2,t}$$

(10)

where $G(z_{t-2};\gamma,c)=\frac{1}{1+\exp[-\gamma(z_{t-2}-c)]}$ is a logistic smooth transition function, and the disturbance vector, $(\varepsilon_{1,t},\varepsilon_{2,t})'$, is independently and identically distributed. This process can be estimated by NLLS.

(iii) Resample the residuals (in tandem) in order to generate a T+1,000 pseudo-sample of observations for r_t and z_t matching the original sample size. By drawing the disturbances in this way, We assume that they are *i.i.d.*. Discard the first 1,000 transient observations.

(iv) We compute and store the t-statistic corresponding to $\widehat{\beta_k}$ for k=1,...,10 for the pseudo-sample. We repeat this process 2,000 times in order to generate the empirical distribution of t-statistics for each k.

(v) Calculate the 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 in step (i).

This allows us to assess the significance of $\widehat{\beta_k}$ in the forecasting regression under the assumption of nonlinear DGP. Estimation results for the predictive regression model based on $z_t=p_t-d_t$ is reported in Table 5.

[Insert Table 5 here]

Note that the $\widehat{\beta_k}$ estimates and their associated t-statistics are computed in the same manner as in Table 1; the only difference in

Table 5 is that we use the modified bootstrap procedure, which assumes that z_t obeys a nonlinear process, to generate the p−values and assess the statistical significance. From Table 5, we see more insignificant evidences of real stock price predictability at every horizon based on the price−dividend ratio. One thing to note from Table 5 is that the p−values are declining as we extend the forecasting horizon when we assume a nonlinear DGP for the price−dividend ratio. As recognized by Kilian (1999), Berkowitz and Giorgianni (2001), and Kilian and Taylor (2001), a nonlinear framework does not necessarily preclude the predictability at long, but not short, horizons. While the predictability at long, but not short, horizons is a logical possibility in a nonlinear framework, we are also interested in whether there is any statistical evidence that the power increases at long horizons in a nonlinear framework. In order to investigate this for the data, we conduct the Monte Carlo simulations. The Monte Carlo simulation algorithm for the size test in the long horizon regression with the nonlinear valuation ratio specification can be described as follows.

(i) We assume the following DGP under the null hypothesis of no predictability in order to capture the nonlinear dynamics in z_t,

$$r_t = a_0 + \varepsilon_{1,t}$$

$$z_t = (\phi_{1,0} + \phi_{1,1} z_{t-1} + \phi_{1,2} z_{t-2} + \phi_{1,3} z_{t-3})(1 - G(z_{t-2};\gamma, c)) + (\phi_{2,0} + \phi_{2,1} z_{t-1} + \phi_{2,2} z_{t-2} + \phi_{2,3} z_{t-3})\, G(z_{t-2};\gamma, c) + \varepsilon_{2,t} \quad (11)$$

where $G(z_{t-2};\gamma,c)=\frac{1}{1+\exp[-\gamma(z_{t-2}-c)]}$ is a logistic smooth transition function, and the disturbance vector, $(\varepsilon_{1,t},\varepsilon_{2,t})'$, is independently and identically distributed. This process can be estimated by NLLS with the original data set.

(ii) Generate the pseudo-data under the null of no-predictability by resampling the estimated residuals with replacement.

(iii) Estimate the following long horizon regression with the generated pseudo-data set.

$$r^{k}_{t+k}=a_{k}+b_{k}z_{t}+\varepsilon^{k}_{t+k} \qquad (12)$$

where r_t is the continuously compounded real stock return. For each k of forecasting horizon, construct the test statistic.

(iv) For the generated data set in (ii), we assume the following DGP under the null hypothesis of no predictability in order to capture the nonlinear dynamics in z_t,

$$r_{t}=a_{0}+\varepsilon_{1,t}$$

$$z_{t}=(\phi_{1,0}+\phi_{1,1}z_{t-1}+\phi_{1,2}z_{t-2}+\phi_{1,3}z_{t-3})(1-G(z_{t-2};\gamma,c))$$
$$+(\phi_{2,0}+\phi_{2,1}z_{t-1}+\phi_{2,2}z_{t-2}+\phi_{2,3}z_{t-3})G(z_{t-2};\gamma,c)+\varepsilon_{2,t}$$

(13)

where $G(z_{t-2};\gamma,c)=\frac{1}{1+\exp[-\gamma(z_{t-2}-c)]}$ is a logistic smooth transition function, and the disturbance vector, $(\varepsilon_{1,t}, \varepsilon_{2,t})'$, is independently and identically distributed. This process can be estimated by NLLS.

(v) Resample the residuals in tandem in order to generate a pseudo-sample of observations for r_t and z_t matching the original sample size. Discard the first 1,000 transient observations.

(vi) We compute and store the t-statistic corresponding to $\widehat{\beta_k}$ for k=1,...,10 for the pseudo-sample by estimating equation (12).

(vii) Repeat steps (v)-(vi) 1,000 times in order to generate the empirical distribution of t-statistics for each k. Calculate a p-value for each k as the proportion of the simulated t-statistics that are less than the t-statistic calculated in step (vi).

(viii) Repeat steps (ii)-(vii) 1,000 times. Calculate the size for each k as the proportion of the simulated p-values that are less than 0.10, the nominal size. This allows us to assess the effective size for each k in the long horizon prediction regression under the assumption of a nonlinear DGP. NLLS estimation was carried out using the Broyden, Fletcher, Goldfarb and Shanno (BFGS) optimization algorithm in the OPTMUM subroutine library of GAUSS 3.2. We use the Moore-Penrose pseudo-inverse if the matrix is singular or near-singular when we invert the $X'X$ matrix. From Table 5, we see that the modified bootstrap procedure is remarkably accurate, as the effective size of each test statistic is very close to the nominal size of 0.10 and remains fairly constant across forecast horizons for the actual sample size of 130. The size distortion is not a concern in testing the predictability using the nonparametric bootstrap with nonlinear DGP. From the empirical distribution of the t-statistic, the degree of predictability becomes higher as we lengthen the forecasting horizon.

However, we cannot find any decisive evidence in stock return predictability. Our test results indicate that the stock return cannot be well predicted by the log of price dividend process over the random walk even if we correctly specify the data generating process.

4. The predictability in each regime

In this section, we split the sample into two subsamples, one with regime 0 when the transition function $G(z_{t-2};\gamma,c)$ takes the value of zero, and the other with regime 1 when the transition function $G(z_{t-2};\gamma,c)$ takes the value of one.

According to our estimation results in Table 3, the log of price−dividend ratio process can be represented as $z_t=(0.9088+0.6589z_{t-1}-0.0551z_{t-2}+0.1040z_{t-3})+\varepsilon_{1,t}$ in regime 0. In regime 1, on the other hand, this process is given as

$$z_t=-0.9618+1.1336z_{t-1}-0.5957z_{t-2}+0.7461z_{t-3}+\varepsilon_{2,t}.$$

Armed with estimates in each regime, we perform the nonparametric bootstrap in order to investigate the predictability in each regime. Each nonparametric bootstrap procedure performs 5,000 replications. Table 6 reports the value of $\widehat{\beta_k}$ along with its t_A and p−value.

[Insert Table 6 here]

t_A is a t−statistic calculated using Andrews' procedure when we correct for heteroskedasticity and autocorrelation. Parameter estimates of the forecasting variable increase in absolute value from −0.138 when the forecasting horizon is 1 year to −0.715 when we extend the forecasting horizon to 10 years. p−values show the statistical significance of the parameter estimate under 10% significance level when the p−value for the forecasting horizon of 1 year is 0.11. Table 7 shows stock return predictability using the log of price−dividend ratio as a forecasting variable by the method of nonparametric bootstrap.

[Insert Table 7 here]

As we extend the forecasting horizon from 1 year to 10 years, the parameter estimates increase in absolute value from −0.072 to −1.4580, and t−statistics also increase in absolute value from −0.845 to −3.989. Even though the bootstrap p−values decrease from 0.24 for k=1 to 0.11 for k=10, the exact significance level shows that no parameter estimates are statistically significant under 10% significance level. This result can be compared to the previous result when we attempt the nonparametric bootstrap within regime 0 where all the parameter estimates with the exception of 1−year horizon shows statistically significant predictability. When we include the log of the price−dividend ratio as a regressor in the long−horizon prediction regression for stock returns, we implicitly assume that the regressor is stationary over long time horizon. However, empirical evidence is weakly supportive of the hypothesis of cointegration between the log of the stock price and the log of the dividend. The persistency in the regressor leads to several difficulties when making inference from long−horizon regressions. The least squares estimate of the slope parameter is biased away from zero in small samples. With the sample of size 47 when the log of the price−dividend ratio process is in regime 1, if there is no statistical relationship between the log of the stock price and the log of the dividend, the long−horizon regressions become close to a classical spurious regression as we lengthen the forecasting horizon.

5. Concluding remarks

Various factors contribute to the biased estimate. Serial correlation induced into the disturbance terms by the inconsistency between the sampling period and forecasting horizons can be taken care of if we use the heteroskedasticity and autocorrelation consistent Newey−West estimate of covariance combined with Andrews' selection procedure for truncation lag.

Basing inference on the asymptotic distribution, however, still provides a poor approximation to the exact distribution of the test statistic. It has been noted that the parameter estimate from the long horizon predictive regressions of stock returns is biased since the work of Stambaugh (1986). Moreover, we have also found that the log of the price−dividend ratio used as a forecasting variable in predictive regressions of stock returns is subject to a unit root test. It is difficult to reject a unit root in the log of price−dividend ratio. Even though the unit root is able to be rejected, the persistence of regressor has been of concern. The more autocorrelated is the regressor, the more underestimated autoregressive coefficient we have. In turn, this bias will be transmitted to the bias in the coefficient of forecasting variable.

The local–to–unity framework has been used to derive the asymptotic null distribution of the t–statistic, where the explanatory variable is assumed to follow an autoregression with a root near to unity.

Other methodology for studying these biases is the simulation method through which we can get empirical distribution of test statistic. In section 2, we explore the long–horizon predictability tests using linear valuation model. There, we find that the nonparametric bootstrap procedure still leaves size distortion behind and that the evidence for the stock return predictability weakens compared to the extant literature if we extend the sample period to year 2001. In section 3, we incorporate the LSTAR specification in the DGP and exercise the long–horizon predictability tests. The LSTAR model specification shows that the log of the price–dividend process has two regimes, one with stationary process and the other with nonstationary process. Our test results indicate that the stock return cannot be well predicted by the log of the price–dividend process over the random walk and the evidence survives even after we correctly specify the nonlinear dynamics inherent in the forecasting variable. In section 4, we split the sample into two accordingly, and implement the long–horizon predictability tests in each regime. Results show that the stock return is predictable in the stationary regime, while the test statistic under the null of unpredictability is insignificant in the nonstationary regime.

Table 1. Estimation results for the predictive regression model, under the assumption of a linear DGP

				$z_t = p_t - d_t$				
Horizon(k)	$\widehat{\beta_\kappa}$	t_{20}	t_A	lag	p-value	R^2	size	power
1 year	−0.008	−0.177	−0.182	1	0.830	0.000	0.252	0.656
2 years	−0.050	−0.440	−0.455	7	0.770	0.004	0.258	0.650
3 years	−0.047	−0.284	−0.284	11	0.813	0.003	0.258	0.630
4 years	−0.138	−0.631	−0.625	14	0.740	0.015	0.254	0.626
5 years	−0.235	−0.982	−0.970	17	0.659	0.033	0.228	0.610
6 years	−0.277	−1.102	−1.111	21	0.621	0.039	0.228	0.586
7 years	−0.337	−1.340	−1.427	25	0.561	0.050	0.232	0.582
8 years	−0.469	−1.943	−2.183	28	0.396	0.082	0.222	0.582
9 years	−0.527	−2.147	−2.426	28	0.344	0.092	0.226	0.564
10 years	−0.560	−2.172	−2.453	31	0.374	0.094	0.228	0.558

Notes: t_{20} and t_A are the heteroskedasticity autocorrelation consistent t-statistics for the slope coefficient in the long horizon regression with a fixed truncation lag of 20 and a truncation lag by Andrews(1991) procedure respectively. The truncation lag by Andrews procedure is reported in the column "lag". The nonparametric bootstrap distribution of the test statistic is based on 5,000 iterations.

Table 2. Lagrange multiplier test for serial correlation in residuals

Order	LM stat.	p-value
1	0.834	0.361
2	3.220	0.200
3	4.637	0.200
4	4.921	0.296
5	4.898	0.428

Notes: We estimate the model $z_t = b_0 + \sum_{i=1}^{I} b_i z_{t-i} + \varepsilon_{2,t}$ by allowing up to 5 lags and choose the one with $p=3$ which produces the minimum AIC. Lagrange Multiplier test for serial correlation in the estimated residual shows that we fail to reject the null of no serial correlation for up to fifth-order serial correlation.

Table 3. Log price−dividend ratio nonlinear specification

	LSTAR Model		
Variable	Estimate	St. Error	HAC−St. Error
constant	0.9088	0.3387	0.3363
z_{t-1}	0.6589	0.0958	0.1054
z_{t-2}	−0.0551	0.1248	0.1497
z_{t-3}	0.1040	0.0954	0.1131
constant	−0.9618	0.4569	0.3649
z_{t-1}	1.1336	0.1938	0.1720
z_{t-2}	−0.5957	0.3010	0.2478
z_{t-3}	0.7461	0.1925	0.1405
γ	500.0000	1127.93	1896.28
c	3.2389	0.3252	0.5469
SSR		3.9588	
S.E. of Residual		0.1780	
AIC		−3.292	
HQ		−3.200	
BIC		−3.066	

Notes: We estimate the logistic smooth transition autoregressive model in equation (5) by quasi−maximum likelihood method. Akaike information criteria and Hannan−Quinn criteria both select the third−order serial correlation in the data, so we set AR(3) in each regime. In the last panel, SSR (sum of squared residuals) decrease from 4.7376 for linear specification to 3.9588 for LSTAR specification.

Table 4. Log price−dividend ratio linear specification

	Linear Model		
Variable	Estiamte	St. Error	HAC−St.Error
constant	0.2050	0.1832	0.1831
z_{t-1}	0.8429	0.0866	0.1035
z_{t-2}	−0.1734	0.1136	0.1336
z_{t-3}	0.2711	0.0915	0.1049
SSR		4.7376	
S.E. of Residual.		0.1947	
AIC		−3.209	
HQ		−3.172	
BIC		−3.118	

Notes: We estimate the linear autoregressive model by quasi−maximum likelihood method. Akaike information criteria and Hannan−Quinn criteria both select the third−order serial correlation in the data, so we set the model as AR(3).

Table 5. Estimation results for the predictive regression model, under the assumption of a nonlinear DGP

		□ $= p_t - d_t$		
Horizon(k)	$\widehat{\beta_k}$	t−statistic	p−value	size
1 year	−0.008	−0.182	0.421	0.100
2 years	−0.050	−0.455	0.348	0.102
3 years	−0.047	−0.284	0.419	0.100
4 years	−0.138	−0.625	0.355	0.122
5 years	−0.235	−0.970	0.313	0.120
6 years	−0.277	−1.111	0.316	0.122
7 years	−0.337	−1.427	0.295	0.122
8 years	−0.469	−2.183	0.226	0.126
9 years	−0.527	−2.426	0.215	0.120
10 years	−0.560	−2.453	0.219	0.130

Notes: We use the modified bootstrap procedure, which assumes that z_t obeys a nonlinear process as in Table 3, to generate p−values and assess statistical significance. The exact significance levels for the null hypothesis of unpredictability are 0.42 for 1 year horizon and decrease to 0.22 for 10 year horizon as we extend the forecasting horizon. The effective size is very close to the nominal size for every forecasting horizon.

Table 6. Estimation results for the predictive regression model, under the assumption of a linear DGP: Regime 0

		$z = p_t - d_t$	
Horizon(k)	$\widehat{\beta_k}$	t_A	p−value
1 year	−0.138	−1.501	0.105
2 years	−0.243	−2.315	0.053
3 years	−0.344	−2.319	0.075
4 years	−0.537	−3.569	0.022
5 years	−0.610	−3.875	0.022
6 years	−0.592	−3.918	0.026
7 years	−0.654	−4.344	0.021
8 years	−0.655	−3.950	0.036
9 years	−0.656	−4.106	0.036
10 years	−0.715	−4.095	0.048

Notes: t_A is the heteroskedasticity autocorrelation consistent t−statistics for the coefficient in the long horizon regression with a truncation lag by Andrews(1991) procedure. The nonparametric bootstrap distribution of the test statistic is based on

5,000 iterations. The exact significance levels for various forecasting horizons range from 0.02 to 0.11.

Table 7. Estimation results for the predictive regression model, under the assumption of a linear DGP: Regime 1

	$z_t = p_t - d_t$		
Horizon(k)	$\widehat{\beta_k}$	t_A	p−value
1 year	−0.072	−0.845	0.236
2 years	−0.105	−0.559	0.371
3 years	−0.082	−0.276	0.440
4 years	−0.138	−0.348	0.430
5 years	−0.232	−0.571	0.386
6 years	−0.621	−1.522	0.255
7 years	−0.679	−1.926	0.228
8 years	−0.814	−2.420	0.184
9 years	−1.137	−3.262	0.137
10 years	−1.458	−3.989	0.111

Notes: t_A is the heteroskedasticity autocorrelation consistent t−statistics for the coefficient in the long horizon regression with a truncation lag by Andrews(1991) procedure. The nonparametric bootstrap distribution of the test statistic is based on 5,000 iterations. According to the empirical distribution of test statistic by nonparametric bootstrap, with the conventional significance levels, we find no evidence of predictability beyond for any forecasting horizons. The exact significance levels for various forecasting horizons range from 0.11 to 0.44.

Figure 1. Historical log price−dividend ratio and the value of transition function $G(s_t;\gamma,c)$

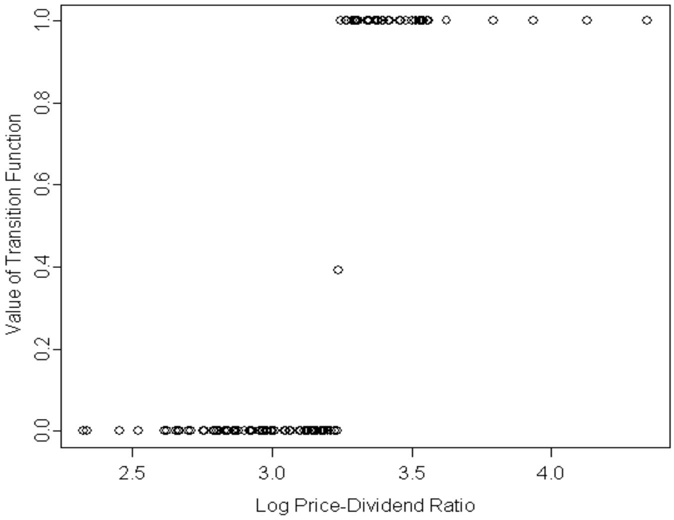
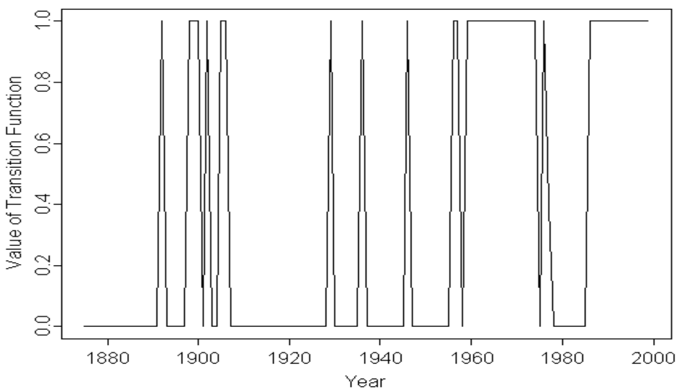


Figure 2. Time plots of estimated transition function $G(s_t;\gamma,c)$ from 1871−2001



References

Andrews, D. W. K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59, 807−858.

Berben, R. P. and D. van Dijk, 1998, Does the absence of cointegration explain the typical findings in long horizon regressions?, Report 9814, Econometrics Institute, Erasmus University.

Berkowitz, J. and L. Giorgianni, 2001, Long−horizon exchange rate predictability?, Review of Economics and Statistics 83, 81−91.

Campbell, J. Y. and R. J. Shiller, 1998, Valuation ratios and the long−run stock market outlook, Journal of Portfolio Management 24, 11−26.

Campbell, J. Y. and R. Shiller, 2002, Valuation ratios and the long−run stock market outlook: An update, Cowles Foundation Discussion Paper No. 1295.

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

Elliot, G. and J. Stock, 1994, Inference in time series regression when the order of integration of a regressor is unknown, Econometric Theory 10, 672−700.

Escribano, A. and O. Jorda, 1999, Improved testing and specification of smooth transition regression models, in Nonlinear Time Series Analysis of Economic and Financial Data, 289−319, eds., P. Rothman, Kluwer Academic Publishers.

Granger, C. W. J. and T. Teräsvirta, 1993, Modelling Nonlinear Economic Relationships, Oxford University Press.

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. and M. P. Taylor, 2001, Why is it so difficult to beat the random walk forecast of exchange rates?, Manuscript, University of Michigan.

Kothari, S. P. and J. Shanken, 1997, Book−to−market, dividend yield, and expected market returns: A time series analysis, Journal of Financial Economics 44, 169−203.

Lanne, M., 2002, Testing the predictability of stock returns, The Review of Economics and Statistics 84, 407−415.

Lee, H−J., 2006, Nonlinearity in valuation ratio and its implications on long−horizon stock return predictability, Journal of Economic Research 11, 167−201.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann, 1990, Noise trader risk in financial markets, Journal of Political Economy 98, 703−738.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, Journal of Finance 45, 379−395.

Luukkonen, R., P. Saikkonen and T. Teräsvirta, 1988, Testing linearity against smooth transtion autoregressive models, Biometrika 75, 491−499

Mark, N. C., 1995, Exchange rates and fundamentals: evidence on long−horizon predictability, American Economic Review 85, 201−218.

Nelson, C. R. and M. J. Kim, 1993, Predictable stock returns: The role of small sample bias, Journal of Finance 48, 641−661.

Rapach, D. and M. Wohar, 2005, Valuation ratios and long−horizon stock price predictability, Journal of Applied Econometrics 20, 327−344.

Shaman, P. and R. Stine, 1988, The bias of autoregressive coefficient estimators, Journal of the American Statistical Association 83, 842–848.

Stambaugh, R. F., 1986, Bias in regressions with lagged stochastic regressions, Working Paper, University of Chicago.

Stambaugh, R. F., 1999, Predictive regressions, Journal of Financial Economics 54, 375–421.

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

Torous, W., R. Valkanov, and S. Yan, 2004, On predicting stock returns with nearly integrated explanatory variables, Journal of Business 77, 937–966.

Wooldridge, J. M., 1991, On the application of robust, regression–based diagnostics to models of conditional means and conditional variances, Journal of Econometrics 47, 5–46.