Assessing the Power of Long-Horizon Predictive Tests in Models of Bull and Bear Markets

Alex Maynard*
University of Guelph

Dongmeng Ren
University of Guelph

October 2, 2014

Abstract

We compare the finite sample power of short and long-horizon tests in nonlinear predictive regression models of regime switching between bull and bear markets, allowing for time varying transition probabilities. As a point of reference, we also provide a similar comparison in a linear predictive regression model without regime switching. Overall, our results do not support the contention of higher power in longer horizon tests in either the linear or nonlinear regime switching models. Nonetheless, it is possible that other plausible nonlinear models provide stronger justification for long-horizon tests.

Key words: predictive regression, long-horizon regression, regime-switching, nonlinear, stock return predictability

*Corresponding author. Department of Economics, College of Business and Economics, University of Guelph, McKinnan Building, Guelph, ON, Canada, N1G 2W1. *E-mail:* maynarda@uoguelph.ca, *Fax:* 519-763-8497, *Phone:* 519-824-4120, extension 53014. This paper was prepared for the 2013 Advances in Econometrics Conference: Essays in Honor of Peter C. B. Phillips held at the Southern Methodist University. We thank the conference participants, participants at the 2014 Meeting of the Canadian Economic Association, Erik Hjalmarsson, and Leo Michelis for useful discussion. We are grateful to the Editors, Joon Park and Yoosoon Chang, and to an anonymous referee for very helpful comments and suggestions that have substantially improved the paper. We gratefully acknowledge the use of code provided by Erik Hjalmarsson and code made available online by Zhuanxin Ding and Marcelo Perlin. Maynard thanks the SSHRC for research funding.

1 Introduction

Long-horizon predictive regression tests have been the subject of substantial interest in both empirical finance and financial econometrics, following influential applications to stock return predictability, [Fama and French (1988*b*), Campbell and Shiller (1988)], exchange rate prediction [Mark (1995), Chinn and Meese (1995)] and the term structure of interest rates [Fama and Bliss (1987), Cutler *et al.* (1991)]. The strong empirical results, including large R-squares and t-statistics, observed in long-horizon stock return regressions have been particularly influential in the finance literature, as one of several important findings leading to the conclusion by the Royal Swedish Academy of Sciences (2013) Nobel Press release that

> There is no way to predict the price of stocks and bonds over the next few days or weeks. But it is quite possible to foresee the broad course of these prices over longer periods, such as the next three to five years.

Similarly, in his survey article, Cochrane (1999, "New Facts in Finance") cites the predictability of stock returns at long horizons as one of the three most important new facts in finance.

This paper asks whether there is any long-horizon power advantage in empirically plausible nonlinear models using valid econometric tests. In particular, we compare the power of short and long-horizon predictive regressions under the assumption that the data is generated by nonlinear Hamilton (1989) style regime switching models. These models have been successfully used for some time to model long term swings in foreign currency exchange rates [Engel and Hamilton (1990), Bollen *et al.* (2000), Dewachter (2001), Cheung and Erlandsson (2005)]. A more recent literature develops similar models to capture bull and bear runs (or multiple states) in stock markets [Maheu and McCurdy (2000), Chauvet and Potter (2000), Guidolin and Timmermann (2005), Guidolin and Hyde (2010), Maheu *et al.* (2012)]. Schaller and Norden (1997) and Guidolin and Timmermann (2005) estimate predictive regressions with regime switches. To our knowledge, no one has considered the power of long-horizon regressions in these models, although Kilian (1999) mentions this avenue.

Overall, our results are not supportive of long-horizon regression. Long-horizon tests can show modest size-adjusted power gains using a standard t-test in the linear model. Using more powerful tests, we find no power advantages at long horizons. Short-horizon tests have as good or better power. Long-horizons tests fare no better (and often worse) in bull/bear regime shift models, even when the regressor is predictive of the regime shift itself. In the context of long-horizon IVX tests, we find that the power comparison depends on both the horizon and the tuning parameter determining the degree of filtering used to create the instrument. Nonetheless, we find that with a good choice of this tuning parameter, the short run IVX can provide as good or better power as its long run counterpart in both linear and regime switching predictive regression settings.

The remainder of the paper is organized as follows: Section 2 reviews the literature and provides the background and motivation for our study. In Section 3 we introduce the regime shifting predictive regression model that is used as a data generating process for our power comparison. Section 4 describes the data. Section 5 presents the empirical estimates of the regime shift model used to anchor our simulations. Section 6 describes the short and long-horizon specifications and tests employed in the power comparisons. Section 7 provides an extensive simulated power comparison of short and long-horizon regressions using two different long-horizon specifications and several long-horizon tests.

2 Literature, Background, and Motivation

The dependent variables in long-horizon regressions are typically multi-year asset returns observed at a monthly or quarterly frequency, defined by

$$r^k_{t+k} = \sum_{j=1}^{k} r_{t+j}, \tag{1}$$

where k is the horizon period and r_t is the one period return. The long-horizon return is regressed on a pre-determined predictor x_t, such as a dividend or earnings price ratio, resulting in the long-horizon regression specification

$$r^k_{t+k} = \beta_0(k) + \beta_1(k)x_t + \varepsilon^k_{1,t+k} \tag{2}$$

with tests for long-horizon predictability based on the restriction

$$H_0 : \beta_1(k) = 0. \tag{3}$$

The special case in which $k = 1$ corresponds to the short-horizon predictive regression

$$r_{t+k} = \beta_0 + \beta_1 x_t + \varepsilon_{1,t+1}, \tag{4}$$

which has also generated substantial interest. Although predetermined, valuation predictors, such as the dividend and earnings price ratios, tend to be both persistent and highly endogenous. Following Cavanagh *et al.* (1995), they have been traditionally modelled using the local-to-unity process [Phillips (1987), Chan and Wei (1987)] and this has more recently been generalized to include mildly integrated processes [Phillips and Magdalinos (2007), Phillips and Lee (2012)]. The long horizon, persistence, and endogeneity combine to create a nonstandard inference problem. They have thus generated an increasing interest in financial econometrics, including the development of new long-horizon tests [Valkanov (2003), Liu and Maynard (2007), Hjalmarsson (2012), Phillips and Lee (2013)].

Although his interest in this particular problem appears fairly recent, it seems fair to say that Professor Peter C.B. Phillips has made at least three very important contributions to this literature. The first contribution is an indirect contribution. In fact, the long-horizon regression provides one of many nice examples of the useful insights that can be gained from the nonstationary asymptotics developed in very large part by Professor Phillips. The empirical results from these regressions appear remarkably strong and are of clear importance to economics and finance. At the same time, early simulation studies detect substantial size distortion [Kim and Nelson (1993), Goetzmann and Jorion (1993)]. Allowing the horizon to increase with k and modelling the predictor as a near unit root, Valkanov (2003) uses the tools of (near) nonstationary asymptotics to provide a clear explanation for the size distortion observed in earlier studies. In his framework, the regressor x_t on the right hand side of (2) is near integrated, while the dependent variable (1) is partially summed since k grows with the sample size, T. Consequently, under the null hypothesis in (3), the asymptotics resemble in certain respects the spurious regression asymptotics of Phillips (1986), helping to explain both the size distortion and large t and R-squared statistics in empirical applications.

Further insights are provided by Hjalmarsson (2012), who shows that when the regression coefficient from the long-horizon regression is scaled by k it shares the same second order endogeneity bias as the short-horizon regression estimator of β_1 in (4). Often referred to as the Stambaugh (1999) bias, following Cavanagh *et al.* (1995) this bias is perhaps best understood by viewing the large sample behavior of the predictive regression in (4) as a special case of cointegrating regression asymptotics [Park and Phillips (1988), Park and Phillips (1989), Phillips and Hansen (1990), Phillips (1991), inter alia], generalized to near unit roots [Phillips (1987), Phillips (1988), Chan and Wei (1987)]. Thus, the underlying theoretical frameworks developed earlier by Professor Phillips have been employed to provide a great deal of insight into this important empirical application.

More recently, Professor Phillips has made two important direct contributions to this literature. One of the challenges in predictive regression is that, if the predictor is modelled as a near unit root process, the asymptotic distribution of the test statistic typically depends on the local-to-unity parameter which cannot be estimated. Until recently, many of the econometric methods that have been designed to provide inference in predictive regression, including long-horizon regressions, rely on a Bonferroni bound employing a first stage confidence interval for the local-to-unity parameter based on the inversion of a unit root test as in Stock (1991). Phillips (2014) has recently uncovered serious concerns with this approach, showing that the Stock (1991) confidence interval for the local-to-unity parameter has zero asymptotic coverage when the data generating process is stationary. This can lead to the invalidity of the resulting Bonferroni based tests when the true process is mildly integrated or stationary.

In a new major breakthrough, Phillips and Magdalinos (2007) generalize the local-to-unity model to allow for mildly integrated processes and Magdalinos and Phillips (2009) use this new theory to propose an alternative solution to the predictive regression known as the IVX method, which uses a mildly filtered version of the predictor as an instrument. Their solution altogether eliminates the endogeneity bias and dependence on an unknown local-to-unity parameter without reliance on Bonferroni methods. It continues to work for stationary and mildly integrated regressors and is easily extended to multivariate settings.[1] Phillips and Lee (2013) have recently proposed a long-horizon version of the IVX method, using a rearranged version of the long-horizon specification suggested in Jegadeesh (1991).

While the econometric literature discussed above is primarily concerned with understanding or eliminating the size distortion in long-horizon predictive regression, in this paper we are interested in the power of long-horizon specifications relative to their short run counterparts in (4). This has been a controversial question, with the presumed power advantages of long-horizon regressions both strongly criticized in Boudoukh *et al.* (2008)["The Myth of Long-Horizon Predictability"], and valiantly defended by Cochrane (2008).

There seems to be an unstated assumption in much of the early empirical literature that predictive regressions are more powerful at longer horizons, perhaps due to economic intuition or simply as a result of the stronger empirical results reported at longer horizons, see e.g. Campbell *et al.* (1997, Chapter 7). This proposition is stated more directly in Campbell (2001). Several works provide support for this conjecture [Campbell (2001), Kilian and Taylor (2003), Wohar and Rapach (2005), Mark and Sul (2006), Cochrane (2008)], while others question the power advantage of longer horizons [Ang and Bekaert (2007), Boudoukh *et al.* (2008), Hjalmarsson (2008), Hjalmarsson (2012)].

With a few exceptions, the power properties of long-horizon regressions have mainly been examined assuming that the true model for returns is given by a short-horizon linear model similar to (4) [Campbell (2001), Valkanov (2003), Mark and Sul (2006), Cochrane (2008), Hjalmarsson (2008), Hjalmarsson (2012)]. In this case, as both Campbell (2001) and Hjalmarsson (2012) remark, the short-horizon regression in (4) is a correctly specified model. This may lead one to question the advantage of (2), relative to the correctly specified regression in (4). Nonetheless, as Campbell (2001), Mark and Sul (2006), and Cochrane (2008) all astutely note, the short-horizon regression loses power in the realistic case in which the valuation predictors are both highly persistent and contemporaneously correlated with the residual. They argue that the long-horizon regression offers power improvements in this case.

In fact, this combination of near unit root regressor with contemporaneous endogeneity, gives rise to second order biases in (4), resulting in both size distortion and power loss [Phillips and Hansen (1990), Phillips (1991), Campbell and Yogo (2006)]. However, this arguably reflects only an inefficient use of the information in (4) by OLS based tests that are sub-optimal in the combined presence of endogeneity and (near) nonstationarity. Since more efficient short run

[1]It also solves the much more general problem of cointegration with near integrated regressors, see Elliott (1998).

estimators and more powerful short run tests are now available [Jansson and Moreira (2006), Campbell and Yogo (2006), Magdalinos and Phillips (2009), inter alia] one can still argue that there is no inherent power advantage to the long-horizon specification when the true model is the linear short-horizon predictive regression. In our simulated power comparisons we do find some small size-adjusted power improvements using OLS based tests when the true model is linear, but these improvements disappear when more powerful tests are used. Hjalmarsson (2012) also provides some results to support this contention.

In our view, any inherent advantage to the long-horizon specification could only arise under nonlinear alternatives to (4). In this case, neither (4) nor (2) is correctly specified, so that the long run specification in (2) may plausibly provide a better approximation to the true, but unknown, nonlinear alternative. Kilian (1999) similarly argues that "the observed pattern of p-values is inconsistent with a linear model". Yet, we are aware of relatively little literature which investigates the power of long-horizon specifications when the data generating process is nonlinear. Kilian and Taylor (2003) and Wohar and Rapach (2005) simulate power using a nonlinear ESTAR model for the predictor finding power advantages at longer horizons. On the other hand, Ang and Bekaert (2007) simulate power in a nonlinear present value model, but find long-horizon regressions to be less powerful. Marmer (2008) shows that predictability is present only in the short run in the case when the predictive component of returns is a general integral transformation of a nonstationary predictor.[2]

We compare the power of long and short-horizon predictive tests in the case when the true model is characterized by regime switching between bull and bear markets. We assume that the econometrician does not know the true model and tests predictability using standard linear predictive models without regime shifting in both the short and long-horizon specifications. Since neither linear regression is correctly specified, it is no longer obvious that a powerful short-horizon test should outperform a long-horizon test. In fact, since the states are themselves persistent, with a probability of switching between states that depends on the predictor, it seems a priori plausible that the long-horizon regressions might outperform the short-horizon regressions without the Markov-switching component.

3 Regime Switching Predictive Regressive Models

We consider the following regime switching model

$$r_{t+1} = \beta_0(s_t) + \beta_1(s_t)x_t + \sigma_1(s_t)\varepsilon_{t+1}, \quad (5)$$

[2] Kasparis *et al.* (2012) and Cai and Gao (2013) find evidence of nonlinear predictive power in the dividend yield for stock returns, while the results in Juhl (2011) are suggestive, but insignificant. Gonzalo and Pitarakis (2012) find evidence of structural breaks in the predictive regression relationship. However, none of these papers considers the implications for long-horizon predictive tests.

where the state $s_t \in 1 \ldots N$ is a first order Markov process with transition probabilities

$$p_{ij} \quad = \quad \text{Prob}\left(s_t = j | s_{t-1} = i\right). \tag{6}$$

When $N = 1$ and there is just one single state, then the regime switching model specializes to the standard linear predictive model

$$r_{t+1} \quad = \quad \beta_0 + \beta_1 x_t + \sigma_1 \varepsilon_{t+1}. \tag{7}$$

In order to allow the predictor to predict transitions between states, as well as returns within states, we focus our attention on a model with time varying transition probabilities. Following Ding (2012), we model them as (where Φ is the CDF of $N(0,1)$):

$$\begin{aligned} p_{ij,t} &= \Pi_{l=1}^{j-1}\left(1 - q_{il,t}\right) q_{ij,t} \\ q_{ij,t} &= \begin{cases} \Phi\left(\theta_{0,ij} + \theta_{1,ij} x_{t-1}\right) & \text{for} \quad 1 < j < N-1 \\ 1 & \text{for} \quad j = N. \end{cases} \end{aligned} \tag{8}$$

This formulation allows us to model the transition probabilities as a function of the predictor, while still ensuring that these probabilities appropriately sum to one. For example, in the two-state model $N = 2$, this simplifies to

$$p_{i1,t} \quad = \quad q_{i1,t} = \Phi\left(\theta_{0,i1} + \theta_{1,i1} x_{t-1}\right) \tag{9}$$

$$p_{i2,t} \quad = \quad 1 - p_{i1,t}, \tag{10}$$

in which x_t may help predict transitions between bull and bear states.

To provide an economic interpretation, we follow the literature in referring to the states as bull and bear states in the two state model and as bull, normal, and bear states in the three state model. To avoid any ambiguity, we define the bull and bear states, respectively, as the states with highest and lowest values of the estimated intercept $\hat{\beta}_0(s_t)$ in (5).

3.1 Persistent, Endogenous Regressors

Although predetermined, common valuation predictors, such as the dividend or earnings price ratios, are both persistent and endogenous. It is typically difficult to reject unit roots in these variables. Yet, taken literally, a unit root in these predictors may not seem natural from a financial modelling perspective. Moreover, time series unit root tests are, by design, not consistent against near unit root alternatives. Accordingly, the local-to-unity or near unit root models [Phillips (1987), Phillips (1988), Chan and Wei (1987)] have been increasingly used to model the predictor. This class of models has been recently generalized by Phillips and Magdalinos (2007) to include mildly integrated series. Consider, for simplicity, a first order autoregressive process for the predictor

$$x_t = \rho_0 + \rho_1 x_{t-1} + v_t. \tag{11}$$

A very general formulation for ρ_1 is provided by Phillips and Lee (2013), who model ρ_1 as

$$\rho_1 = 1 + c/T^{\eta}. \tag{12}$$

Equation (12) specializes to a unit root process, in the case when $c = 0$, a near integrated process when $\eta = 1$ and $c < 0$ and a mildly integrated series when $0 < \eta < 1$.[3]

In order to realistically capture the endogeneity in the data, we follow the linear predictive regression literature in modelling the innovations as contemporaneously correlated with correlation $\delta = \text{corr}\,(\varepsilon_t, v_t)$:

$$(\varepsilon_t, v_t/\sigma_2)' \sim \text{i.i.d.} \left(0, \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix}\right). \tag{13}$$

This error structure is common in both the predictive regression and regime shift literatures. In the absence of regime shift, it nonetheless has the disadvantage of imposing homoskedasticity. In our context, we allow for the conditional return variance $\sigma_1^2(s_t)$ to vary across states. Incorporating time varying conditional heteroskedasticity within states could further enhance the realism of the model, although possibly at the cost of complicating its estimation.

4 Data and Variables

Our study is based on 1048 monthly observations of the excess return, including dividends, and the dividend price ratio on the S&P 500 price index starting in July 1926 and ending in October, 2013. The stock return and dividend data are maintained by Robert Shiller and are combined with the monthly risk free rate maintained by Kenneth French. The monthly dividends are linearly interpolated from yearly totals. Defining the level of the S&P 500 price index by P_t, dividends by D_t, and the risk free rate by i_t, we define the excess return as

$$r_t = ln\left(\frac{P_t + D_t}{P_{t-1}}\right) - i_t. \tag{14}$$

Following common practice [see for example, Fama and French (1988*a*) and Chapter 7 of Campbell *et al.* (1997)], we sum the monthly dividends over the past twelve months to obtain

$$dp_t = ln\left(\frac{D_t + D_{t-1} + \ldots D_{t-11}}{P_t}\right). \tag{15}$$

Dividends are often summed as in (15) on account of their strong seasonality. Although not standard practice, this could also be achieved by a yearly average of dividends. It is possible that smoothing the dividends may also mitigate any bias from the linear interpolation of yearly dividends.[4]

[3] Locally ($c > 0$, $\eta = 1$) and mildly ($c > 0$, $0 < \eta < 1$) explosive processes are also allowed for in (12).

[4] Recent research by Ghysels and Miller (2014) uncovers potentially serious bias in the use of interpolated data in the context of cointegration testing.

Following (Schaller and Norden 1997, footnote 14, page 14), who estimate a regime switching predictive model with fixed coefficients, we then standardize dp_t to have zero-mean and unit variance in order to define our predictor as

$$x_t = \left(dp_t - \bar{dp}_t\right)/\hat{\sigma}_{dp_t}, \tag{16}$$

where $\bar{dp}_t$ and $\hat{\sigma}_{dp_t}$ are the sample mean and standard deviation of dp_t respectively. Although this normalization is not typical in the linear predictive regression literature, we also find that it leads to a somewhat better fit and more easily interpretable parameters when estimating a regime switching predictive regression model.

If dp_t is (near) nonstationary, it will have an infinite range, tending to take values far from zero. Without the normalization in (16), this would lead the transition probabilities, defined as an asymptotically homogeneous function of the predictor in (8),[5] to cluster at zero or one depending on the signs of θ_1 and x_t. By normalizing the standard error, we avoid this problem. On the other hand, this normalization would present additional challenges for asymptotic inference in the regime shift model, since x_t is a function of the dividend price ratio process in (15), which may itself be characterized by a near unit root, in which case the standard deviation in (16) is increasing in sample size. We leave these developments for future work. In this paper, we simply employ the parameter estimates of the regime switching model to parameterize the data generating process used in our simulated power analysis.

5 Empirical Estimates

Before turning to the regime switching regression model, we first estimate a regime switching model for the mean and variance of the excess return in (14). In this way, we explore the distinct properties of the returns in the bear, bull, and normal states of the market. This also allows us to assess the significance of the coefficients and the fit of the regime switching model, without the additional complications induced by time varying transition probabilities and the inclusion of a persistent predictive regressor.

Tables 1 and 2 respectively show the estimated coefficients and transition coefficients in the pure regime switching model without predictors. Table 1 provides the estimates of $\beta_0(s_t)$ and $\sigma_1(s_t)$ for each state s_t, when restricting $\beta_1(s_t) = 0$ in (5). Likewise, Table 2 provides estimates of the intercepts for the transition probabilities in (8), which are non-time varying in the absence of a predictor.[6] Standard errors are given in parenthesis.

For the purposes of comparison, we include the results for the single state model in column

[5]See Park and Phillips (1999).

[6]Omitting the predictor restricts $\theta_{1,ij} = 0$ in (8) for all i, j.

2 of Table 1.[7] This shows a mean monthly (log) excess return for the S&P500 of 0.0051, or approximately six percent on an annual basis, over the entire sample period and a return variance of 0.0014. Columns 3 and 4 show the estimated bear and bull state values for the two state model. The mean return in the bull state is more than substantially higher than in the bear state, whereas the variance of the bear state is approximately nine times as high as that of the bull state, reflecting the greater turbulence that occurs during market downturns. In both cases, the difference between the estimates in the two states is several times bigger than the larger of the two standard errors, lending statistical support to the observation that the market behaves differently during bear and bull states. In the three state model (columns 5-7), the bull and bear states capture the more substantial rallies and downturns, with calmer periods absorbed by the normal state. Consequently, the three state models shows a much larger differential between average returns in the bull and bear markets. The bear market again shows the greatest volatility, while the bull state has the lowest return variance.

Columns 2-3 of Table 2 show the the estimated intercepts in the model for the transition probabilities between bear and bull states in the two state model. Both states are quite persistent, but the probability of switching from bear to bull state in any period is substantially higher than the probability of switching from bull to bear. This reflects the greater duration of bull states. Columns 4-6 of the same table provide the coefficients of the transition probabilities for the three state model. All three states are again persistent, but the normal and bull states are more persistent than the bear state. In all cases, the standard errors suggest that the estimates are reasonably precise.

The last row of Table 1 shows the Hannan and Quinn (1979) information criteria (H-Q IC) for each model, with lower, more negative values denoting a better penalized fit.[8] Both regime switching models are strongly favored relative to the single state model and the two state model shows a slight improvement over the three state model.[9]

We now include the dividend price ratio as a state-dependent predictor for the returns themselves in (5).[10] This arguably allows for a more realistic specification of the predictive regression model, which incorporates the difference in expected return and variance across the bear and bull states. It also allows for the predictive power of the dividend price ratio to vary across states.

[7]We are grateful to Marcelo Perlin for use of his MATLAB code for the estimation of the fixed transition probability Markov Switching Models. His code is available online and documented in Perlin (2012).

[8]Guidolin and Hyde (2010) similarly employ the H-Q IC to aid the selection of the number of states in a regime switching model.

[9]In additional results, omitted for brevity, a marginally better H-Q IC value was obtained in the four state, but at the cost of large standard errors, resulting in the poor identification of a number of the model coefficients.

[10]We are grateful to Zhuangxin Ding for use of his MATLAB code. As documented in Ding (2012), this code extends Perlin (2012)'s original code in order to allow for time varying transition probabilities. In order to verify that global maximums were obtained, we compared estimation results and likelihood values across several perturbations to our initial starting values for the optimizations.

In addition, we allow the dividend price ratio to help predict the transition probability between bear and bull states via its inclusion in the model for the time-varying transition probabilities in (8). To the extent that valuation variables may predict transitions between bull and bear markets, this may have stronger practical implications than return predictions within a given state.

Drawing inference in the regime shift predictive regression model is presumably complicated by a similar second order bias as in the linear predictive regression model. Deriving techniques that avoid second order bias and size distortion in this highly nonlinear model is a worthy goal for future research, but lies beyond both the scope and needs of our current paper, in which we require only consistent model estimates in order to anchor our simulated power comparisons in Section 7. In a related context, Gonzalo and Pitarakis (2012) develop an IVX based predictive regression test with structural breaks. Although their model differs in a number of important respects from that in (5) and (8), their finding offers general support for the existence of breaks or regime changes in predictive regression.

Table 3 shows the parameters estimates and standard errors in parenthesis in the regime switching predictive regression model (5) with time varying transition probabilities given by (8). The results for the two state model are shown in columns 3-4. As in Table 1, the bull state has a higher mean return, whereas the bear state has a higher standard deviation. The predictive power of the dividend yield (if significant) is also about seven times as large in the bear state as in the bull state. Gonzalo and Pitarakis (2012) recently come to a similar conclusion in a model of predictive regression with structural breaks. Interestingly, in the three state model shown in columns 5-7, it is the bear state in which the dividend yield is most strongly predictive, with the coefficient on the dividend price ratio positive and estimated with a large standard error in the bear state.

Table 4 shows the estimates and standard errors for the parameters in (8) governing the transition probabilities. Here, we focus our comments on the two state model in columns 2-3, whose parameters are the most readily interpretable, as seen from (9). The first panel provides the estimates of the intercepts (the $\theta_{o,ij}$ in 8). The diagonal element of the intercepts ($\hat{\theta}_{o,11}$) is large and positive, while the off-diagonal element ($\hat{\theta}_{o,21}$) is negative. This indicates that the states are again persistent, with a relatively small probability of switching in any single period. The second panel in Table 4 shows the slope coefficients (the $\theta_{1,ij}$ in 8). The positive estimate of $\hat{\theta}_{1,11}$ in the second column suggests that the conditional probability of remaining in a bull market is highest when the dividend price ratio is large. Put another way, a switch from bull to bear market, possibly via a market crash, appears more likely when the stocks are highly valued relative to dividends (low dividend price ratio). The negative estimate of $\hat{\theta}_{1,21}$ suggests that the transition from bear to bull market is also more likely when stocks are relatively higher valued. This may reflect the observation that bull markets sometimes end in sudden crashes, whereas the transition from bear to bull often follows a more gradual market rally, which allows the stock prices to partially recover prior to the onset of the next bull market.

The last row of Table 3 shows the Hannan and Quinn (1979) information criteria (H-Q IC) for each model. The inclusion of the dividend price ratio leads to a lower H-Q IC in all models, although it is unclear if the difference is significant. Among the models with the dividend price ratio included in Table 3, the two state model has the lowest H-Q IC. Its value is substantially lower than the single state model or linear predictive regression. The two state model also has a somewhat better fit than the three state model. We therefore employ the two state model with the dividend price ratio as our baseline model in which to anchor our simulations in Section 7. Figure 1 shows the fitted probabilities of a bear state in this model, with the calendar date shown on the horizontal axis and the major market downturns listed in the figure caption. Although during a majority of the time the probability of the bear state is low, the model appears to identify nearly all of the major historical downturns and/or market crashes as bear states. Indeed, it would appear to do quite a good job in capturing the major periods that most market observers would classify as bear states.

6 Long-Horizon Regression Specifications and Tests

In the simulated power exercise that follows, the empirical researcher is not assumed to know the true model. Instead it is assumed that either a short or long-horizon linear predictive regression is employed as an empirical specification. In particular, we consider two long-horizon specifications from the literature, both of which share the same null hypothesis, but which differ in the formulation of their alternative. Both involve the choice of a horizon length (k), which in some asymptotic analysis is modelled as a function of the sample size T [Valkanov (2003), Phillips and Lee (2013)] and in others taken to be arbitrarily large, but fixed, as in Hjalmarsson (2011). A fairly general formulation covering both of these possibilities is given by

$$k = \lambda T^{v} \quad \text{for} \quad 0 \leq v \leq 1, \quad \lambda > 0, \text{ and } \quad \lambda \mathbf{I}(v \neq 0) \leq 1 \tag{17}$$

where $\mathbf{I}()$ denotes the indicator function. In (17), $v = 0$ corresponds to a fixed horizon as in Hjalmarsson (2011), ($v = 1$, $0 < \lambda < 1$) corresponds to a horizon that grows at a fixed fraction of sample size as in Valkanov (2003), and ($\lambda = 1$, $0 < v < 1$) allows a horizon that grows more slowly than T, as in Phillips and Lee (2013).

6.1 Standard Long-Horizon Specification

One of the most common versions of the long-horizon specification is given by (2), in which r^{k}_{t+k} is the long-horizon (k period) return in (1). The null hypothesis of non-predictability is tested via a test of the parameter restriction in (3) under the maintained hypothesis[11] that

$$E_t \varepsilon_{1,t+k}^k = 0. \tag{18}$$

6.2 Rearranged Long-Horizon Regression

We also consider a rearrangement of this regression originally proposed by Cochrane (1991) and Jegadeesh (1991) and given by

$$r_{t+k} = \beta_0^*(k) + \beta_1^*(k) x_t^k + \varepsilon_{t+k}^{*,k} \tag{19}$$

in which a one period return is regressed on the long-horizon (or partially summed) regressor

$$x_t^k = \sum_{j=1}^{k} x_{t+k-j}. \tag{20}$$

The null hypothesis of non-predictability is then tested using the coefficient restriction

$$H_0 : \beta_1^*(k) = 0 \tag{21}$$

under the maintained assumption that

$$E_t \varepsilon_{t+1}^{*,k} = 0. \tag{22}$$

Liu and Maynard (2007) use this same rearranged regression to propose an exact nonparametric test of long-horizon predictability based on a sign test. Phillips and Lee (2013) employ this specification to provide a long-horizon version of the IVX solution to the predictive regression problem.

Since x_t^k is realized at time $t+k-1$, the rearranged regression in (19) can also be viewed as a short-horizon regression that uses a long history of the predictor x_t as its regressor. Note, however, that the temporal distance between the one period component returns and one period component predictors in the two regressions, (2) and (19), are the same. That is, $(r_{t+k}, x_t^k) = (r_{t+k}, x_{t+k-1} + x_{t+k-2} + \ldots + x_t)$ and $(r_{t+k}^k, x_t) = (r_{t+1} + r_{t+2} + \ldots r_{t+k}, x_t)$ both involve spacings of $1, 2, \ldots$ and k periods between r_t and x_t. Therefore, when (r_t, x_t) are jointly covariance stationary the rearrangement in (19) can be motivated by expressing $\beta_1(k) = \mathrm{cov}(r_{t+k}^k, x_t)\,/\mathrm{var}(x_t)$ and $\beta_1^*(k) = \mathrm{cov}(r_{t+k}, x_t^k)\,/\mathrm{var}(x_t^k)$ and noting that the covariances in the numerator are equivalent. In other words, the null hypotheses that $\beta_1(k) = 0$ and that $\beta_1^*(k) = 0$ impose the same orthogonality restriction, viz. $\mathrm{cov}(r_{t+k}^k, x_t) = \mathrm{cov}(r_{t+k}, x_t^k) = 0$. More generally, even when x_t is not stationary, under the maintained hypothesis in (22), Liu

[11] The maintained hypothesis in (18) ensures the r_{t+k} is unpredictable under the null in (3). We refer to it as a maintained hypothesis, since it is not usually explicitly tested in this literature.

and Maynard (2007) and Phillips and Lee (2013) show that the null hypothesis (21) of the rearranged specification implies the original null hypothesis in (3), together with the maintained assumption (18).[12]Therefore, the two formulations test the same null hypothesis, but may have different power implications under the alternative hypothesis.

6.3 Long-Horizon Tests for the Standard Specification

We are not currently aware of a valid existing method of testing (3) using the traditional long-horizon formulation in (2) under the general specifications (12) and (17) for the persistence of predictor, x_t and the choice of the horizon k. Nevertheless, this specification is of particular interest since it is perhaps the mostly widely used in the empirical literature. Within the confines of the local-to-unity model, which imposes $\eta = 1$ in (12), Valkanov (2003) and Hjalmarsson (2011) provide long-horizon predictive tests based on (2) employing local-to-unity asymptotics and Bonferroni bounds procedures.

Valkanov (2003) provides the asymptotic distribution of the standard t-test for testing the restriction (3) in (2) under a local-to-unity asymptotic specification for x_t (12, with $\eta = 1$) and a horizon that grows at a fixed fraction of the sample size ($v = 1$ in 17). For a given value of the local to unity parameter c, this provides for a correct critical value for the standard test statistic, scaled by the square-root of the sample size. However, since the critical value depends on c, which cannot be consistently estimated in a time series context, a Bonferroni procedure based on a first stage confidence interval is employed. This approach is therefore similar in spirit to the short-horizon predictive test of Cavanagh *et al.* (1995). Both employ a standard test statistic (possibly after rescaling) with corrected critical values.

Hjalmarsson (2011) instead corrects the test statistic by providing a second order endogeneity correction to the estimator of $\beta_1(k)$ on which the test is based. This is accomplished via an augmented regression of the type first introduced by Phillips (1991)[13]

$$\begin{aligned} r_{t+k}^k &= \beta_0(k) + \beta_1(k)x_t + \phi(k)\Delta_c x_t^k + \varepsilon_{1,t+k}^{+,k} \quad \text{where} \\ \Delta_c x_t^k &= \sum_{j=1}^k \Big(x_{t-k+j} - (1+\frac{c}{T})x_{t-k+j-1}\Big) = \sum_{j=1}^k v_{t-k+j}. \end{aligned} \tag{23}$$

The inclusion of the term $\Delta_c x_t^k$ in (23)[14] removes the endogeneity bias in the estimate of $\beta_1(k)$ in (23), denoted by $\widehat{\beta}_1^+(k)$. Although the distribution of $\widehat{\beta}_1^+(k)$ still depends on the horizon length k, it does so in a simple way. As is well explained in Boudoukh *et al.* (2008), under the

[12]As explained in Liu and Maynard (2007), (21,22) imply $E_{t+j-1}r_{t+j} = \beta_0^*(k)$ and therefore $E_t r_{t+k} = k\beta_0(k)$, which is equivalent to (3,18). See also Phillips and Lee (2013) for an insightful explanation of the relationship between the two long-horizon specifications.

[13]In (23) the estimator $\beta_1^+(k)$ is equivalent to a local-to-unity version of the fully modified Phillips and Hansen (1990) estimator specialized to the context of the long-horizon predictive regression.

[14]Due to the inclusion of this term, we include a plus in the superscript of $\varepsilon_{1,t+k}^{+,k}$ to distinguish it from $\varepsilon_{1,t+k}^k$.

null hypothesis in (3), estimates of $\beta_1(k)$ are approximately proportional to k. Consequently, Hjalmarsson (2011) shows that when rescaled by the horizon length, $\widehat{\beta}_1^+(k)/k$ has a mixed normal null limiting distribution whose variance no longer depends on k. Likewise, letting $t_k^+(c)$ denote the standard t-statistic associated with the test of $H_0 : \beta_1(k) = 0$ in (23), he shows that

$$t_k^+(c)/\sqrt{k} \rightarrow_d N(0,1) \tag{24}$$

under the null hypothesis, justifying the use of standard critical values.

In the simulations below, we will refer to $t_k^+(c)/\sqrt{k}$ as the 'oracle' version of Hjalmarsson (2011)'s test. It is infeasible in practice, because the augmentation $\Delta_c x_t^k$ depends on the unknown value of c. In practice, a Bonferroni bound based on a first-stage confidence interval for c is therefore required in order to draw inference. Denoting this confidence interval, with confidence level $1-\alpha_1$ by $\left(\underline{c}_{\alpha_1}, \bar{c}_{\alpha_1}\right)$ a feasible, yet conservative version of the test rejects if

$$t_{k,min}^+/\sqrt{k} = \inf_{c \in \left(\underline{c}_{\alpha_1}, \bar{c}_{\alpha_1}\right)} t_k^+(c)/\sqrt{k} > z_{\alpha_2}, \tag{25}$$

where z_{α_2} is the standard normal critical value associated with the significance level α_2, such that $\alpha_1 + \alpha_2 = \alpha$, where α is the desired significance level of the test. In other words, a rejection occurs for the Bonferroni bounds test only if it occurs for every possible value of c in the first stage confidence interval. The requirement that $\alpha_1 + \alpha_2 = \alpha$ can lead to overly conservative tests and in practice adjustments to α_1 are typically recommended in order to keep the test from becoming too conservative [Cavanagh *et al.* (1995), Campbell and Yogo (2006)]. In implementing this test, we select α_1 according to the recommendations in Hjalmarsson (2011), to which we refer the reader for further details. In the linear predictive regression context, Hjalmarsson (2012) finds that his test has better power properties than the earlier test by Valkanov (2003). Thus, we employ this version of the test of the restriction (3) in (2).

In an important recent paper, Phillips (2014) has shown that the Stock (1991) confidence interval for the local-to-unity parameter c has zero asymptotic coverage when the data generating process is stationary. This can lead to either invalidity or excessive conservatism of Bonferroni procedures based on a first stage confidence interval for c, when the true process is mildly integrated ($\eta < 1$ in (12)) or stationary. Although Hjalmarsson (2011) employs an alternative first stage bound, based on Chen and Deo (2009), his simulations do indicate that his test becomes increasingly conservative for large negative values of c, particularly at long horizons.

We address this in several respects. First, for all of our simulations, we compare the power of the test across the oracle version (24), in addition to the Bonferroni version in (25). Although the oracle version of the test is not feasible, this allows us to assess whether any of the power comparisons across horizons is driven primarily by the implementation of the Bonferroni bound. More importantly, we include power comparisons for the new IVX based long-horizon test of Phillips and Lee (2013). This is the first regression based long-horizon test that we are aware of which does not require Bonferroni bounds. Since it is based not on (2), but rather on the rearranged regression (19), we discuss its implementation in the next section. Currently, we are not aware of any IVX test based on the original long-horizon specification in (2). Given the emphasis in the empirical literature on the specification in (2), we feel that it is important to include feasible tests based on this specification in our power comparisons.

6.4 Long-Horizon IVX Test for the Rearranged Specification

Phillips and Lee (2013) develop an IVX estimator and predictive test for use in the long-horizon context, employing the rearranged regression in (19). They refer to their test as the long-horizon IVX (LHIVX) test. As mentioned above, unlike other long-horizon regression based tests, the LHIVX test does not rely on Bonferroni bounds or first stage confidence intervals for c. Instead it provides a single test statistic with a standard normal limit distribution yielding simple inference. Consequently, its validity is not restricted to the local-to-unity model, and it is easily applied in multivariate contexts. Indeed, Phillips and Lee (2013) allow for a very general model of the persistence in x_t as in (12).

IVX based tests employ an instrument that is a mildly filtered version of the original regressor x_t. The instrument is filtered enough so that the endogeneity term is not present in the limit distribution of the instrumental variable estimator, but not so much that it unduly reduces the instrument strength. An important theoretical foundation for its development is Phillips and Magdalinos (2007)'s generalization of the local-to-unity framework to allow for mildly integrated variables. The instrument is designed to be a mildly integrated regressor ($\eta < 1$ in (12)). This intermediate persistence of the instrument is crucial to the success of the IVX approach, since an I(1) or local-to-unity based instrument would still be subject to an endogeneity bias that depends on c, whereas an I(0) instrument would only be weakly correlated with an I(1) regressor.

We will focus on the case of a single regressor, since this is the case we consider in Section 7. In the short-horizon case, when $k = 1$ in (19), their IVX instrument is defined as the mildly filtered series

$$\tilde{z}_t = \sum_{j=1}^{T} \left(1 + c_z/T^{\delta_z}\right)^{(t-j)} \Delta x_j \tag{26}$$

where $c_z < 0$ and $0 < \delta_z < 1$ are specified by the researcher. In the context of the long-horizon regression when $k > 1$ in (19), Phillips and Lee (2013) propose the long-horizon instrument

$$\tilde{z}_t^k = \sum_{j=1}^{k} \tilde{z}_{t+j-1}, \tag{27}$$

which is partially summed in the same way as the regressor in (20). Under an additional rate condition that guides the choice of c_z and δ_z, Phillips and Lee (2013) show that their estimator is asymptotically mixed normal, resulting in standard normal or Chi-squared inference over the wide range of persistence levels for x_t in (12), including mildly integrated, nearly integrated, integrated, mildly explosive and locally explosive models.

7 Power of Long-Horizon Regression in Linear and Regime Switching Predictive Regression Models

We provide a simulation based comparison of the power of short and long-horizon linear regression predictive tests when the true model is given by the regime switching model in (5) with persistent endogenous regressors as in (11,12,13). As a point of comparison, we also provide results when the data is generated from the short-horizon linear predictive regression (7).

The regime switching model has a good number of parameters associated with it. In order to ensure that we simulate from a realistic parameterization of the model, we use the estimates in Tables 3 to set the default values of the parameters in (5). Similarly, we use the estimates in Table 4 to simulate the model for the time varying transition probabilities in (8). We focus on the two regime ($N = 2$) model, which appears to have the best fit, but also compare it to the linear predictive regression ($N = 1$).[15] In the model for x_t, the parameters in (12) are not estimable and there is a genuine uncertainty regarding the persistence of valuation predictors such as the dividend price ratio. We consider two specifications of (12): ($c = -2.5, \eta = 1$) and ($c = -20, \eta = 1$). The first represents a case in which the series is almost as persistent as a unit root, whereas the second represents the case of a much less persistent series. The innovations for return and predictor series are drawn from (13) using a multivariate normal distribution with correlation $\delta = -0.95$, which is typical of valuation predictors, such as the dividend price ratio.[16] We consider $T = 200$ as a fairly small sample and $T = 1048$, the sample size in our data set described in Section 4. All simulation results are based on two thousand replications.

We restrict certain parameters under the null hypothesis. In the linear predictive regression it is typical to test the null hypothesis that $\beta_1 = 0$ in (7) under the maintained assumption that $E_t \varepsilon_{t+1} = 0$, which is equivalent to the non-predictability condition $E_t r_{t+1} = \beta_0$.[17] We enforce the same non-predictability condition when imposing the null hypothesis in the regime switching model. This implies both zero slope coefficients in all states and intercepts that do not vary across states, i.e.

$$H_0 : \begin{pmatrix} \beta_0(s_t) \\ \beta_1(s_t) \end{pmatrix} = \begin{pmatrix} \beta_{0,0} \\ 0 \end{pmatrix} \qquad \text{for all states } s_t. \tag{28}$$

[15] We have also re-run many of our results using a three regime version of our model, based on earlier estimates, obtaining qualitatively similar results.

[16] We estimate $\delta = -0.9424$ in the linear predictive regression using our data.

[17] Since, following much of the literature, we do not impose $\beta_0 = 0$, this could be more precisely referred to as the null hypothesis of non-time-varying-predictability. The non-zero intercept allows for a non-time varying risk premium and accommodates the positive average returns that have been recorded over long historical periods.

For $\beta_{0,0}$, the null value of β_0, we take a weighted average of our empirical estimates of $\beta_0(s_t)$ across the N states, where the weights are given by the unconditional state probabilities implied by the transition probabilities in (8) evaluated at the mean value of the predictor $\bar{x}$.

In order to provide a power curve, rather than simply power at a single point in the alternative, we allow the parameters that are restricted under the null hypothesis to drift back towards their default (estimated) values under the alternative hypothesis, i.e.

$$H_A : \begin{pmatrix} \beta_0(s_t) \\ \beta_1(s_t) \end{pmatrix} = \begin{pmatrix} \beta_{0,0} \\ 0 \end{pmatrix} + \frac{\gamma}{T^{(1+v_2)/2}} \begin{pmatrix} \beta_0^*(s_t) - \beta_{0,0} \\ \beta_1^*(s_t) \end{pmatrix}, \tag{29}$$

where $\beta_0^*(s_t)$ and $\beta_1^*(s_t)$ are the default values of the parameters taken from the empirical estimates in each state shown in Table 4. The value of γ determines the distance from the null hypothesis and is allowed to vary across the horizontal axis. The value of v_2 is specified below for each test.

Although the data is generated from the regime switching predictive regression model, we do not assume that the empirical researcher knows that this is the true model. Rather, our interest lies in comparing the power of the most common short and long-horizon linear predictive regressions when the data is generated by a plausible nonlinear model, such as the regime switching predictive regression model. We compare the power of the long and short-horizon regressions using the two well known long-horizon specifications, the standard long-horizon regression (2) and the rearranged long-horizon regression (19), discussed in Section 6. We consider only the empirically relevant positive, one-sided alternatives $H_A : \beta_1(k) > 0$ or $H_A : \beta_1^*(k) > 0$ respectively to the null restrictions (3) and (21). To keep our comparisons easily viewable, in each figure we compare the power curve for the short-horizon specification $k = 1$ to two long-horizon specifications, where k is chosen according to (17). The choice of λ and v are discussed separately below for each test.

In the case of the standard long-horizon regression (2), we simulate power comparisons across three different tests:

- Although the OLS based test is invalid in our framework and highly sized distorted, it has nonetheless been influential in the literature. It is therefore of some interest to know whether longer horizon versions of the test lead to any power improvements or whether they merely aggravate size distortion. Thus we compare size-adjusted power of the OLS based test between long and short-horizon tests.
- As an alternative test that has conservative size within the confines of the local-to-model, we consider the feasible, Bonferroni version of the rescaled, endogeneity corrected tests of Hjalmarsson (2011) using the test statistic $t^+_{k,min}/\sqrt{k}$ in (25). Since this test is not

oversized, we compare power rather than size-adjusted power in this case.[18]

- Since Phillips (2014) has pointed out some important shortcomings with many of the existing Bonferroni based approaches to predictive testing and Hjalmarsson (2011) reports that when c is large negative, the Bonferroni version of his test becomes more conservative at larger horizons, we also provide results for the oracle version of the same test in (24). Although the oracle version is of course infeasible, it helps us to ensure that none of the comparison are solely driven by the influence of the Bonferroni bounds in the feasible version of the test.

In the case of the rearranged regression, we compare two additional tests.

- We compare size-adjusted power in an OLS based test of the rearranged regression.
- More interestingly, we also compare the power (without size-adjustment) of the IVX and LHIVX tests of Phillips and Lee (2013) discussed in section 6.

Since this results in a great many power curves, we include only a few of the curves to illustrate our main findings below. Additional results, that generally confirm the main conclusions arrived at below, are available from the authors upon request. These include results using both the smaller sample size of $T = 200$ and the lower persistence parameter of $c = -20$.

7.1 Power Comparisons Based on the Standard Long-Horizon Specification

We first compare the size-adjusted power of short and long-horizon tests based on (2) for the OLS based test. We also compare power for both the Bonferroni and oracle versions of Hjalmarsson (2011)'s test discussed in Section 6.3. Since these tests have non-trivial power against $O_p\left(T^{-1}\right)$ alternatives, we set $v_2 = 1$ in (29). Similarly, we set $v = 1$ in (17) when selecting k and take $\lambda = 0.05$ and 0.010, resulting in values of $k = 10$ and $k = 20$ when $T = 200$ and $k = 52$ and $k = 104$ when $T = 1048$. In all cases we compare to $k = 1$.

7.1.1 True Model is a Linear Predictive Regression

Figure 2 (using $T = 1048$) provides the size adjusted power curves of standard t test when $k = 1$ and $k = 52$ and $k = 105$ when the true model is the linear predictive model in (7). The three curves are generally quite similar. A small power advantage to the longer horizon tests is detectable towards the middle of the power curve ($50 < \gamma < 150$). The OLS based test does not make efficient use of the endogeneity between the regression error and regressor innovation,

[18] In principle, it is not clear whether any of the tests we consider here are correctly sized when the residual variance $\sigma_1^2(s_t)$ varies across states under the null hypothesis. However, this does not appear to be an important practical concern in any of our simulations.

which may explain why there is room for some slight power advantages to the long-horizon specification. We next turn to the oracle version of Hjalmarsson (2011)'s test in Figure 3, which makes optimal use of this endogeneity within the confines of the linear predictive model with local-to-unity predictor. With the use of this more powerful test, the slight advantage of the long-horizon test disappears. The short-horizon is more powerful throughout the power curve. The dominance of the short-horizon test remains even after the test is made feasible using its Bonferroni implementation in Figure 4. This confirms that when the true model is linear and the short run regression is therefore correctly specified, the long-horizon tests should not be able to provide power improvements to efficient versions of the short-horizon test.

7.1.2 True Model is a Regime Switching Predictive Regression

We now consider the same comparisons in the regime switching predictive regression model, in which the regressor can be predictive for both the transition probabilities of switching between regimes and the return outcomes within regimes. In this case, both the short and long-horizon models are inherently misspecified and act only as linear approximations to the true nonlinear model. In principle, there is no reason to argue that the long-horizon model couldn't provide an improved approximation and better power than the short-horizon model. Indeed, it would seem surprising if there were not some nonlinear models for which this is the case. The practical question that we now address is whether, using simulations based on empirical estimates of the regime switching model, we obtain power improvements at longer horizons. Put another way, if the regime switching model were the true model, would we be more likely to detect the resulting return predictability at longer horizons?

Our results suggest that this is not the case. Long-horizons tests fare no better (and sometimes worse) when the data is generated from nonlinear regime shift models, capturing bear/bull markets. We find this result even when allowing the regressor to be predictive of the regime shift itself. The OLS test, shown in Figure 5 (using $T = 1048$), again shows modest size-adjusted power advantages at long-horizons close to the null hypothesis, whereas the short-horizon test has the best power for distant alternatives. However, the short-horizon test continues to show the best power over the entire alternative in both the oracle and Bonferroni versions of Hjalmarsson (2011)'s test, shown in Figures 6 and 7. Thus, even in this nonlinear model, the long-horizon tests do not provide a power advantage over their short-horizon counterparts.

7.2 Power Comparisons Based on Rearranged Long-Horizon Specification

We now compare power across horizons in the rearranged long-horizon empirical specification in (19), with our primary focus on the IVX and LHIVX estimators discussed in Section 6. For the short-horizon IVX, $k = 1$. For the LHIVX, Phillips and Lee (2013) consider return horizons that grow more slowly than sample size ($v < 1$ in (17)). We fix $\lambda = 1$ and consider both $v = 0.80$ and $v = 0.90$, resulting in values of $k = 69$ and $k = 117$ when $T = 200$ and $k = 260$ and $k = 522$ when $T = 1048$. Phillips and Lee (2013) show that their test is divergent against alternatives of the rate $T^{-1/2}k^{-1/2} = T^{-(1+v)/2}$. Therefore, we set $v = 0.80$ in the specification of the local alternative in (29), where γ again varies across the horizontal axis. Following both the recommendations of Phillips and Lee (2012) and Phillips and Lee (2013) we set $c_z = -5.0$. The selection of δ_z in (26) requires slightly more attention. For the short-horizon IVX, Phillips and Lee (2012) recommend a value of δ_z between 0.75 and 0.95. For the long-horizon regression, rate restrictions require $\delta_z < v$ in (17) and Phillips and Lee (2013) suggest $\delta_z = v - 0.05$. Therefore, we employ $\delta_z = 0.75$ when $v = 0.80$ and $\delta_z = 0.85$ when $v = 0.90$. We compare these to two versions of the short-horizon IVX: $k = 1$ with $\delta_z = 0.75$ and $k = 1$ with $\delta_z = 0.85$. With these choices, we use recommended values of δ_z in all cases and we can compare the short and long-horizon regressions holding δ_z fixed.

The LHIVX is a newly introduced test and we are not aware of any finite sample power results for the test even in the linear model, so it is interesting to compare its performance across horizons and choices of δ_z. Figure 8 provides this comparison in the linear predictive model for $T = 1048$ and $c = -2.5$. Of the two long-horizon tests, it is interesting to note that using $v = 0.80$ ($k = 261$) with $\delta_z = 0.75$ has both better size and power than the test employing $v = 0.90$ ($k = 525$). This result was not expected, but it may be that $v = 0.90$ comes too close to the restriction that $v < 1$. In fact, the LHIVX test with $v = 0.80$ and $\delta_z = 0.75$ has very good local power, dominating the short-horizon IVX test with the same value of δ_z for $\gamma < 200$. However, when using a larger value of $\delta_z = 0.85$, the short-horizon test does as well as the long-horizon test near the null hypothesis and provides better power further into the alternative. These results are interesting in that they highlight the important interplay between the choice of δ_z and the horizon. Yet, overall they still suggest that with a good choice of δ_z the short-horizon IVX can provide as much or more power than its long-horizon counterpart.

In Figure 9, we next investigate the power of the LHIVX test when the data is generated by the regime switching model. The general conclusions do not change. The LHIVX test with $v = 0.80$ ($k = 250$) and the short-horizon IVX test with $\delta_z = 0.85$ are equally good and improve on the other two tests near the null hypothesis, while the short-horizon test with $\delta_z = 0.85$ appears to be the best further along the power curve.

Overall, when the short-horizon test is efficient or based on a good selection of δ_z, our simulations show little evidence of power improvements from longer horizon tests in either the linear predictive regression or the regime switching predictive regression model.

8 Conclusion

There has been much debate in the literature regarding the presumed power advantages of long-horizon regressions. With a few notable exceptions, most of this debate has taken place within the context of a linear predictive model as the true data generating process. In this case, long-horizon regression tests may have power advantages over short-horizon tests that do not fully exploit the information in the model, especially the contemporaneous endogeneity between the regressor innovation and the regression residual. On the other hand, it seems unlikely that they would have advantages over more efficient short-horizon tests when a linear short-horizon model is the true model.

The class of nonlinear models is so large that it would not be surprising if one could find nonlinear models which are better approximated by long-horizon than short-horizon regression models. However, this would only seem to provide practical support for the long-horizon approach if both the nonlinear model that supports it and its parameter values are realistic. This argues in favor of a comparison based on simulations from an estimated nonlinear model that has already received support in the empirical literature.

In this paper we compare the power of short and long-horizon tests in both the linear model and an empirically plausible nonlinear model, involving regime switching predictive regression with time varying transition probabilities that are also a function of the predictor. Previous empirical literature has shown this model to be successful in capturing the bull and bear states long observed by market participants. We use our empirical estimates to anchor our simulations in a realistic manner.

We then compare the power of short and long-horizon tests when the true model is given by the estimated regime switching predictive regression model. As a point of reference, we provide similar comparisons when the data is generated by a linear regression model. We make these comparisons using two different versions of the long-horizon regressions: the standard long-horizon regression in (2), which has a long-horizon dependent variable and the rearranged long-horizon regression in (19), which uses a long-horizon predictor. Accordingly, we also consider two different recent long-horizon tests, the Hjalmarsson (2011) test and the Phillips and Lee (2013) long-horizon IVX (LHIVX) test. In addition, we compare size adjusted power for the influential, though size-distorted, OLS based tests used in much of the earlier empirical work. We also make comparisons using an infeasible, oracle version of the Hjalmarsson (2011) test.

Taken as a whole, our results are not overly supportive of the long-horizon regressions. In the case of the standard long-horizon specification, we find that that long-horizon regressions can provide at best a modest improvement in size-adjusted power using OLS based tests. However, in more powerful versions of the tests, such as the oracle version of the Hjalmarsson (2011) test, or even its Bonferroni counterpart, the short-horizon test shows better power in both the linear and regime switching predictive tests.

In the context of the rearranged regression, the power of the LHIVX test depends on both the horizon and the parameter δ_z, which determines the extent of the mild filtering used in the creation of the instrument. The theoretical results provide more flexibility in selecting this parameter at short horizons than in long horizons. For some choices of δ_z the long-horizon IVX test can provide more power close to the null hypothesis than its short-horizon counterpart. However, with a better choice of δ_z, we find that the short-horizon IVX is generally more powerful than the LHIVX, even close to the null hypothesis. The advantage to the short-horizon test increases as we move further from the null hypothesis.

Overall, our results support the contention that the higher power of long-horizon regressions may simply be a "myth" as argued by Boudoukh *et al.* (2008). On the other hand, regime switching models are not the only plausibly nonlinear models that could be considered in making these comparisons. Contrary to our results, improved power has been found in long-horizon regressions when using a nonlinear ESTAR model as the data generating process [Kilian and Taylor (2003), Wohar and Rapach (2005)]. Models incorporating financial bubbles [Evans (1991), Phillips *et al.* (2014)] might also support the conjecture of improved power at longer horizons and this could be an interesting avenue for future research. More general theoretical results on the relative power of long horizon tests in nonlinear models might be possible in the general framework of Park and Phillips (1999).

References

Ang, A. and G. Bekaert (2007). Stock return predictability: Is it there? *Review of Financial Studies* **20**(3), 651–707.

Bollen, N. P. B., S. F. Gray and R. E. Whaley (2000). Regime switching in foreign exchange rates: Evidence from currency option prices. *Journal of Econometrics* **94**(1-2), 239–276.

Boudoukh, J., M. Richardson and R. Whitelaw (2008). The myth of long-horizon predictability. *The Review of Financial Studies* **21**(4), 1577–1605.

Cai, B. and J. Gao (2013). Hermite series estimation in nonlinear cointegrating models. Monash University, Department of Econometrics and Business Statistics, Working Paper 17/13.

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

Campbell, J. Y. (2001). Why long horizons? A study of power against persistent alternatives. *Journal of Empirical Finance* **8**, 459–491.

Campbell, J. Y. and M. Yogo (2006). Efficient tests of stock return predictability. *Journal of Financial Economics* **81**(1), 27–60.

Campbell, J. Y. and R. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* **1**(3), 195–228.

Cavanagh, C. L., G. Elliott and J. H. Stock (1995). Inference in models with nearly integrated regressors. *Econometric Theory* **11**, 1131–1147.

Chan, N. H. and C. Z. Wei (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* **15**(3), 1050–1063.

Chauvet, M. and S. Potter (2000). Coincident and leading indicators of the stock market. *Journal of Empirical Finance* **7**(1), 87–111.

Chen, W. W. and R. S. Deo (2009). The restricted likelihood ratio test at the boundary in autoregressive series. *Journal of Time Series Analysis* **30**(6), 618–630.

Cheung, Y.-W. and U. G. Erlandsson (2005). Exchange rates and Markov switching dynamics. *Journal of Business & Economic Statistics* **23**, 314–320.

Chinn, M. D. and R. A. Meese (1995). Banking on currency forecasts: How predictable is change in money? *Journal of International Economics* **38**(1-2), 161–178.

Cochrane, J. (1991). Volatility tests and efficient markets: A review essay. *Journal of Monetary Economics* **27**, 463–487.

Cochrane, J. H. (1999). New facts in finance. *Economic Perspectives, Federal Reserve Bank of Chicago* **23**(3), 36–58.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies* **21**(4), 1533–1575.

Cutler, D. M., J. M. Poterba and L. H. Summers (1991). Speculative dynamics. *Review of Economic Studies* **58**(3), 529–546.

Dewachter, H. (2001). Can Markov switching models replicate chartist profits in the foreign exchange market? *Journal of International Money and Finance* **20**(1), 25–41.

Ding, Z. (2012). *An Implementation of Markov Regime Switching Model with Time Varying Transition Probabilities in MATLAB*. Available at SSRN: http://ssrn.com/abstract=2083332 or http://dx.doi.org/10.2139/ssrn.2083332.

Elliott, G. (1998). On the robustness of cointegration methods when regressors have almost unit roots. *Econometrica* **66**, 149–158.

Engel, C. and J. D. Hamilton (1990). Long swings in the dollar: Are they in the data and do markets know it? *American Economic Review* **80**(4), 689–713.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review* **81**(4), 922–930.

Fama, E. and K. French (1988*a*). Dividend yields and expected stock returns. *Journal of Financial Economics* **22**, 3–25.

Fama, E. and K. French (1988*b*). Permanent and temporary components of stock prices. *Journal of Political Economy* **96**, 246–273.

Fama, E. F. and R. R. Bliss (1987). The information in long-maturity forward rates. *American Economic Review* **77**(4), 680–692.

Ghysels, E. and J. I. Miller (2014). On the size distortion from linearly interpolating low-frequency series for cointegration tests. Unpublished Manuscript.

Goetzmann, W. N. and P. Jorion (1993). Testing the predictive power of dividend yields. *Journal of Finance* **48**(2), 663–679.

Gonzalo, J. and J.-Y. Pitarakis (2012). Regime-specific predictability in predictive regressions. *Journal of Business & Economic Statistics* **30**(2), 229–241.

Guidolin, M. and A. Timmermann (2005). Economic implications of bull and bear regimes in U.K. stock and bond returns. *Economic Journal* **115**(500), 111–143.

Guidolin, M. and S. Hyde (2010). Can VAR models capture regime shifts in asset returns? A long-horizon strategic asset allocation perspective. *Journal of Banking and Finance* **36**(3), 695–716.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**, 357–384.

Hannan, E. J. and B. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistal Society Series B* **41**, 190–195.

Hjalmarsson, E. (2008). Interpreting long-horizon estimates in predictive regressions. *Finance Research Letters* **5**(2), 104–117.

Hjalmarsson, E. (2011). New methods for inference in long-horizon regressions. *Journal of Financial and Quantitative Analysis* **46**(3), 815–839.

Hjalmarsson, E. (2012). Some curious power properties of long-horizon tests. *Finance Research Letters* **9**, 81–91.

Jansson, M. and M. J. Moreira (2006). Optimal inference in regression models with nearly integrated regressors. *Econometrica* **74**, 681–714.

Jegadeesh, N. (1991). Seasonality in stock price mean reversion: Evidence from the U.S. and the U.K. *Journal of Finance* **46**, 1427–1444.

Juhl, T. (2011). A nonparametric test of the predictive regression model. Department of Economics, The University of Kansas.

Kasparis, I., E. Andreou and P. C. B. Phillips (2012). Nonparametric predictive regression. Cowles Foundation Discussion Papers 1878. Cowles Foundation for Research in Economics, Yale University.

Kilian, L. (1999). Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions? *Journal of Applied Econometrics* **14**(5), 491–510.

Kilian, L. and M. P. Taylor (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics* **60**(1), 85–107.

Kim, M. J. and C. R. Nelson (1993). Predictable stock returns: The role of small sample bias. *Journal of Finance* **48**, 641–662.

Liu, W. and A. Maynard (2007). A new application of exact non-parametric methods to long-horizon predictability tests. *Studies in Nonlinear Dynamics & Econometrics* **11**(1), Article 7.

Magdalinos, T. and P. C. B. Phillips (2009). Econometric inference in the vicinity of unity. CoFie Working Paper (7), Singapore Management University.

Maheu, J. M. and T. H. McCurdy (2000). Identifying bull and bear markets in stock returns. *Journal of Business & Economic Statistics* **18**(1), 100–112.

Maheu, J. M., T. H. McCurdy and Y. Song (2012). Components of bull and bear markets: Bull corrections and bear rallies. *Journal of Business & Economic Statistics* **30**(3), 391–403.

Mark, N. C. (1995). Exchange rates and fundamentals: Evidence on long-horizon predictability. *American Economic Review* **85**(1), 201–218.

Mark, N. C. and D. Sul (2006). The power of long-horizon predictive tests. Technical report. Department of Economics, Ohio State University.

Marmer, V. (2008). Nonlinearity, nonstationarity, and spurious forecasts. *Journal of Econometrics* **142**(1), 1–27.

Park, J. and P. C. B. Phillips (1999). Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory* **15**, 269–298.

Park, J. Y. and P. C. B. Phillips (1989). Statistical inference in regressions with integrated processes: Part 2. *Econometric Theory* **5**, 95–131.

Park, J. Y. and P. C. B. Phillips (1988). Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory* **4**(03), 468–497.

Perlin, M. (2012). *MS_Regress – The MATLAB Package for Markov Regime Switching Models*. Available at SSRN: http://ssrn.com/abstract=1714016 or http://dx.doi.org/10.2139/ssrn.1714016.

Phillips, P. C. B., S. Shi and J. Yu (2014). Specification sensitivity in right-tailed unit root testing for explosive behavior. *Oxford Bulletin of Economics and Statistics* **76**(3), 315–333.

Phillips, P. C. B. (1986). Understanding spurious regression in econometrics. *Journal of Econometrics* **33**, 311–340.

Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika* **74**, 535–547.

Phillips, P. C. B. (1988). Regression theory for near-integrated time series. *Econometrica* **56**, 1021–1043.

Phillips, P. C. B. (1991). Optimal inference in cointegrated systems. *Econometrica* **59**, 283–306.

Phillips, P. C. B. (2014). On confidence intervals for autoregressive roots and predictive regression. *Econometrica* **82**(3), 1177–1195.

Phillips, P. C. B. and B. E. Hansen (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* **57**, 99–125.

Phillips, P. C. B. and J. H. Lee (2012). Robust econometric inference with mixed integrated and mildly explosive regressors. Technical report. Department of Economics, Yale University.

Phillips, P. C. B. and J. H. Lee (2013). Predictive regression under various degrees of persistence and robust long-horizon regression. *Journal of Econometrics* **177**(2), 250–264.

Phillips, P. C. B. and T. Magdalinos (2007). Limit theory for moderate deviations from a unit root. *Journal of Econometrics* **136**(1), 115–130.

Royal Swedish Academy of Sciences (2013). The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2013: Eugene F Fama, Lars Peter Hansen, Robert J Shiller. Press Release.

Schaller, H. and S. V. Norden (1997). Regime switching in stock market returns. *Applied Financial Economics* **7**(2), 177–191.

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

Stock, J. H. (1991). Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series. *Journal of Monetary Economics* **28**, 435–439.

Valkanov, R. (2003). Long-horizon regressions: theoretical results and applications. *Journal of Financial Economics* **68**, 201–232.

Wohar, M. E. and D. E. Rapach (2005). Valuation ratios and long-horizon stock price predictability. *Journal of Applied Econometrics* **20**(3), 327–344.

Table 1: Estimated Model Coefficients in Regime Switching Models without Predictors

	Single State Model	Two State Model		Three State Model		
s_t	N/A	Bull	Bear	Bull	Normal	Bear
$\hat{\beta}_0(s_t)$	0.0051	0.0103	-0.0223	0.0105	0.0027	-0.0284
	(0.0014)	(0.0010)	(0.0078)	(0.0011)	(0.0042)	(0.0123)
$\hat{\sigma}_1^2(s_t)$	0.0021	0.0008	0.0074	0.0008	0.0021	0.0095
	(0.2184)	(0.0001)	(0.0000)	(0.0000)	(0.0003)	(0.0009)
H-Q IC	-3.3378	-3.6928		-3.6797		

The table shows the estimated expected return $\hat{\beta}_0(s_t)$ and variance $\hat{\sigma}_1^2(s_t)$ for each possible value of the state s_t. In other words, it shows the estimated coefficients in a restricted version (5), in which $\beta_1(s_t) = 0$. Standard errors are given in parenthesis. The bottom line provides the value of the Hannan and Quinn (1979) information criteria (H-Q IC), with a lower value indicating a better penalized fit.

Table 2: Transition Probability Model Coefficients in Regime Switching Model without Predictors

Intercepts					
	Two State Model		Three State Model		
s_t/s_{t-1}	Bull (i=1)	Bear (i=2)	Bull (i=1)	Normal (i=2)	Bear (i=3)
	$\widehat{\theta}_{0,11}$	$\widehat{\theta}_{0,21}$	$\widehat{\theta}_{0,1j}$	$\widehat{\theta}_{0,2j}$	$\widehat{\theta}_{0,3j}$
Bull (j=1)	1.8780	-1.0058	2.2325	-1.9328	-1.5921
	(0.1264)	(0.1459)	(0.1326)	(0.3365)	(0.3114)
Normal (j=2)			-1.9567	2.2392	-1.4603
			(3.4577)	(0.6587)	(0.3455)

The table entries show the estimated intercepts in (8), the model for the transition probabilities in (5), which are non-time varying in the absence of a predictor. Standard errors are given in parenthesis.

Table 3: Coefficient Estimates in Regime Switching Predictive Regression Model with Time Varying Transition Probabilities

	Single State Model	Two State Model		Three State Model		
$s_t =$:	N/A	Bull	Bear	Bull	Normal	Bear
$\hat{\beta}_0(s_t)$	0.0057	0.0110	-0.0274	0.0153	0.0028	-0.0028
	(0.0015)	(0.0012)	(0.0089)	(0.0013)	(0.0034)	(0.0095)
$\hat{\beta}_1(s_t)$	0.0018	0.0020	0.0144	0.0036	0.0032	0.0065
	(0.0013)	(0.0010)	(0.0070)	(0.0012)	(0.0029)	(0.0062)
$\hat{\sigma}_1^2(s_t)$	0.0021	0.0009	0.0074	0.0007	0.0014	0.0075
	(0.0010)	(0.0001)	(0.0010)	(0.0001)	(0.0002)	(0.0007)
H-Q IC	-3.3396	-3.6911		-3.6478		

The table entries show the estimates of the coefficients in the regime switching predictive regression model (5) for each possible value of the state s_t. Standard errors are given in parenthesis. The bottom line provides the value of the Hannan and Quinn (1979) information criteria (H-Q IC), with a lower value indicating a better penalized fit. Note that the inclusion of the dividend price ratio as a predictor may induce a predictive regression problem, a solution for which has yet to be provided in this regime switching model. Therefore, standard errors can only be relied on as a rough indicator of the estimation variance, but should not be employed to draw formal inferences.

Table 4: Time Varying Transition Probability Model Coefficients in Regime Switching Predictive Regression Model

Intercepts					
	Two State Model		Three State Model		
s_t/s_{t-1}	Bull (i=1)	Bear (i=2)	Bull (i=1)	Normal (i=2)	Bear (i=3)
	$\widehat{\theta}_{0,11}$	$\widehat{\theta}_{0,21}$	$\widehat{\theta}_{0,1j}$	$\widehat{\theta}_{0,2j}$	$\widehat{\theta}_{0,3j}$
Bull (j=1)	2.2589	-1.0872	2.0002	-2.0033	-1.9870
	(0.2272)	(0.2035)	(0.2090)	(0.3422)	(0.2724)
Normal (j=2)			-1.9965	2.0148	-1.9790
			(2.6658)	(0.3991)	(0.3856)
Slope Coefficients					
	Two State Model		Three State Model		
s_t/s_{t-1}	Bull (i=1)	Bear(i=2)	Bull (i=1)	Normal (i=2)	Bear (i=3)
	$\widehat{\theta}_{1,11}$	$\widehat{\theta}_{1,21}$	$\widehat{\theta}_{1,1j}$	$\widehat{\theta}_{1,2j}$	$\widehat{\theta}_{1,3j}$
Bull (j=1)	0.3356	-0.3296	0.0139	0.0063	-0.0145
	(0.1478)	(0.1732)	(0.1675)	(0.4277)	(0.3006)
Normal (j=2)			-0.0033	-0.0023	-0.0095
			(1.9531)	(0.3386)	(0.4061)

The table entries show the estimates of the coefficients in (8), the model for the time transition probabilities in the regime switching predictive regression model (5). Standard errors are given in parenthesis. Note that the inclusion of the dividend price ratio as a predictor may induce a predictive regression problem, a solution for which has yet to be provided in this regime switching model. Therefore, standard errors can only be relied on as a rough indicator of the estimation variance, but should not be employed to draw formal inferences.

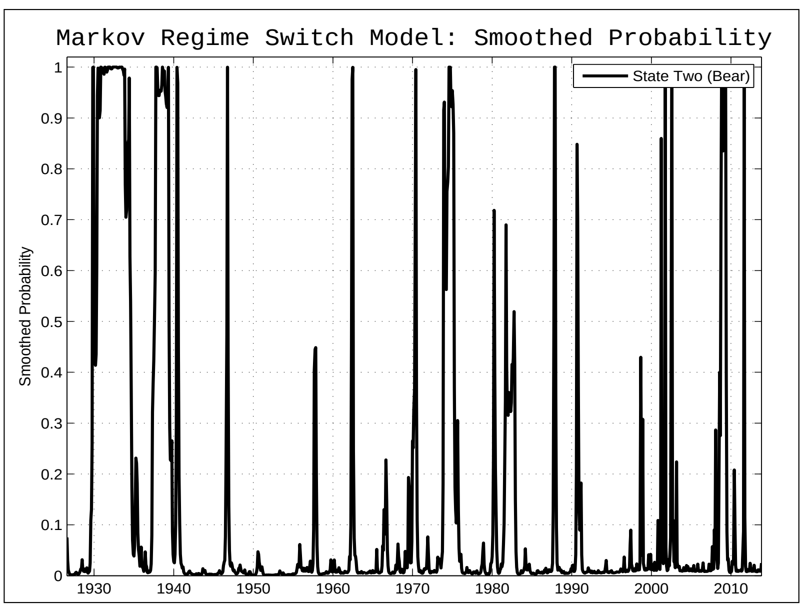

Figure 1: Fitted Bear State Probabilities in Two State Regime Switching Predictive Regression Model with Time Varying Transition Probabilities.

Major Market Downturns: 1929-1933 The Great Depression; 1937-1938 Recession; 1945 End of the War Recession; 1960-1961 Recession; 1969-1970 Recession; 1973-1975 Oil Crisis; Early 1980s Recession; Early 1990s Recession; Early 2000s Recession; Dec 2007-June 2009 Great Recession.

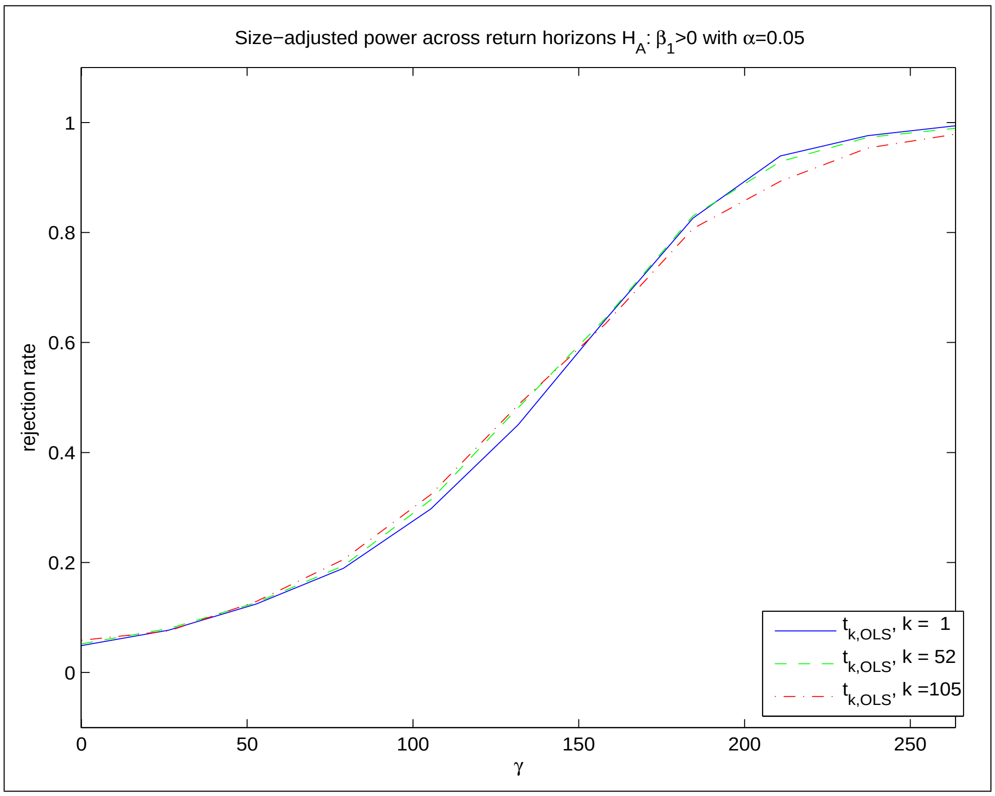

Figure 2: The size-adjusted power of the OLS based predictive test when the return is generated from the linear predictive regression model in (7), the predictor is generated by the local-to-unity process in (11) and (12) with $c = -2.5$, and the residuals are generated from (13) using a multivariate normal distribution with $\delta = -0.95$. We set $T = 1048$.

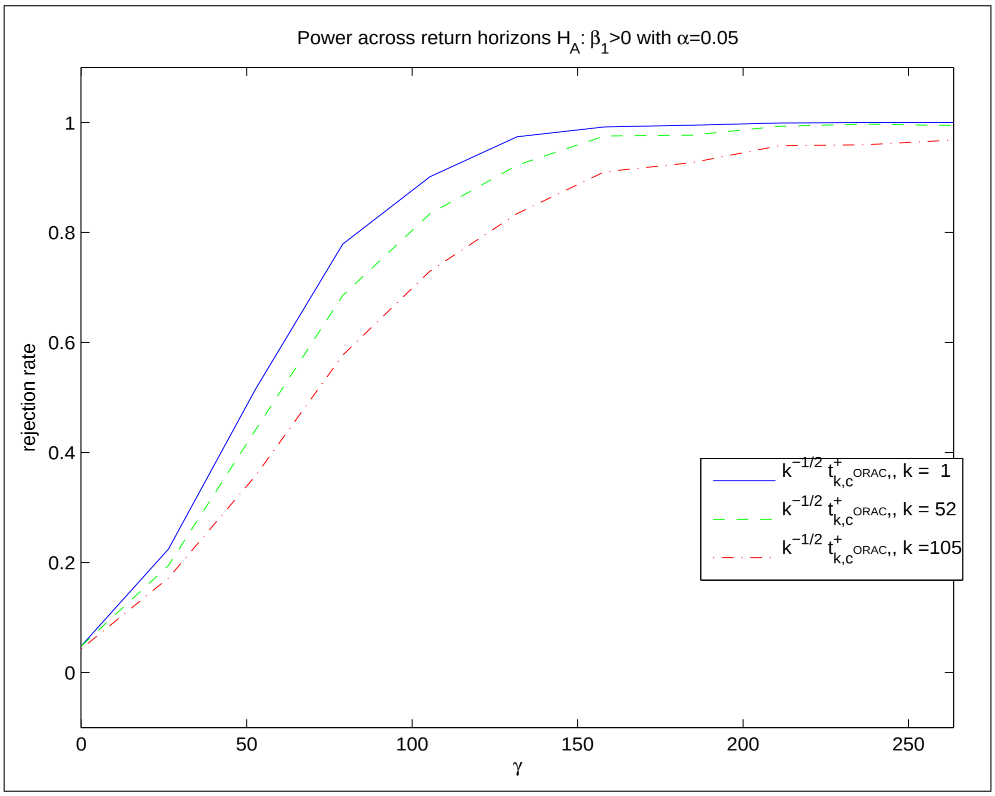

Figure 3: The power of the oracle based predictive test when the return is generated from the linear predictive regression model in (7), the predictor is generated by the local-to-unity process in (11) and (12) with $c = -2.5$, and the residuals are generated from (13) using a multivariate normal distribution with $\delta = -0.95$. We set $T = 1048$.

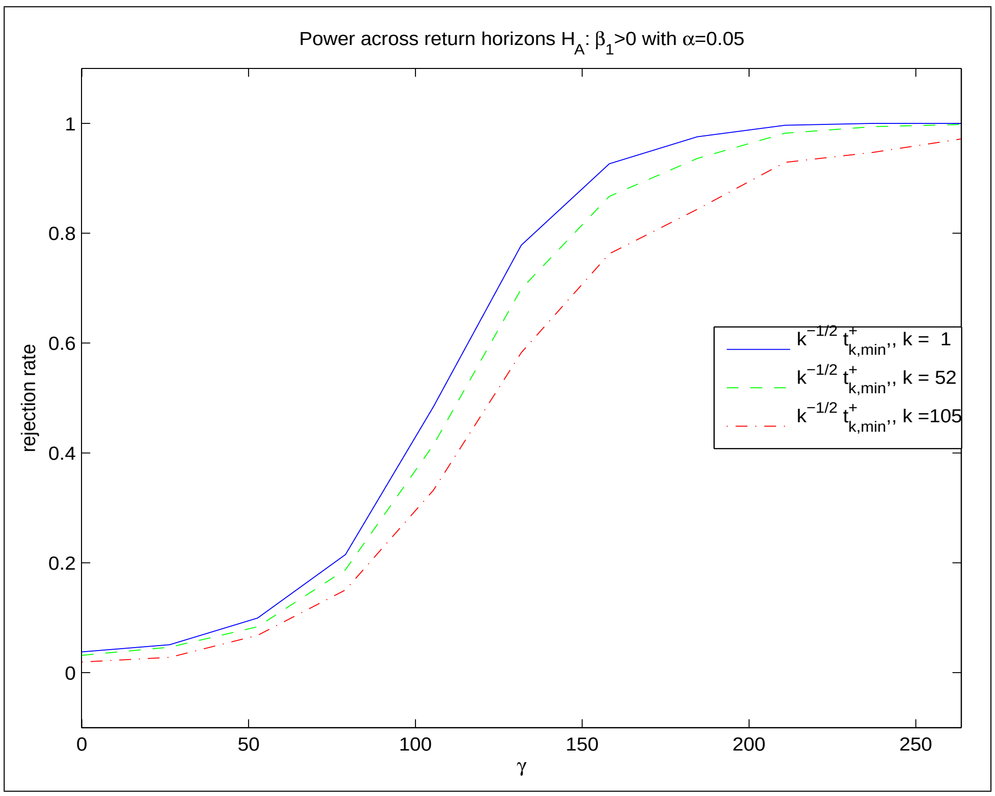

Figure 4: The power of the Bonferroni based predictive test when the return is generated from the linear predictive regression model in (7), the predictor is generated by the local-to-unity process in (11) and (12) with $c = -2.5$, and the residuals are generated from (13) using a multivariate normal distribution with $\delta = -0.95$. We set $T = 1048$.

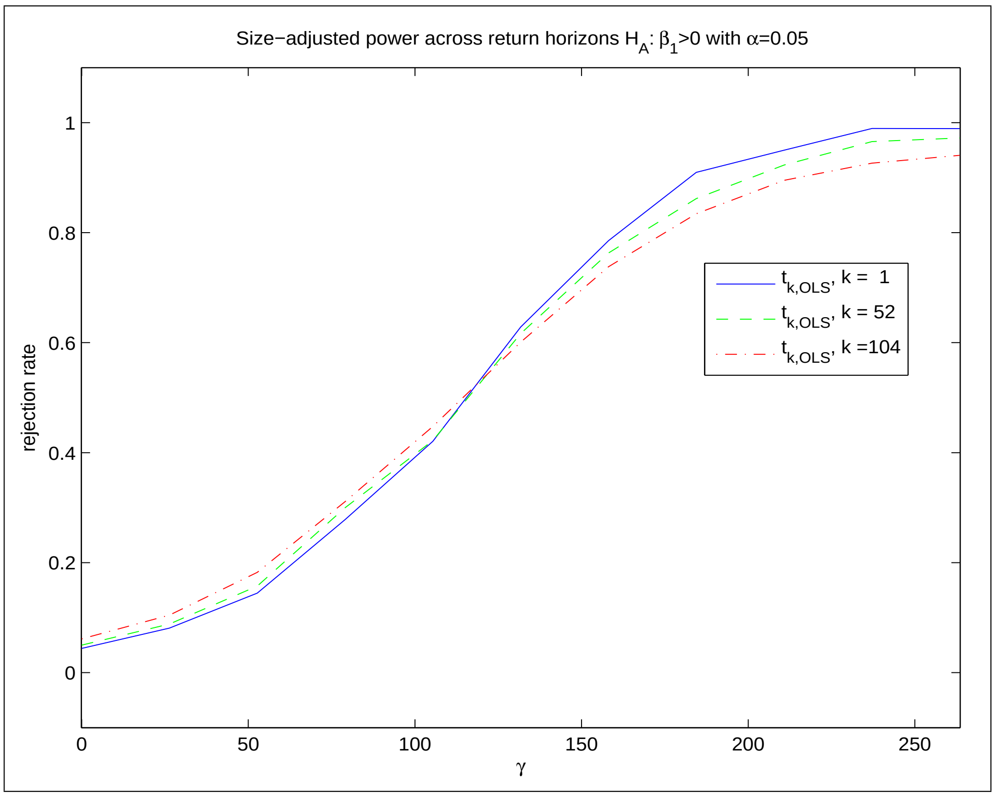

Figure 5: The size-adjusted power of the OLS based predictive test when the true return model is the two-state regime switching model in (5) with time varying transition probabilities given by (8), the predictor is generated by the local-to-unity process in (11) and (12) with $c = -2.5$, and the residuals are generated from (13) using a multivariate normal distribution with $\delta = -0.95$. We set $T = 1048$.

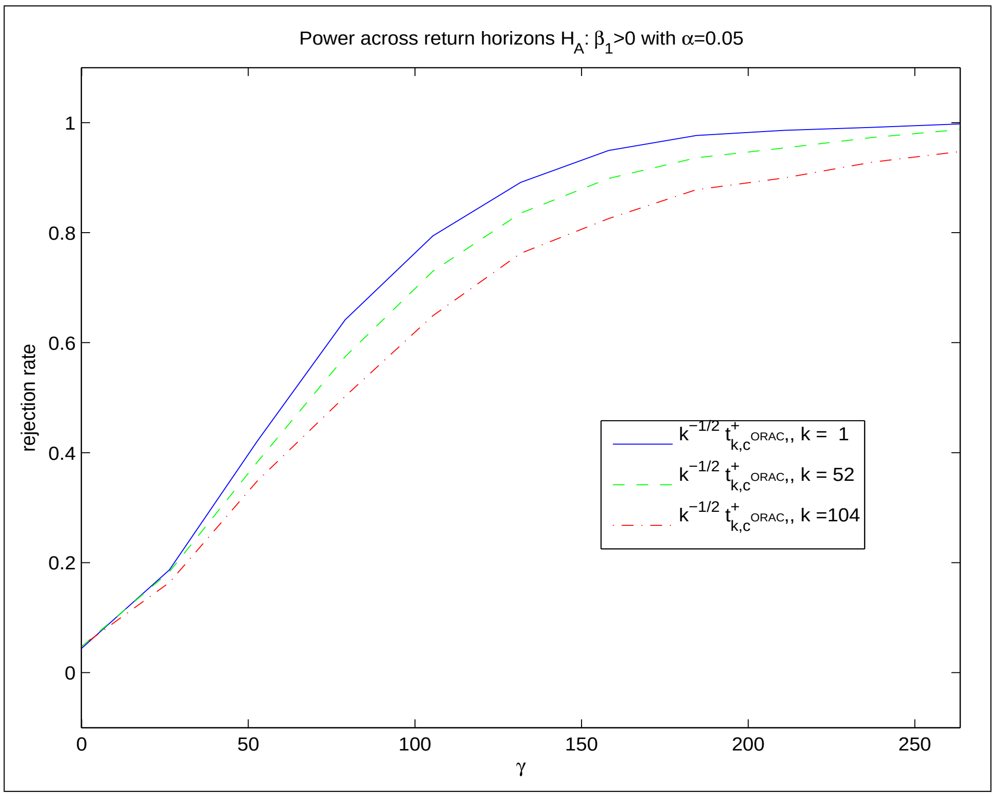

Figure 6: The power of the oracle based predictive test when the true return model is the two-state regime switching model in (5) with time varying transition probabilities given by (8), the predictor is generated by the local-to-unity process in (11) and (12) with $c = -2.5$, and the residuals are generated from (13) using a multivariate normal distribution with $\delta = -0.95$. We set $T = 1048$.

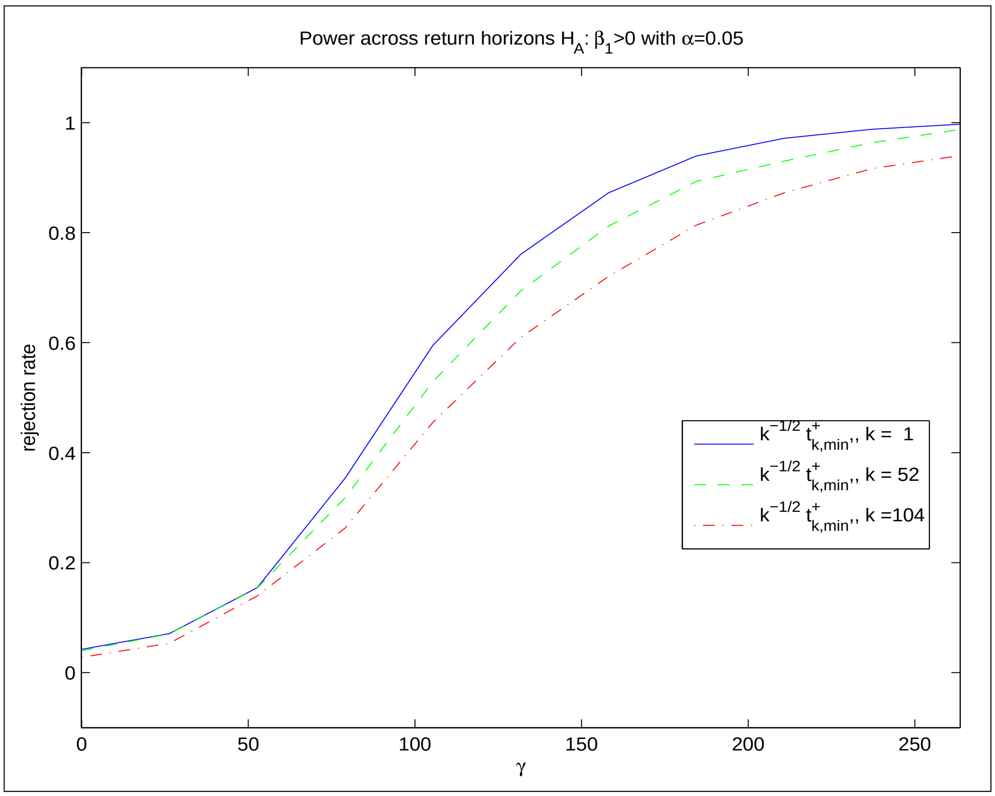

Figure 7: The power of the Bonferroni based predictive test when the true return model is the two-state regime switching model in (5) with time varying transition probabilities given by (8), the predictor is generated by the local-to-unity process in (11) and (12) with $c = -2.5$, and the residuals are generated from (13) using a multivariate normal distribution with $\delta = -0.95$. We set $T = 1048$.

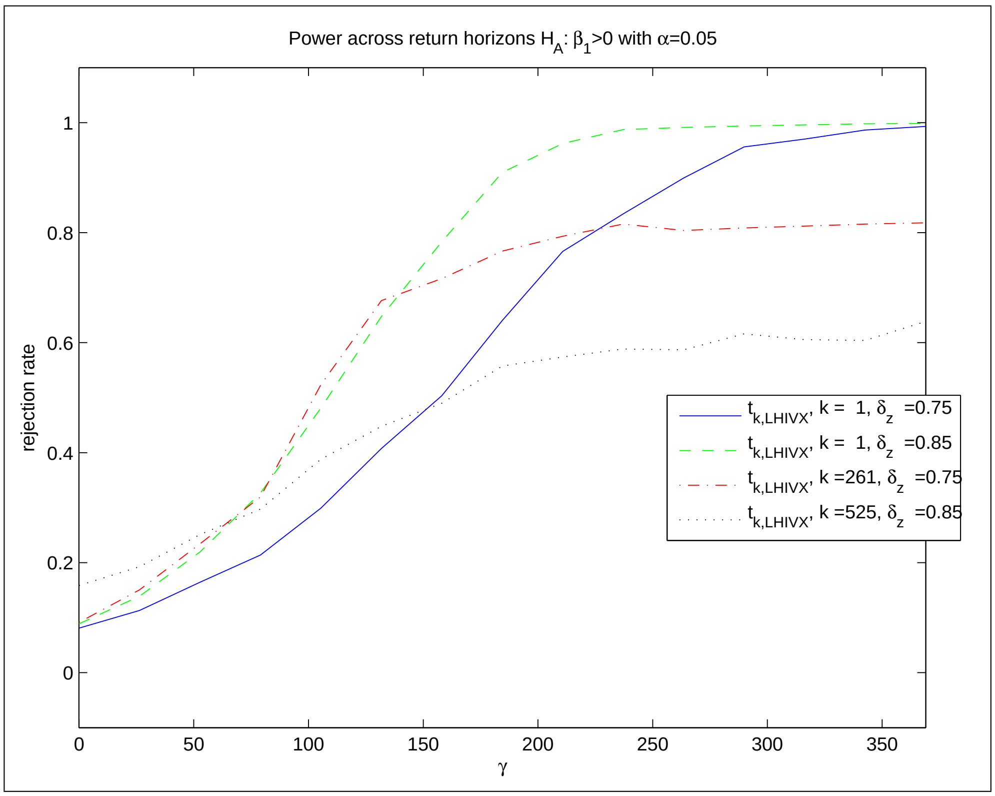

Figure 8: The power of the LHIVX test in a linear predictive regression model. We plot power for both $k = 1$ and for k given by (17) with $\lambda = 1$ and $v = 0.75$ and $v = 0.90$. The horizontal axis shows the value of γ in (29) with $v_2 = 0.80$. Here we set $c = -2.5$, $c_z = -5.0$, and $T = 1048$.

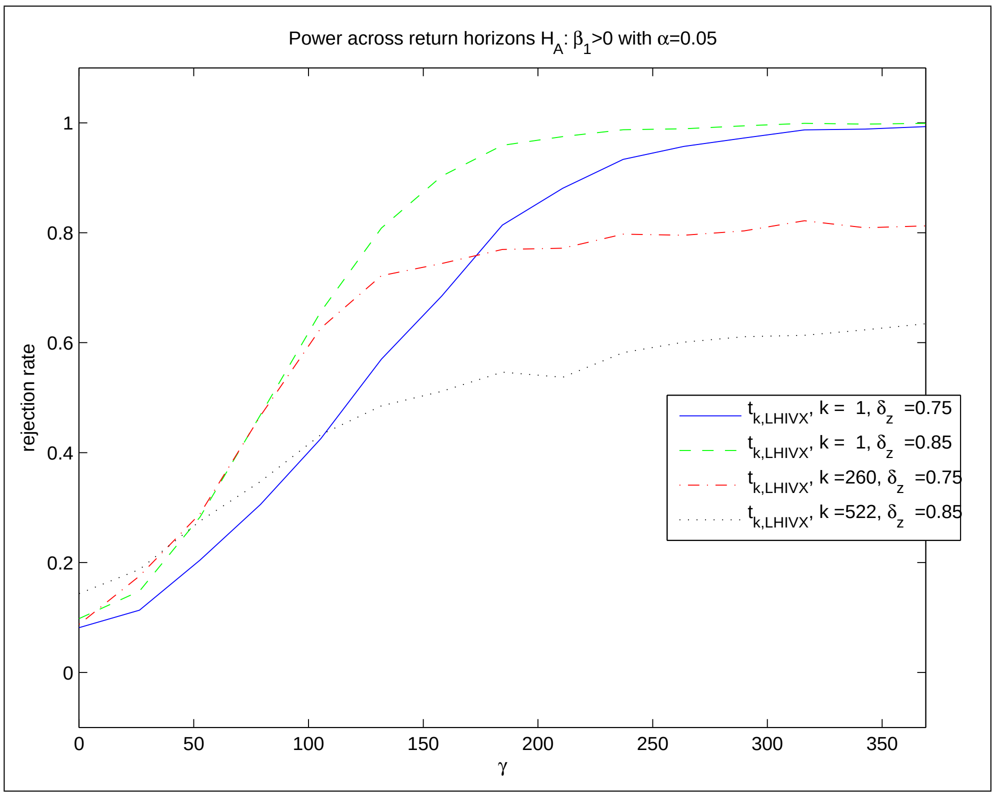

Figure 9: The power of the LHIVX test in the two-state regime switching model with time varying transition probabilities. We plot power for both $k = 1$ and for k given by (17) with $\lambda = 1$ and $v = 0.75$ and $v = 0.90$. The horizontal axis shows the value of γ in (29) with $v_2 = 0.80$. Here we set $c = -2.5$, $c_z = -5.0$, and $T = 1048$.