Unified Inference for Long-Horizon Predictive Regressions including mildly integrated and explosive cases*

Seok Young Hong[†] Daniel Tsvetanov[‡]

January 10, 2024

Preliminary and Incomplete

Comments Welcome

Abstract

We propose a unified procedure for testing the predictability of asset returns based on the empirical likelihood method. We make novel econometric contributions by allowing the predictor variable in our unified test to be mildly integrated or mildly explosive in addition to the usual persistence classes permitted in the literature: stationary, locally integrated, and unit root cases. Moreover, we extend this robust procedure to study the long-horizon predictive regression model, which has received much attention as a suitable alternative to examine return predictability. We report results from an empirical application on the US stock market, where we found evidence of predictability by the Treasury bills and the inflation rate. A simulation study confirms our proposed test performs very well in finite sample, exhibiting much better size properties than the IVX approach of Phillips and Lee (2013) and Kostakis, Magdalinos, and Stamatogiannis (2023).

JEL Classification: C12; C32; C51; C52

Keywords: Predictive Regression; Time Series; Empirical Likelihood

*We are very grateful to Oliver Linton for providing much useful advice. We also thank helpful comments and suggestions from participants at the 2023 Symposium on Economics and Finance in the Asia-Pacific.

[†]Nanyang Technological University, Singapore 639798, Singapore ; seokyoung.hong@ntu.edu.sg.

[‡]Norwich Business School, University of East Anglia, Norwich NR4 7TJ, UK; D.Tsvetanov@uea.ac.uk.

1. Introduction

The predictability of asset returns has been a crucial topic in financial economics, and its test has been a longstanding subject of research. Indeed, the conditional first moment properties of financial time series offer an insight into many fundamental issues and the empirical puzzles in finance. Consequently, there has been a huge demand for robust econometric methods for the study of return predictability (see, *inter alia*, Pesaran and Timmermann (1995)).

A standard approach for conducting inference on predictability is to employ a linear univariate predictive regression model, where a response variable Y is predicted by some lagged predictor X. In most empirical application, Y would typically represent the excess return on the stock market portfolio, in which case X may be the earnings-price ratio, the dividend yield, the inflation rate, the interest rates, and so on. The reader is referred to Phillips (2015) for an excellent survey. Specifically, the one-period liner predictive model stipulates that the true data generating process (DGP) for the vector of observations $\{(Y_t, X_t)\}_{t\in\mathbb{Z}}$ is given by:

$$Y_{t+1} = \alpha + \beta X_t + u_{t+1}, \tag{1}$$

$$X_{t+1} = \theta + \rho X_t + \varepsilon_{t+1} \tag{2}$$

where u_t and ε_t are some error terms. The object of interest is to test the null hypothesis of no predictability $H_0: \ \beta = 0$, and/or to construct a confidence interval for β. As an extension of this one-period model, there has been a great deal of interest in the long-horizon framework, where the predictability over a multiple-time period is examined, since the seminal work of Campbell and Shiller (1988), Mishkin (1990), and Boudoukh and Richardson (1993). Prediction over a long-horizon can be much useful in studies involving the forward premium, dividends, and stock returns. From an econometric point of view, the importance of long-horizon models also stems from the fact that many empirical works on short-horizon predictive regression report rather inconclusive findings and low explanatory power, Phillips and Lee (2013). Perhaps due to the additional complications involved, which we shall revisit in Section 2.1 below, the econometric literature on long-horizon predictability has been relatively small, and this is one of the contributions we aim to add in this paper.

Despite their wide applicability and usefulness, predictive regression models face several econometric challenges. When the sequence $\{X_t\}$ is stationary (i.e. $|\rho| < 1$), one may straightforwardly implement the simple least squares estimation method. However, as Stambaugh (1999) points out, when there is dependence between the two error terms u_t and ε_t the least squares estimator for β is biased in finite sample and therefore the inference becomes unreliable. See also Amihud and Hurvich (2004), Amihud, Hurvich, and Wang (2009), and Chen and Deo (2009) for further discussions and the bias-correction methods they propose. On the other hand, when $\{X_t\}$ is nonstationary (say unit root, nearly integrated, mildly integrated), which is often the case for macroeconomic predictors, the limiting behavior of the sample means $T^{-1}\sum_{t=1}^{T} X_{t-1}^k$, $k = 1, 2$ is completely different from that of the stationary case, see Campbell and Yogo (2006) and Cai and Wang (2014). Lastly, the asymptotic limit is also heavily affected by whether the variance of X is finite or infinite.

In practice, it is utterly difficult to know the exact temporal properties of X *a priori*. Therefore, it is imperative for the econometrician to have a unified test/estimation theory that is valid irrespective of whether X is stationary or nonstationary, or has a finite or infinite variance. Consequently, several groups of work have been proposed to construct inference procedures robust to different types of persistence. The first ones are the Bonferroni t-test by Cavanagh, Elliott, and Stock (2009), and the Bonferroni Q-test by Campbell and Yogo (2006). As Phillips and Lee (2013) point out however, it is difficult for Bonferroni type methods to allow for multiple number of regressors, and they have undesirable finite sample properties due to having non-standard limit distribution. Furthermore, they require a joint normality assumption on the error terms u and ε in (1) and (2), which can be restrictive in practice.

The second strand of research on the unified methods is the instrumental variable estimation (IVX) method by Magdalinos and Phillips (2009). The approach consists of filtering the predictor to construct an instrumental variable, whose degrees of persistence are explicitly controlled, and this variable is then used in the predictive regression of interest. The test follow a standard chi-squared limit irrespective of the degree of persistence of the original variable. Kostakis, Magdalinos, and Stamatogiannis (2015) study the IVX based Wald test in the context of predictive regression, on which Demetrescu, Georgiev, Rodrigues, and Taylor (2023) make extensions, and Phillips and Lee (2016) consider the cases of local unit roots in the explosive direction and mildly explosive roots. Demetrescu and Rodrigues (2022) study the bias correction analogous to Amihud and Hurvich (2004) in the IVX method. Yang, Long, Peng, and Cai (2020) investigate a new instrumental variable based Wald test (IVX-AR) which accounts for serial correlation and heteroscedasticity in the error terms. Recently, Demetrescu, Rodrigues, and Taylor (2023) develop IVX tests for long-horizon predictability. The IVX method however, has practical limitations in the sense that a tuning parameter needs to be chosen and the rate of convergence is rather slow.

An alternative method was proposed by Zhu, Cai, and Peng (2014), who use an empirical likelihood procedure using some weighted score equations. It is a nonparametric approach based on the seminal work by Owen (1988, 1990), and has been proven to be a very effective tool in this context. The test is extremely useful as it has a chi-squared limit regardless of the degree of persistence of the predictor X_t, does not need the selection of the tuning parameters, and has a faster convergence rate than the IVX methods. Some extensions have been made thereafter within the context of a short horizon predictability. Li, Li, and Peng (2017) allow the error term u_t to follow an AR(p) process, Liu, Yang, Cai, and Peng (2019) include the difference of the predicting variable in the regression model, and lastly Yang, Liu, Peng, and Cai (2021) permit the existence of a lagged predicted variable.

Despite the advantages of the empirical likelihood method, to the best of our knowledge, there is no paper that allows for mildly integrated or mild explosive regressors, which limits the applicability of the approach. Furthermore, the long-horizon predictive regression model, which has received much attention as a suitable alternative to examine return predictability, has not yet been explored. In this paper, we fill this gap by proposing a unified empirical likelihood test for long-horizon predictability which is valid regardless of whether the predictor is stationary, nearly stationary, unit root, mildly integrated or mildly explosive, which cover virtually all possible scenarios the econometrician may face.

Section 2 presents a detailed description of the method and the related asymptotic analysis, Section 3 reports simulation results, and Section 4 concludes. All proofs are contained in the Appendix. As for notations, we denote by $\Longrightarrow$ and $\xrightarrow{d}$ weak convergence in the Skorohod space $\mathcal{D}[0,1]$ and convergence in law, respectively, and take the term 'stationarity' to mean strict stationarity. Throughout, C (or C', C'') refers to some generic constant that may take different values in different places unless defined otherwise.

2. Methodological framework and asymptotic results

2.1 Long-horizon predictive regressions

Since Fama and French (1988) and Campbell and Shiller (1988), the return predictability over a multi-step time horizon has received a considerable interest in finance. Several econometric methods have been proposed in the literature thereafter. Volkanov (2003) studies an asymptotically valid procedure and asserts that long-horizon regressions always return "significant" results.

In the econometric analysis that follows, we assume that the true data generating process (DGP) for the vector of observations $\{(Y_t, X_t)\}_{t\in\mathbb{Z}}$ is given by the one-period liner predictive model given in (1) and (2) above. We make the following very general assumption about u_t and

ε_t, which allows for a wide range of empirical features that are often encountered in practice.

Assumption 1. *(i) The errors* u_t *and* ε_t *are characterised as*

$$u_t = \phi v_t + \epsilon_t \tag{3}$$

$$\varepsilon_t = b_1 \varepsilon_{t-1} + \cdots + b_{p-1} \varepsilon_{t-p+1} + v_t \tag{4}$$

where the lag polynomial $B(L) := 1 - b_1 L - \cdots - b_p L^p$ *is invertible;*

(ii) $(\epsilon_1, v_1), \ldots, (\epsilon_T, v_T)$ *are independent and identically distributed (iid) random vectors and* $\mathrm{E}(u_1) = 0$, $\mathrm{E}(\epsilon_1) = 0$, $\mathrm{E}(|u_1|^{2+q} + |\epsilon_1|^{2+q}) < \infty$ *for some* $q > 0$.

To allow for a wide range of possibilities, it is common in the literature to consider the case where X_t is a scalar and to define the autoregressive coefficient $\rho = \rho_T$ in equation (2) as

$$\rho \;=\; 1 + \frac{c}{T^a}, \tag{5}$$

where c and a are some constants, with a being non-negative. It is evident that the time series characteristics of X_t are determined by the values of the pair (c, a).

In this study, we will consider the following five cases:

C1 : $c < 0$ and $a = 0$; X_t is **stationary**;

C2 : $c < 0$ and $a \in (0, 1)$; X_t is **mildly integrated**;

C3 : $c \neq 0$ and $a = 1$; X_t is **near integrated**;

C4 : $c = 0$; X_t is **integrated**;

C5 : $c > 0$ and $a \in (0, 1)$; X_t is **mildly explosive**;

These cases cover a large spectrum of possible behavior that X_t may exhibit, and therefore, offer a considerable generality that is useful in both empirical and theoretical work. In particular, within the context of empirical likelihood method, there has been no study in the literature that allowed for mildly integrated **C2** and mildly explosive **C5** cases.

A standard approach in the literature to investigate the predictability of Y_t is to conduct inference about β using a short-horizon (one-period) regression model as the specification in equation (1). In this paper, we investigate the multiple horizon predictability via a long-run predictive regression specification that results from the h-period ($h > 1$) temporal aggregation of (1). That is,

$$Y_t(h) \;=\; \alpha_h + \beta_h X_t + e_{t+h}, \quad t = 1, \ldots, T - h \tag{6}$$

where $Y_t(h) = \sum_{j=1}^{h} Y_{t-h+j}$. It is important to note that this specification is a fitted regression and not a DGP (see Hjalmarsson, 2011). The process in equation (1) continues to be the true DGP for Y_t, and $Y_t(h)$ is simply its h-period accumulation.

If Y_t is unpredictable and the true value of β is zero, then $\beta_h = 0$ for every h. Similarly, $\beta_h \neq 0$ for every h whenever $\beta \neq 0$. The regression specification in equation (3) is therefore empirically useful, since a statistically significant estimate of b_h can be interpreted as evidence of long horizon predictability. Also, notice that if we set $h = 1$, equation (3) becomes a short-horizon predictive regression specification. We shall impose the following condition for the asymptotic derivation.

Assumption 2. *The horizon $h = h_T \to \infty$ satisfies $h/T^{q/2} \to 0$ as $T \to \infty$.*

As aforementioned in the introduction, there are several statistical obstacles in (6) that affect the asymptotic inference for the presence of predictability. First, the time series characteristics of X_t is unknown and it is often difficult in practice to determine the exact type of persistence class that it belongs to. In fact, when X_t belongs to class **C2**, **C3**, **C4**, or **C5**, standard inference methods are not valid, since the limit theory of the regression parameter estimators depends on the localizing constant c, which is not consistently estimable (see, e.g., Phillips, 2015). Second, when the innovations in equation (2) are highly correlated with the innovations in equation (1), the least squares estimator of β based on the short-horizon predictive regression is biased. This bias disappears asymptotically when X_t belongs to class **C1** but is present even in the limit when X_t is non-stationary. Moreover, when X_t belongs to either class **C2** or class **C3**, the bias cannot be corrected, since it is a function of the localizing constant c (Phillips and Lee, 2013). Finally, in the long horizon predictive regressions, the presence of overlapping observations due to the accumulation of Y_t generates a serial correlation in the innovations, which is not present in the short-horizon specification, and it causes the estimator of β_h to be inconsistent (Kostakis, Magdalinos, and Stamatogiannis, 2023).

Various econometric methods have been proposed in the literature to overcome the aforementioned issues. In this paper, we adopt and extend the method proposed by Zhu, Cai, and Peng (2014), who used an empirical likelihood procedure based on some weighted score equations to develop a test of short-horizon predictability. The test has a chi-squared limit regardless of the degree of persistence of the predicting variable X_t. We show that a unified inference can be extended even when X_t belongs to **C2** or **C5**, and also within the context of long-horizon predictability. A detailed description of the method and the related asymptotic analysis is provided in the following section.

2.2 Empirical likelihood: Test description and asymptotic results

To fix ideas, let $h = 1$ in the regression model (6) and consider the following estimating equations:

$$\sum_{t=1}^{T-h} \left(Y_{t+h}(h) - \alpha_h - \beta_h X_t\right) = 0 \tag{7}$$

and

$$\sum_{t=1}^{T-h} \left(Y_{t+h}(h) - \alpha_h - \beta_h X_t\right) w\left(X_t\right) = 0 \tag{8}$$

where $w(\cdot)$ is some weight function. Following Zhu, Cai, and Peng (2014), we specify $w(X_t) = X_t/\sqrt{1 + X_t^2}$, since in this case $T^{-1}\sum_{t=1}^{T}|w(X_t)| \xrightarrow{p} 1$ as $T \to \infty$ for X_t being both stationary and non-stationary. When X_t belongs to class **C1** (i.e. when it is stationary), solving these equations yields the weighted OLS estimates of α_h and β_h. However, when X_t is mildly integrated, near integrated, unit root, or mildly explosive (i.e., cases **C2**-**C5**), the joint limit of $(T-h)^{-1}\sum_{t=1}^{T-h}(Y_{t+h}(h) - \alpha_h - \beta_h X_t)$ and $(T-h)^{-1}\sum_{t=1}^{T-h}(Y_{t+h}(h) - \alpha_h - \beta_h X_t)w(X_t)$ does not follow a bivariate normal distribution, since $(T-h)^{-1}\sum_{t=1}^{T-h} e_{t+h}w(X_t)$ converges in distribution to a random variable as $T \to \infty$ because of the constant (see, e.g., Chan, Li, and Peng, 2012). Thus, Wilks' theorem[1] in this case does not hold.

To overcome this issue, Zhu, Cai, and Peng (2014) proposed to split the sample and to difference the data with a large lag. In particular, let $m = \lfloor T/2 \rfloor$, where $\lfloor\cdot\rfloor$ denotes the floor function, and define $\widetilde{Y}_t = Y_{t+m} - Y_t$, $\widetilde{X}_t = X_{t+m} - X_t$, and $\widetilde{e}_t = e_{t+m} - e_t$, for $t = 1, \ldots, m$. Then, the regression model in equation (6) can be rewritten as

$$\widetilde{Y}_{t+h}(h) = \beta_h \widetilde{X}_t + \widetilde{e}_{t+h}, \quad \text{for} \quad t = 1, \ldots, m-h, \tag{9}$$

where $\widetilde{Y}_{t+h}(h) = \sum_{j=1}^{h} \widetilde{Y}_{t+j}$. Based on this model and under the conditions in Assumption 1, it can be shown that Wilks' theorem holds for the empirical likelihood method regardless of the degree of persistence of X_t when $h = 1$. If however $h > 1$, then the dynamics of $\widetilde{e}_{t+h}$ would change and they would no longer be independent. To address this difficulty, we instead apply

[1]Wilk's theorem states that the logarithm of the empirical liklihood ratio has an asymptotic chi-squared distribuion (see, for example, Hall and La Scala, 1990)

the empirical likelihood method, where $h \geqslant 1$, based on the following regression specification:

$$\widetilde{Y}_{t+h} = \beta_h^{\text{rev}} \widetilde{X}_t(h) + \widetilde{u}_{t+h}, \tag{10}$$

where $\widetilde{X}_t(h) = \sum_{j=0}^{h-1} \widetilde{X}_{t+j}$. This specification is a reverse regression of the h-period value of the dependent variable on the h-period sum of the regressor and has been used in the context of long-horizon predictability by Phillips and Lee (2013) and Wei and Wright (2013), among others. Note that the error term $\widetilde{u}_{t+h}$ is independent, which is necessary to ensure that Wilks' theorem holds.

Based on the regression in equation (10), we define the empirical likelihood function for β_h^{rev} as

$$\widetilde{L}_T\left(\beta_h^{\text{rev}}\right) = \sup\left\{\prod_{t=1}^{m-h}((m-h)\pi_t):\ \pi_1 \geqslant 0, \ldots, \pi_{m-h} \geqslant 0,\ \sum_{t=1}^{m-h}\pi_t = 1,\ \sum_{t-1}^{m-h}\pi_t \widetilde{Z}_t\left(\beta_h^{\text{rev}}\right) = 0\right\}, \tag{11}$$

where $\widetilde{Z}_t(\beta_h^{\text{rev}}) = [\widetilde{Y}_{t+h} - \beta_h^{\text{rev}}\widetilde{X}_t(h)]\widetilde{X}_t(h)/\sqrt{1+\widetilde{X}_t(h)^2}$. By the Lagrange multiplier technique, we have

$$\tilde{\ell}_T\left(\beta_h^{\text{rev}}\right) = -2\log\widetilde{L}_T\left(\beta_h^{\text{rev}}\right) = 2\sum_{t=1}^{m}\log\left\{1+\lambda\widetilde{Z}_t\left(\beta_h^{\text{rev}}\right)\right\}, \tag{12}$$

where $\lambda = \lambda(\beta_h^{\text{rev}})$ satisfies

$$\sum_{t=1}^{m}\frac{\widetilde{Z}_t\left(\beta_h^{\text{rev}}\right)}{1+\lambda\widetilde{Z}_t\left(\beta_h^{\text{rev}}\right)} = 0.$$

The following result shows that Wilks' theorem holds for the proposed empirical likelihood method.

Theorem 1. *Suppose that (i) the data is generated according to the process in equations* (1)*-*(2)*, (ii) Assumption 1 is satisfied, and (iii) the predictive variable belongs to either class* ***C1****,* ***C2****,* ***C3****,* ***C4****, or* ***C5****. Then,*

$$\tilde{\ell}_T(\beta_{h,0}^{rev}) \xrightarrow{d} \chi_1^2 \tag{13}$$

as $T \to \infty$, where $\beta_{h,0}^{rev}$ denotes the true value of β_h^{rev}.

As a consequence of Theorem 1, we would reject the hypothesis $H_0 : \beta_{h,0}^{\text{rev}} = 0$ at some significance level ϱ if $\tilde{\ell}_T(\beta_h^{\text{rev}}) > \chi^2_{1,1-\varrho}$. Alternatively, an empirical likelihood confidence interval for $\beta_{h,0}^{\text{rev}}$ with level ϱ can be obtained as in Zhu, Cai, and Peng (2014).

3. Simulation results

3.1 Simulation design

In this section, we use Monte Carlo simulations to investigate the finite sample behaviour of the empirical likelihood method discussed above. For all reported experiments, we generate the data from the process in equations (1) and (2), with $\alpha = 1$, $\theta = 1 - \rho$, and $c \in \{-50, -20, -10, 0, 1\}$. The autoregressive process in equation (2) was initialized at $x_0 = 0$. The vector $(\epsilon_t', v_t')'$ is drawn from a bivariate standard normal distribution and we set $\varepsilon_t = v_t$ and $u_t = \phi v_t + \epsilon_t$, with $\phi \in \{-0.95, -0.5, 0\}$. All tests considered are for the null hypothesis of no long run predictability and are run at the 5% nominal level of significance. To investigate their finite sample properties, we set $\beta = b/\sqrt{T}$, where b is allowed to increase from 0 to 10 in increments of 2. For $b = 0$ ($b \neq 0$), the results represent the finite sample size (power) of the tests. We report result for samples of length $T \in \{250, 500, 1000\}$ and prediction horizons $h \in \{1, 4, 12, 60\}$, using 10,000 Monte Carlo replications.

Kostakis, Magdalinos, and Stamatogiannis (2023) point out that the Stambaugh bias is primarily determined by the interaction between the regressor's degree of persistance and its endogeneity. Note that we are able to examine the effect of this bias on the finite sample properties of the proposed test statistics at different predictive horizons, since we consider in the simulation set up various combinations of values for the parameters c, ϕ, and h.[2] Also, as a base for comparison, we report results from a Monte Carlo study for the long-horizon predictability tests of Phillips and Lee (2013) and Kostakis, Magdalinos, and Stamatogiannis (2023), which we will denote as PL and KMS, respectively.

3.2 Size properties

First, we examine the size properties of the empirical likelihood test described in Section 2, with the α in equation (1) treated as either known or unknown, and the IVX-based tests of Phillips and Lee (2013) and Kostakis, Magdalinos, and Stamatogiannis (2023). Rejection probabilities are summarized in Table 1. We can draw few interesting conclusions from these simulated results. PL and KMS tests exhibit similar performance. This is not surprising given that both tests share a common modelling framework. Both tests appear to be slightly undersized in small samples when there is a correlation between the innovations of the predictor and the predicted variables (i.e., for $\phi \in \{-0.95, -0.5\}$) but their size properties tend to improve as the sample size increases. In contrast, the empirical likelihood test that treats the α as known

[2] We have also considered $\phi \in \{0.95, 0.5\}$ but the results are qualitatively similar to $\phi \in \{-0.95, -0.5\}$, so for the sake of brevity we do not report them here. These results are available on request though.

(EL1) exhibits correct size properties even for $T = 250$ and regardless of (i) the presence of correlation between the innovations of the predictor and the predicted variables and (ii) degree of predictor's persistence. Similar conclusion can be drawn for the empirical likelihood test that treats the α as unknown (EL2), except that for $\phi = -0.95$ and $c = 1$, the EL2 test appears to be slightly oversized. In general, however, both EL1 and EL2 tests show reliable size properties at all prediction horizons that we consider, which suggests that their performance is robust to the overlapping nature of observations in long horizons.

3.3 Power properties

We next compare the finite sample power properties of the proposed tests. Figures 1–3 contain power curves for $T = 250$ and $\phi \in \{-0.95, -0.5, 0\}$, Figures 4–6 contain power curves for $T = 500$ and $\phi \in \{-0.95, -0.5, 0\}$, and Figures 7–9 contain power curves for $T = 1000$ and $\phi \in \{-0.95, -0.5, 0\}$. All figures consist of subplots that correspond to the different combination of values that we assume for the persistence parameter c and the prediction horizon h. Overall, we can conclude that all four tests display comparable performance, especially for $T = 1000$. The power of the tests increases with the increase in the value of the localizing coefficient b, which is to be expected, since for small values of b, the departures from the null are too small to be detected. For $T < 1000$ and $c < 0$, the EL1 test tends to exhibit better power properties than the PL and KMS tests, especially when $k > 4$. However, for $c \geqslant 0$ the PL and KMS tests tend to perform slightly better. Also, the EL1 test that treats α as known is more powerful than the EL2 test. This is mainly due to the data splitting technique that we adopt for the EL2 test and is consistent with the analysis of Zhu, Cai, and Peng (2014), Chan, Li, and Peng (2012), Li, Li, and Peng (2017), and Liu, Yang, Cai, and Peng (2019).

In summary, the simulation results suggest that the proposed empirical likelihood method for testing long-horizon predictability can deliver an accurate size and nontrivial power. This holds regardless of the endogeneity issues, the overlapping nature of observations in long horizons, and the uncertainty regarding the exact type of predictor's persistence that econometritians usually face in practice.

4. Empirical application

This section revisits the evidence on the ability of financial and macroeconomic variables to predict stock market returns. Despite the voluminous literature on this subject, there is still a debate as to whether future stock returns are predictable or not. On one hand, studies like Lettau and Ludvigson (2001) argue that "...excess returns are predictable by variables such as dividend-price ratios, earnings-price ratios, dividend-earnings ratios, and an assortment of other financial indicators". On the other hand, however, studies like Welch and Goyal (2008) suggest that "... a healthy skepticism is appropriate when it comes to predicting the equity premium". We aim to shed some light on this debate by conducting a battery of short- and long-horizon predictability tests that we developed in this paper. We collect monthly data on the following eleven variables that are commonly used in the literature as predictors of the aggregate market: the dividend payout ratio; the long-term yield; the dividend yield; the dividend-price ratio; the Treasury bill rate; the earnings-price ratio; the book-to-market value; the default yield spread; the net equity expansion; the term spread; and the inflation rate. The data is obtained from Amit Goyal's website[3] and covers the period from January 1952 to December 2022. The tests are performed for $h \in \{1, 4, 12, 60\}$ and the dependent variable in all predictive regressions is the continuously compounded return of the CRSP value weighted index[4] in excess of the one-month Treasury bill rate.

The p-values of the proposed robust empirical likelihood tests are summarized in Table 2. We can see that the Treasury bill rate is the only significant predictor at the 5% level of significance and this is just for $h = 1$. There is also some weak evidence of predictability from the Treasury bill rate for $h = 4$, from the long term yield for $h = 1$, and the term spred for for $h = \{12, 60\}$. All other p-values are higher than conventional levels of significance. These conclusions are to a large extend consistent with the findings of Kostakis, Magdalinos, and Stamatogiannis (2015) and Kostakis, Magdalinos, and Stamatogiannis (2023). They also find evidence that the Treasury bill rate and the term spread have some predictive ability.

[3]See https://sites.google.com/view/agoyal145.

[4]The data for the CRSP index is obtained from Kenneth French's wensite: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html

References

Amihud, Yakov, and Clifford M. Hurvich, 2004, Predictive regressions: A reduced-bias estimation method, *Journal of Financial and Quantitative Analysis* 39, 813–841.

——— , and Yi Wang, 2009, Multiple-predictor regressions: Hypothesis testing, *Review of Financial Studies* 22, 413–434.

Boudoukh, Jacob, and Matthew Richardson, 1993, Stock returns and inflation: A long-horizon perspective, *American Economic Review* 83, 1346–1955.

Cai, Zongwu, and Yunfei Wang, 2014, Testing predictive regression models with nonstationary regressors, *Journal of Econometrics* 178, 4–14.

Campbell, John Y., and Robert. J Shiller, 1988, Stock prices, earnings , and expected dividends, *Journal of Finance* 43, 661–676.

Campbell, John Y., and Motohiro Yogo, 2006, Efficient tests of stock return predictability, *Journal of Financial Economics* 81, 27–60.

Cavanagh, Christopher L., Graham Elliott, and James H. Stock, 2009, Inference in models with nearly integrated regressors, *Econometric Theory* 11, 1131–1147.

Chan, Ngai Hang, Deyuan Li, and Liang Peng, 2012, Toward a unified interval estimation of autoregressions, *Econometric Theory* 28, 705–717.

Chen, Willa, W., and S. Deo, Rohit, 2009, A note on the stationarity and the existence of moments of the garch model, *Econometric Theory* 25, 1143–1179.

Demetrescu, Matei, Iliyan Georgiev, M. M. Rodrigues, Paulo, and A. M. Taylor, Robert, 2023, Extensions to ivx methods of inference for return predictability, *Journal of Econometrics, forthcoming.*

Demetrescu, Matei, and Paulo M. M. Rodrigues, 2022, Residual-augmented IVX predictive regression, *Journal of Econometrics* 227, 429–460.

——— , and Robert A.M. Taylor, 2023, Transformed regression-based long-horizon predictability tests, *Journal of Econometrics, forthcoming.*

Fama, Eugene F., and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Hall, Peter, and Barbara La Scala, 1990, Methodology and algorithms of empirical likelihood, *International Statistical Review* 58, 109–127.

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

Kostakis, Alexandros, Tassos Magdalinos, and Michalis P. Stamatogiannis, 2015, Robust econometric inference for stock return predictability, *Review of Financial Studies* 28, 1506–1553.

———, 2023, Taking stock of long-horizon predictability tests: Are factor returns predictable?, *Journal of Econometrics*.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238–1287.

Li, Chenxue, Deyuan Li, and Liang Peng, 2017, Uniform test for predictive regression with ar errors, *Journal of Business & Economic Statistics* 35, 29–39.

Liu, Xiaohui, Bingduo Yang, Zongwu Cai, and Liang Peng, 2019, A unified test for predictability of asset returns regardless of properties of predicting variables, *Journal of Econometrics* 208, 141–159.

Magdalinos, Tassos, and Peter C. B. Phillips, 2009, Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors, *Econometric Theory* 25, 482–526.

Mishkin, Frederic S., 1990, What does the term structure tell us about future inflation?, *Journal of Monetary Economics* 25, 77–95.

Owen, Art. B, 1988, Empirical likelihood ratio confidence intervals for a single functional, *Biometrika* 75, 237–249.

———, 1990, Empirical likelihood ratio confidence regions, *Annals of Statistics* 18, 90–120.

Pesaran, M. Hashem, and Allan Timmermann, 1995, Predictability of stock returns: robustness and economic significance, *Journal of Finance* 50, 1201–1228.

Phillips, Peter C. B., 2015, Halbert white jr. memorial jfec lecture: Pitfalls and possibilities in predictive regression, *Journal of Financial Econometrics* 13, 521–555.

———, and Ji Hyung Lee, 2013, Predictive regression under various degrees of persistence and robust long-horizon regression, *Journal of Econometrics* 177, 250–264.

———, 2016, Robust econometric inference with mixed integrated and mildly explosive regressors, *Journal of Econometrics* 192, 433–450.

Stambaugh, Robert F., 1999, Predictive regression, *Journal of Financial Economics* 54, 375–421.

Volkanov, Rossen, 2003, Long-horizon regressions: theoretical results and applications, *Journal of Financial Economics* 68, 201–232.

Wei, Min, and Jonathan H. Wright, 2013, Reverse regressions and long-horizon forecasting, *Journal of Applied Econometrics* 28, 353–371.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Yang, Bingduo, Xiaohui Liu, Liang Peng, and Zongwu Cai, 2021, Unified tests for a dynamic predictive regression, *Journal of Business & Economic Statistics* 39, 684–699.

Yang, Bingduo, Wei Long, Liang Peng, and Zongwu Cai, 2020, Testing the predictability of u.s. housing price index returns based on an ivx-ar model, *Journal of the American Statistical Association* 115, 1598–1619.

Zhu, Fukang, Zongwu Cai, and Liang Peng, 2014, Predictive regressions for macroeconomic data, *The Annals of Applied Statistics* 8, 577–594.

Table 1
Finite-sample sizes

The table reports the probability of rejecting a true null hypothesis for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). Rejection rates for each test correspond to a 5% nominal level and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for samples of length $T \in \{250, 500, 1000\}$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi \in \{-0.95, -0.5, 0\}$, and localizing constant $c \in \{-50, -20, -10, 0, 1\}$.

ϕ	c	h	$T = 250$				$T = 500$				$T = 1000$			
			EL1	EL2	KMS	PL	EL1	EL2	KMS	PL	EL1	EL2	KMS	PL
-0.95	-50	1	0.05	0.05	0.04	0.04	0.05	0.05	0.04	0.04	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.04	0.04	0.05	0.05	0.04	0.04	0.05	0.05	0.05	0.05
		12	0.05	0.06	0.03	0.03	0.05	0.05	0.04	0.04	0.05	0.05	0.04	0.04
		60	0.05	0.05	0.02	0.02	0.05	0.05	0.03	0.03	0.05	0.05	0.04	0.04
	-20	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.05	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.05	0.03	0.03	0.05	0.05	0.03	0.04	0.05	0.05	0.05	0.05
	-10	1	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06
		12	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.05
		60	0.05	0.04	0.03	0.04	0.05	0.05	0.04	0.05	0.05	0.04	0.05	0.05
	0	1	0.05	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06
		4	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06
		12	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.06	0.02	0.03	0.05	0.05	0.03	0.04	0.05	0.05	0.04	0.05
	1	1	0.05	0.07	0.05	0.05	0.05	0.07	0.05	0.05	0.05	0.07	0.05	0.05
		4	0.05	0.06	0.05	0.05	0.05	0.07	0.05	0.05	0.05	0.07	0.05	0.05
		12	0.05	0.06	0.04	0.05	0.05	0.07	0.05	0.05	0.05	0.07	0.05	0.05
		60	0.05	0.06	0.02	0.03	0.05	0.06	0.03	0.04	0.05	0.06	0.04	0.04
-0.50	-50	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.05	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.06	0.03	0.03	0.05	0.05	0.04	0.04	0.05	0.05	0.04	0.05
	-20	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.05	0.04	0.03	0.05	0.05	0.04	0.04	0.05	0.05	0.05	0.05
	-10	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.05	0.03	0.04	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05
	0	1	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.06	0.05	0.05	0.06	0.06
		4	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06
		12	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.06	0.05	0.05	0.05	0.06
		60	0.05	0.05	0.03	0.03	0.05	0.05	0.04	0.04	0.05	0.05	0.05	0.05
	1	1	0.05	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.06	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05
		60	0.06	0.06	0.03	0.03	0.05	0.05	0.03	0.04	0.05	0.05	0.04	0.04

Table 1 – *Continued*

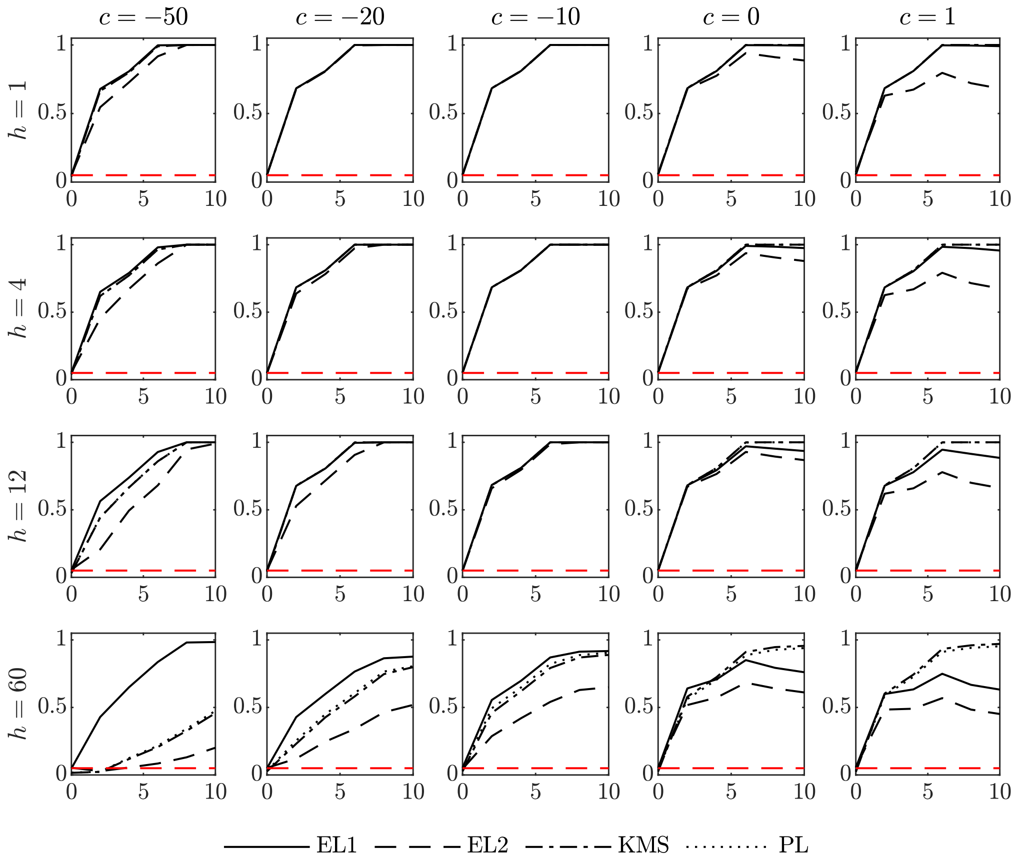
ϕ	c	h	$T = 250$				$T = 500$				$T = 1000$			
			EL1	EL2	KMS	PL	EL1	EL2	KMS	PL	EL1	EL2	KMS	PL
0.00	-50	1	0.05	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.05
		60	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.05
	-20	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	-10	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.06	0.04	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05
	0	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.05	0.04	0.04	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05
	1	1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		4	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		12	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		60	0.05	0.05	0.04	0.04	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05

Table 2
Empirical Results

The table reports p-values from the univariate empirical likelihood test of predictability at horizons $h = \{1, 4, 12, 60\}$. The dependent variable is the continuously compounded return of the CRSP value weighted index in excess of the one-month Treasury bill rate. The sample period is January 1952 - December 2022.

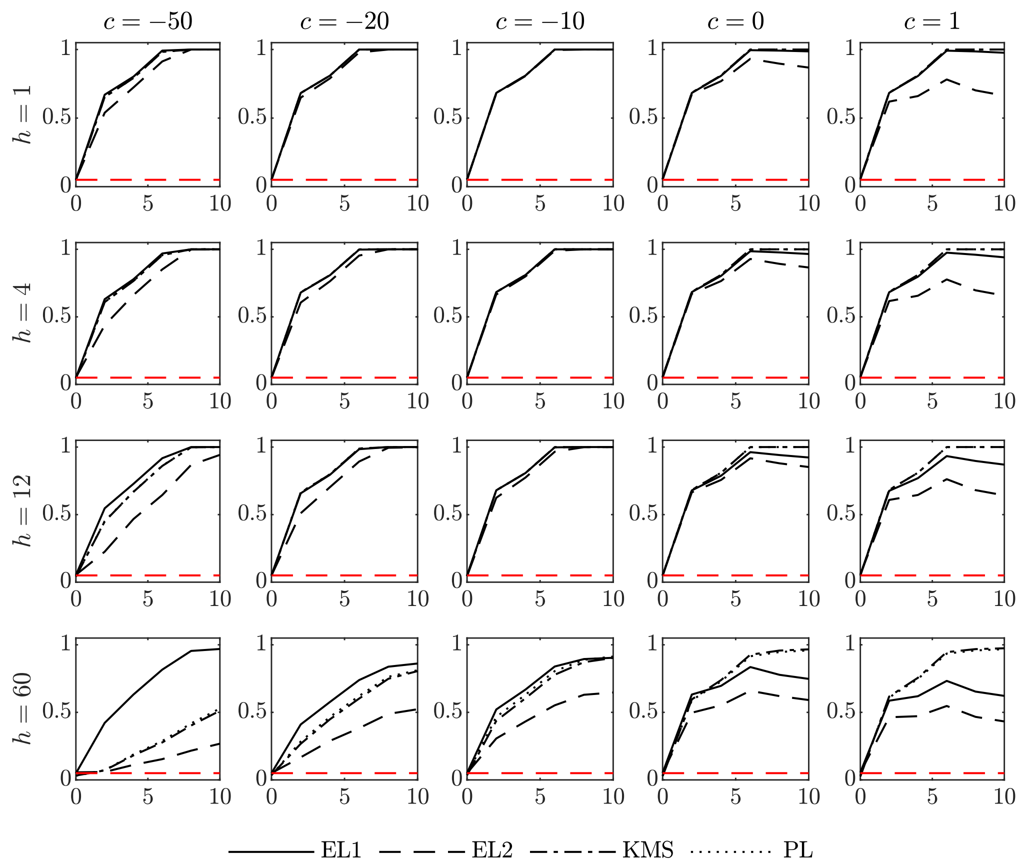
Variable	$h = 1$	$h = 4$	$h = 12$	$h = 60$
Dividend price ratio	0.70	0.91	0.91	0.53
Dividend yield	0.62	0.85	0.89	0.53
Earnings price ratio	0.98	0.80	0.60	0.53
Dividend payout ratio	0.20	0.25	0.48	0.60
Book-to-market ratio	0.93	0.83	0.79	0.53
Net equity expansion	0.61	0.45	0.23	0.37
Treasury bills	0.03	0.07	0.13	0.35
Long term yield	0.08	0.19	0.29	0.61
The term spread	0.21	0.13	0.08	0.08
Default yield spread	0.54	0.34	0.16	0.30
Inflation	0.05	0.04	0.01	0.14

Figure 1
Finite-sample power plots for a sample of length $T = 250$ and degree of endogeneity $\phi = -0.95$



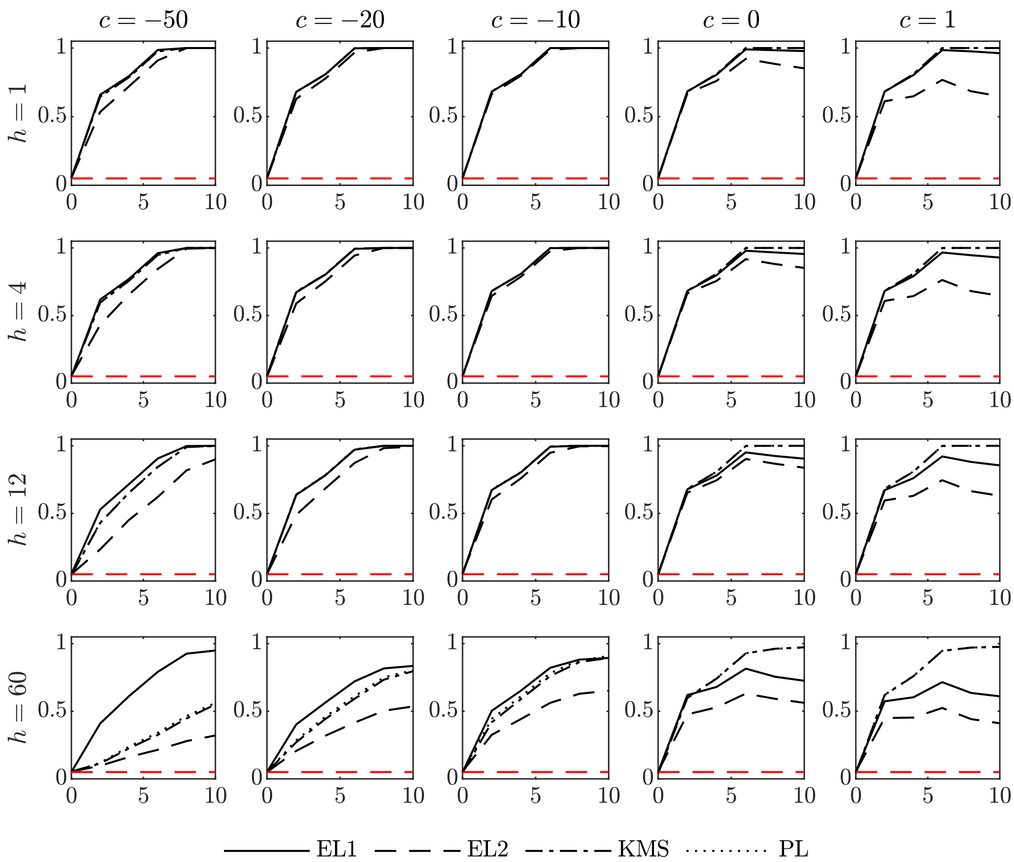
The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 250$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = -0.95$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 2
Finite-sample power plots for a sample of length $T = 250$ and degree of endogeneity $\phi = -0.50$



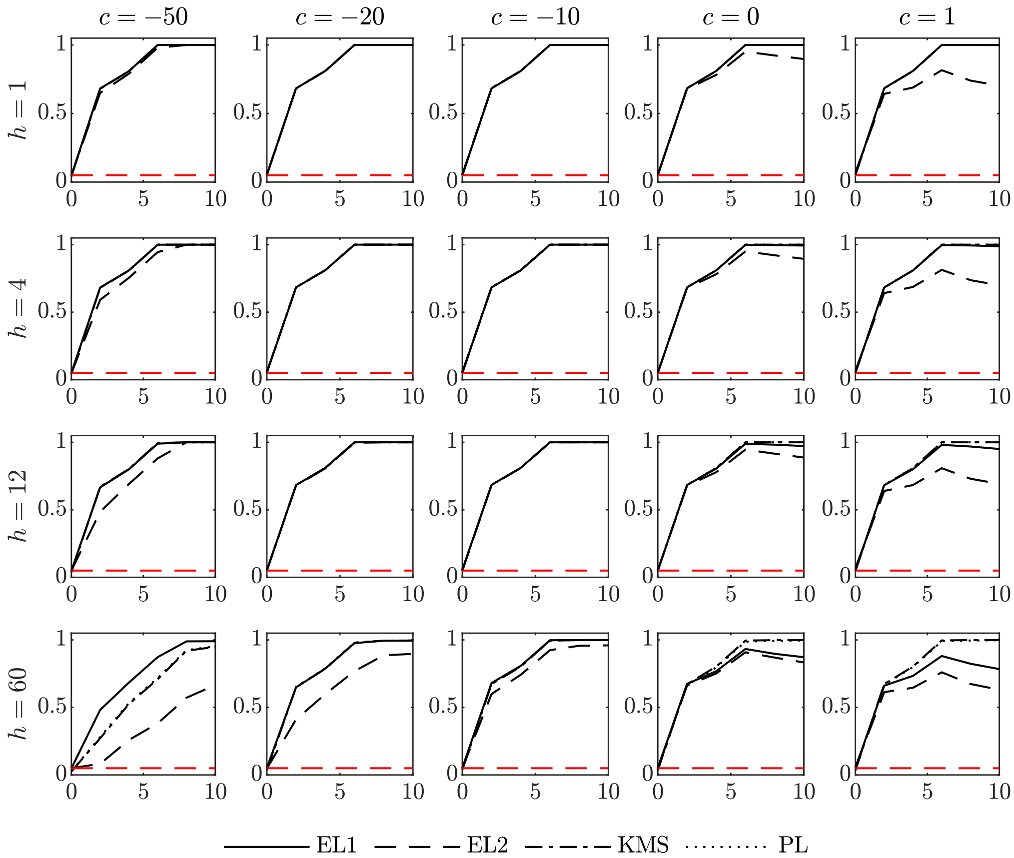
The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 250$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = -0.50$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 3
Finite-sample power plots for a sample of length $T = 250$ and degree of endogeneity $\phi = 0$



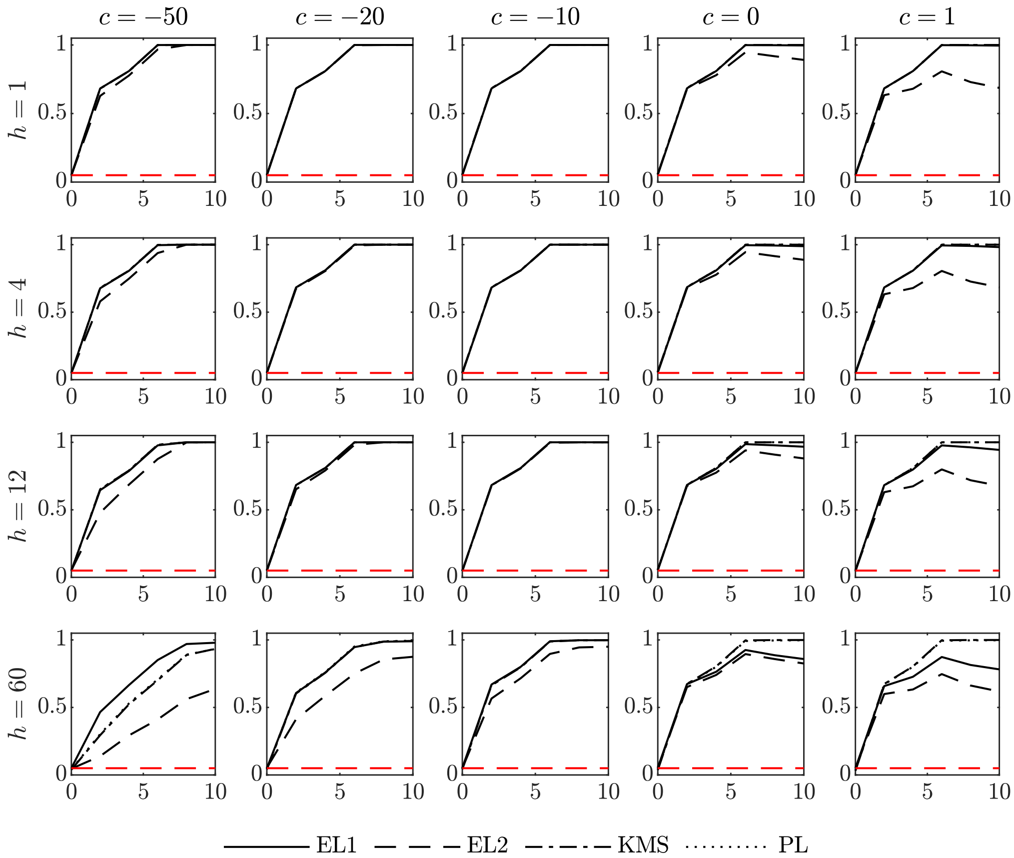
The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 250$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = 0$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 4
Finite-sample power plots for a sample of length $T = 500$ and degree of correlation $\phi = -0.95$



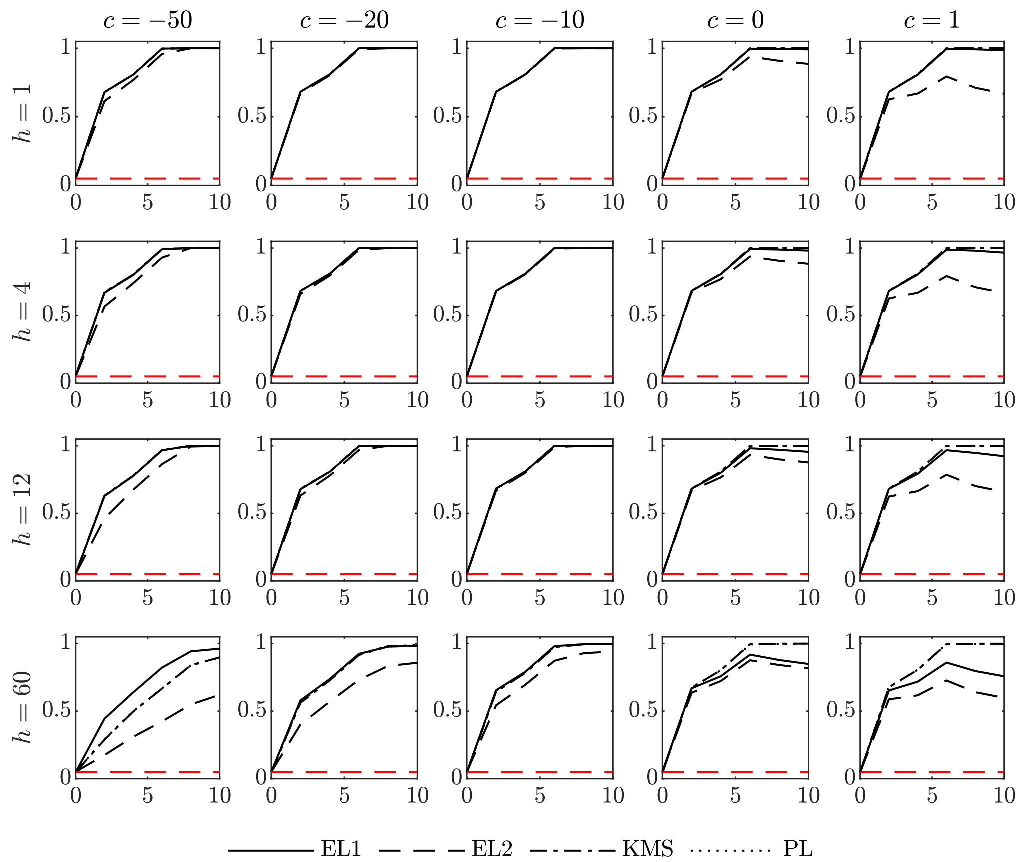
The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 500$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = -0.95$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 5
Finite-sample power plots for a sample of length $T = 500$ and degree of correlation $\phi = -0.50$



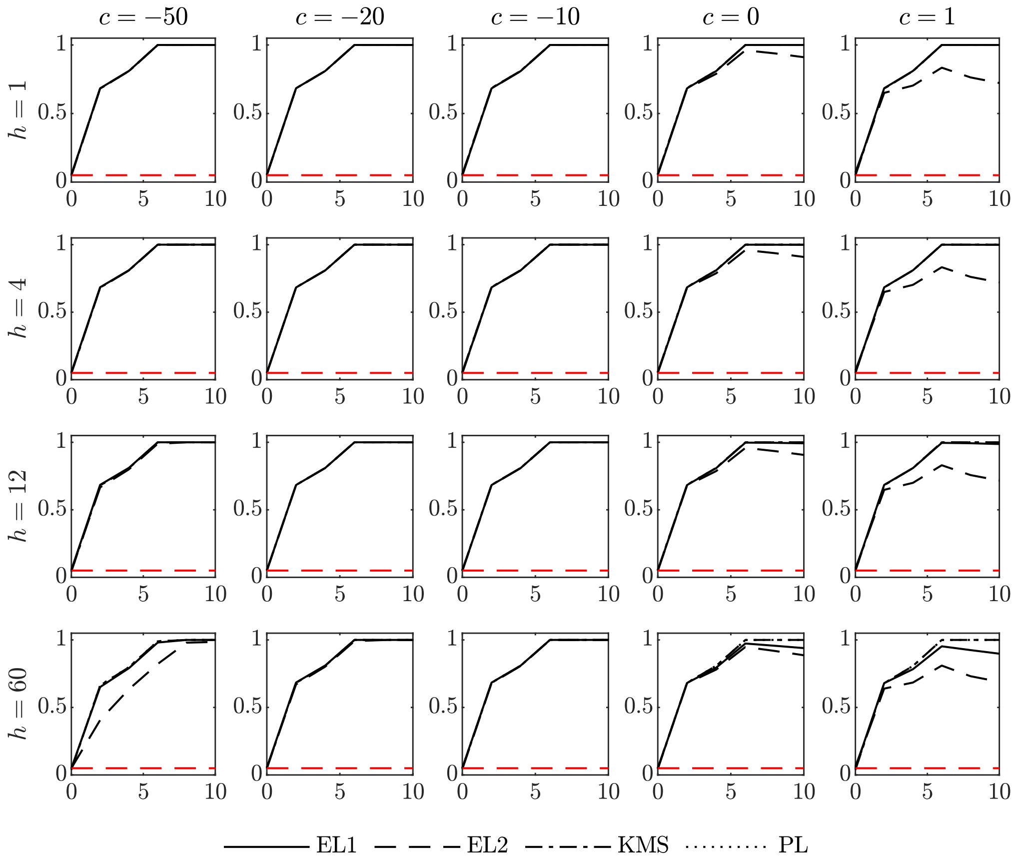
The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 500$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = -0.50$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 6
Finite-sample power plots for a sample of length $T = 500$ and degree of correlation $\phi = 0$



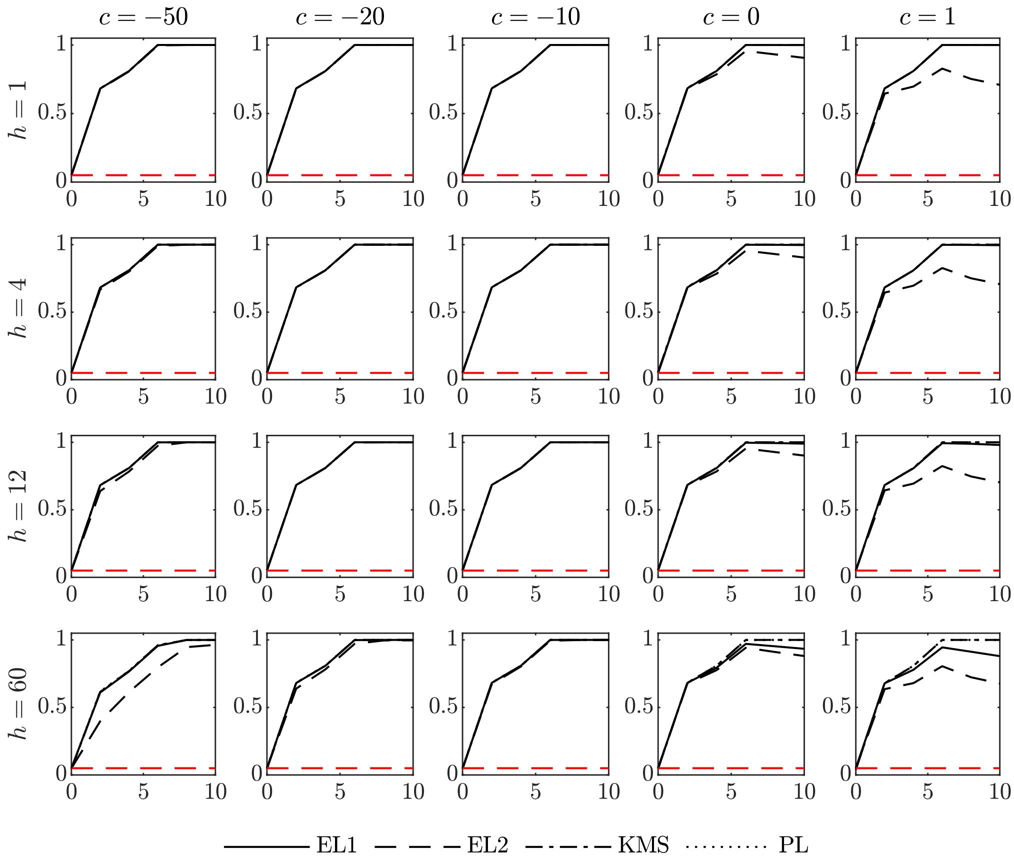
The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 500$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = 0$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 7
Finite-sample power plots for a sample of length $T = 1000$ and degree of correlation $\phi = -0.95$



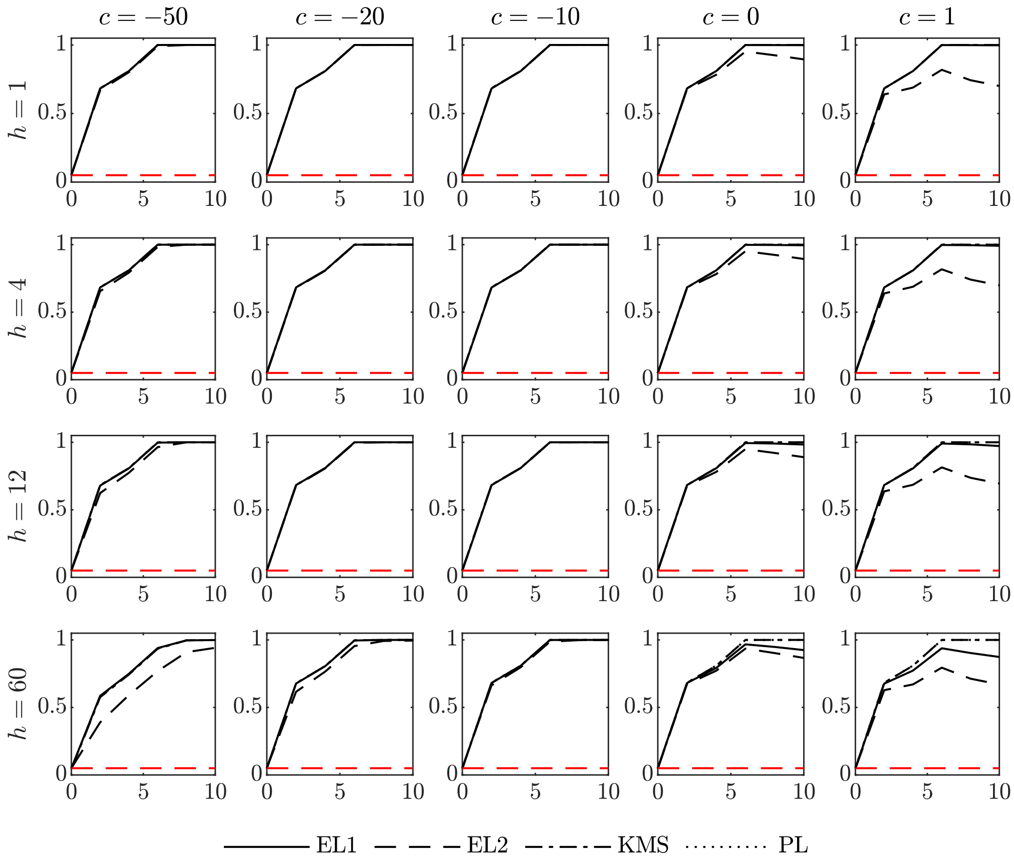
The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 1000$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = -0.95$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 8
Finite-sample power plots for a sample of length $T = 1000$ and degree of correlation $\phi = -0.50$



The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 1000$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = -0.50$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Figure 9
Finite-sample power plots for a sample of length $T = 1000$ and degree of correlation $\phi = 0$



The figure summarizes the probability of rejecting a false null hypothesis (y-axis) for the empirical likelihood test of long-horizon predictability under the scenario that the intercept in the predictive regression model (1) is either known (EL1) or unknown (EL2). As a base for comparison, the table also reports rejection probabilities for the tests of long-horizon predictability proposed by Phillips and Lee (2013) (PL) and Kostakis, Magdalinos, and Stamatogiannis (2023) (KMS). The predictive regression coefficient is set equal to $b/\sqrt{T}$, where the localizing coefficient b (x-axis) takes values that increase from 0 to 10 in increments of 2. $b = 0$ corresponds to the size of the test. Rejection rates for each test correspond to a 5% nominal level (the red dashed line) and are based on 10,000 Monte Carlo simulations (see section 3.1 for detailed description of the simulation design). Results are reported for a sample of length $T = 1000$, forecast horizons $h \in \{1, 4, 12, 60\}$, degree of correlation between the innovations of models (1) and (2) $\phi = 0$, and a localizing constant for the autoregressive coefficient $c \in \{-50, -20, -10, 0, 1\}$.

Appendix for Hong and Tsvetanov (2023)
Proofs of the main results: Version 6 September 2023

Proof of Theorem 1. For the sake of simplicity, we prove under the assumption that $\phi = 0$ and $b_k = 0$ for all $k = 1, \ldots, p-1$ in equations (2) and (3), because the same argument used in e.g. Zhu, Cai, and Peng (2014) in the early part of the proof of Theorem 2 can be employed to deal with the moving average structure in the error terms.

In deriving the asymptotic distribution via the martingale central limit theorem, the first step concerns with the derivation of the limiting variance, that is, the probability limit of the martingale conditional variance, Lévy (1937), Billingsley (1961).

On noting conditional homogeneity and the moment condition of the process $\widetilde{U}_t$, we see that, with respect to the natural filtration $\mathcal{F}_{Tt} = \sigma(X_0, U_0, U_1, \ldots, U_t)$, the conditional variance of the martingale array Z_{Tt} is given by

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\big(Z_{Tt}^2 | \mathcal{F}_{T,t+h-1}\big) &= \frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\left(\left[\frac{(\widetilde{Y}_{t+h} - \beta_h^{\mathrm{rev}}\widetilde{X}_t(h))\widetilde{X}_t(h)}{\sqrt{1+\widetilde{X}_t(h)^2}}\right]^2 \middle| \mathcal{F}_{T,t+h-1}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\left(\frac{\widetilde{U}_{t+h}^2\widetilde{X}_t(h)^2}{1+\widetilde{X}_t(h)^2}\middle| \mathcal{F}_{T,t+h-1}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\left(\frac{\widetilde{U}_{t+h}^2(X_t + X_{t+1} + \cdots + X_{t+h-1})^2}{1+(X_t + X_{t+1} + \cdots + X_{t+h-1})^2}\middle| \mathcal{F}_{T,t+h-1}\right) \\
&= \mathrm{E}(\widetilde{U}_{t+h}^2 | \mathcal{F}_{T,t+h-1}) \cdot \frac{1}{T}\sum_{t=1}^{T} \frac{\widetilde{X}_t(h)^2}{1+\widetilde{X}_t(h)^2} \\
&= \sigma_u^2 \cdot \left\{\frac{1}{T}\sum_{t=1}^{T} \frac{\widetilde{X}_t(h)^2}{1+\widetilde{X}_t(h)^2}\right\} \;=:\; \eta^2. \qquad (14)
\end{aligned}$$

We note that in the mildly integrated case, i.e., $\rho = \rho_T = 1 + c/T^a$ with $0 < a < 1$ and $c < 0$, quoting an earlier version of Phillips and Magdalinos (2007), Phillips and Magdalinos (2005, Theorem 2.1) showed that

$$T^{-a/2} X_{\lfloor T^a t \rfloor} \implies \int_0^t e^{c(t-r)} dW(r), \qquad (15)$$

where W is Brownian motion with variance $\sigma^2 = \mathrm{E}(\varepsilon_t^2)$ and $\Longrightarrow$ refers to weak convergence in the Skorohod space $\mathcal{D}[0, \ell]$ (i.e. the space of the collection of $\mathbb{R}$-valued càdlàg functions on $[0, 1]$), see e.g. Pollard (1984). The initial condition $X_0 = o_p(T^{a/2})$ is imposed, and a finite moment strictly higher than 2 is required for the i.i.d. error term, which are consistent with what we assume. Giraitis and Phillips (2004) has a relevant result when the error is a martingale difference.

Further, in the mildly explosive case, i.e., $\rho = \rho_T = 1 + c/T^a$ with $0 < a < 1$ and $c > 0$, the proof of Aue and Horváth (2007) suggests that for any fixed constant $\ell > 0$ we have

$$\begin{aligned}\frac{1}{\xi_T^{-1/2}(\mathrm{E}(\varepsilon_1^2))^{1/2}} &\sum_{t=1}^{[\ell/\xi_T]} \rho^{-t}\varepsilon_t \\ &= \frac{1}{\xi_T^{-1/2}(\mathrm{E}(\varepsilon_1^2))^{1/2}}\rho^{-[\ell/\xi_T]}X_{[\ell/\xi_T]} \implies e^{-\ell}\mathcal{W}_{\alpha,\beta}(\ell) + \int_0^\ell \mathcal{W}_{\alpha,\beta}(x)dx, \end{aligned} \tag{16}$$

where $\xi_T = \log\rho = \log\rho_T = \log(1 + c/T^a) \to 0$ as $T \to \infty$, and $\mathcal{W}_{\alpha,\beta}$ is a strictly α-stable random variable, Petrov (1975).

In the near integrated case where $a = 1$ and $c \neq 0$, we know from Phillips (1987) that (see also Liu, Yang, Cai and Peng (2019))

$$\frac{1}{\sqrt{T}}X_{[Tr]} \implies \int_0^r e^{-(r-s)\rho}dW(s). \tag{17}$$

as $T \to \infty$.

In all three cases, the denominators of the "multiplier" to X all tend to infinity in LHS of (15), (16) and (17). Therefore, via Skorokhod representation theorem and Lebesgue's dominated convergence theorem we can readily see that for each h,

$$\frac{\widetilde{X}_t(h)^2}{1 + \widetilde{X}_t(h)^2} = \frac{(\sum_{j=0}^{h-1} X_{t+j})^2}{1 + (\sum_{j=0}^{h-1} X_{t+j})^2} \xrightarrow{L_1} 1 \tag{18}$$

as $t \to \infty$. An alternative way to see this is to notice that we have $|X_t| \xrightarrow{p} +\infty$[5], from which we can easily show (18) holds.

Consequently, the stochastic convergence of Cesàro means of random variables, see e.g. Schilling (2017), Bibaut, Luedtke, van der Laan (2020), yields

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\widetilde{X}_t(h)^2}{1 + \widetilde{X}_t(h)^2} \xrightarrow{L_1} 1 \tag{19}$$

as $T \to \infty$, which implies convergence in probability to 1.

Therefore, we finally have

$$\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\left(\left[\frac{(\widetilde{Y}_{t+h} - \beta_h^{\mathrm{rev}}\widetilde{X}_t(h))\widetilde{X}_t(h)}{\sqrt{1 + \widetilde{X}_t(h)^2}}\right]^2 \middle| \mathcal{F}_{T,t+h-1}\right) \xrightarrow{p} 2\mathrm{E}(U_t^2). \tag{20}$$

[5]which is taken to mean $P(X_t > r) \to 1$ for every $r > 0$, see for example Kallenberg (1997).

Meanwhile, in the stationary case, we have, as $T \to \infty$

$$\begin{aligned}
\frac{\widetilde{X}_t(h)^2}{1+\widetilde{X}_t(h)^2} &= \frac{(\sum_{j=0}^{h-1} X_{t+j})^2}{1+(\sum_{j=0}^{h-1} X_{t+j})^2} \\
&= \frac{\left(\sum_{j=0}^{h-1}\sum_{i=1}^{t} \rho^{t+j-i}\varepsilon_i\right)^2}{1+\left(\sum_{j=0}^{h-1}\sum_{i=1}^{t} \rho^{t+j-i}\varepsilon_i\right)^2} \\
&= \lim_{t\to\infty} \mathrm{E}\left\{ \frac{\left(\sum_{i=1}^{t}\sum_{j=0}^{h-1} \rho^{t+j-i}\varepsilon_i\right)^2}{1+\left(\sum_{i=1}^{t}\sum_{j=0}^{h-1} \rho^{t+j-i}\varepsilon_i\right)^2} \right\} \;+\; o_p(1) \\
&=: \theta_0^2 \;+\; o_p(1) \qquad (21)
\end{aligned}$$

because the series converges absolutely almost surely, which can be easily checked using standard arguments, e.g. van de Vaart (2013).

Following the same aforementioned argument that led to (19), we can immediately see that

$$\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\left(\left[\frac{(\widetilde{Y}_{t+h}-\beta_h^{\mathrm{rev}}\widetilde{X}_t(h))\widetilde{X}_t(h)}{\sqrt{1+\widetilde{X}_t(h)^2}}\right]^2 \middle| \mathcal{F}_{T,t+h-1}\right) \xrightarrow{p} 2\mathrm{E}(U_t^2)\cdot\theta_0^2. \qquad (22)$$

as $T \to \infty$.

It now remains to check the set of requirements for establishing the martingale central limit theorem, e.g. Hall and Heyde (1980, Corollary 3.1). Specifically, we check the uniform negligibility condition, see for example Lai and Wei (1982).

With $q > 0$ in the moment condition for $\widetilde{U}_t$, we see that for all $\epsilon > 0$,

$$\begin{aligned}
&\max_{1\leqslant j\leqslant h}\sum_{t=1}^{T}\mathrm{E}\left(\frac{1}{T}\frac{\widetilde{U}_{t+j}^2\widetilde{X}_t(j)^2}{1+\widetilde{X}_t(j)^2}\cdot 1\left\{\frac{1}{T}\frac{\widetilde{U}_{t+j}^2\widetilde{X}_t(j)^2}{1+\widetilde{X}_t(j)^2} > \epsilon\right\}\middle|\mathcal{F}_{T,t+j-1}\right) \\
&\leqslant \max_{1\leqslant j\leqslant h}\sum_{t=1}^{T}\frac{1}{\epsilon^q}\mathrm{E}\left(\left|\frac{1}{\sqrt{T}}\frac{\widetilde{U}_{t+j}\widetilde{X}_t(j)}{\sqrt{1+\widetilde{X}_t(j)^2}}\right|^{2+q}\middle|\mathcal{F}_{T,t+j-1}\right) \\
&= \frac{1}{\epsilon^q T^{1+q/2}}\max_{1\leqslant j\leqslant h}\sum_{t=1}^{T}\mathrm{E}\left(\left|\frac{\widetilde{U}_{t+j}^{2+q}\widetilde{X}_t(j)^{2+q}}{\left(1+\widetilde{X}_t(j)^2\right)^{(2+q)/2}}\right|\middle|\mathcal{F}_{T,t+j-1}\right) \\
&= \frac{\max_{1\leqslant j\leqslant h}\mathrm{E}\left|\widetilde{U}_{t+j}\right|^{2+q}}{\epsilon^q T^{q/2}}\cdot\left\{\max_{1\leqslant j\leqslant h}\frac{1}{T}\sum_{t=1}^{T}\frac{\left|\widetilde{X}_t(j)\right|^{2+q}}{\left(1+\widetilde{X}_t(j)^2\right)^{(2+q)/2}}\right\} \\
&= O(1)\cdot\frac{1}{T^{q/2}}\cdot O_p(h) \qquad (23)
\end{aligned}$$

which converges in probability to 0 in large samples provided $h/T^{q/2} \to 0$.

Now that the conditional Lindeberg condition is met, we see that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{(\widetilde{Y}_{t+h}-\beta_h^{\mathrm{rev}}\widetilde{X}_t(h))\widetilde{X}_t(h)}{\sqrt{1+\widetilde{X}_t(h)^2}} \;=\; \frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z_{Tt}(\beta_0) \implies N\big(0,\eta^2\big), \tag{24}$$

where η^2 is as in (14).

Consequently, the desired result follows from the standard arguments in proving the empirical likelihood method, Owen (1990), Qin and Lawless (1994), Zhu et al. (2014). $\square$