Some curious power properties of long-horizon tests

Erik Hjalmarsson

Queen Mary, University of London, School of Economics and Finance, Mile End Road, London E1 4NS, UK

ABSTRACT

Based on simulations and asymptotic results, I highlight three distinct properties of long-horizon predictive tests. (i) The asymptotic power of long-horizon tests increases only with the sample size relative to the forecasting horizon. Keeping this ratio fixed as the sample size increases does not lead to any power gains asymptotically. (ii) Although the power of long-horizon tests increases with the magnitude of the slope coefficient for alternatives close to the null hypothesis, there are no gains in power as the slope coefficient grows large. That is, the power curve is asymptotically horizontal when viewed as a function of the slope coefficient. (iii) For endogenous regressors—i.e., when the innovations to the regressand are contemporaneously correlated with the innovations to the regressor—traditional tests based on the standard long-run OLS estimator result in power curves that are sometimes decreasing in the magnitude of the slope coefficient.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Long-horizon tests play an important part in empirical finance and economics. In financial economics, it has been very popular to test whether asset returns in general, and stock returns in particular, can be predicted over long horizons with some forecasting variable. The purpose of this study is to shed some further light on the power properties of long-horizon tests, with the focus on documenting and understanding the key features of the power properties of long-run tests when the true data generating process satisfies a simple benchmark model given by the standard linear predictive relationship. In contrast to much previous work, the main goal is not to deduce whether one can achieve greater power at longer horizons relative to shorter horizons (e.g., Campbell, 2001; Rapach and Wohar, 2005), but rather to try to glean some general understanding of the power properties of long-horizon tests.
The study is based on simulations as well as asymptotic results and reveals three interesting facts about the properties of long-run predictive tests. First, the asymptotic power of long-run tests increases only with the sample size relative to the forecasting horizon. Keeping this ratio fixed as the sample size increases does not lead to any power gains asymptotically. Second, although the power of the long-horizon tests increases with the magnitude of the slope coefficient for alternatives close to the null hypothesis, there are no gains in power as the slope coefficient grows large. That is, the power curve is asymptotically horizontal when viewed as a function of the slope coefficient. Third, when the regressors are endogenous, tests that are based on the standard long-run OLS estimator will result in power curves that are sometimes decreasing in the magnitude of the slope coefficient. That is, as the model drifts farther away from the null hypothesis, the power may decrease. This is true both for Valkanov’s (2003) test, but also if one uses, for instance, Newey–West standard errors in a normal $t$-statistic. The test proposed by Hjalmarsson (2011) specifically for the case of endogenous regressors does not suffer from this problem. These findings add an extra note of caution to the use of long-horizon regressions. This is especially true for relatively large forecasting horizons, where the second and third effects mentioned above are most prevalent.

Long-horizon regressions are best interpreted as fitted regressions and not data generating processes, and the power properties are therefore conditional on the specification of the model under the alternative of predictability. In the final section of the paper, I return to a discussion of how sensitive the results are to the exact nature of the data generating process.

# 2. Long-run inference

## 2.1. Model and assumptions

Let the dependent variable be denoted $r_t$, which would typically represent excess stock returns when analyzing stock return predictability, and the corresponding regressor, $x_t$. The long-run fitted regression is given by,

$$ r_{t+q}(q) = \hat{\alpha}_q + \hat{\beta}_q x_t + \hat{u}_{t+q}(q), $$

where $r_{t}(q) = \sum_{j=1}^{q} r_{t-j} x_{t-j}$, and long-run future returns are regressed onto a one-period predictor. $\hat{\beta}_q$ represents the OLS estimator of the slope coefficient in the regression of $r_{t+q}(q)$ onto $x_t$, using overlapping observations, $\hat{\alpha}_q$ is the corresponding estimator of the intercept, and $\hat{u}_{t+q}(q)$ are the fitted residuals. The standard $t$-statistic corresponding to $\hat{\beta}_q$ is denoted by $t_q$.

Eq. (1) is a fitted regression and should not be interpreted as the data generating process (dgp). The dgp more naturally describes the true model for $r_t$, rather than $r_{t}(q)$, since the latter is just a sum of the former. I follow previous work by, for instance, Campbell (2001), and assume that the true dgp for $r_t$ is given by

$$ r_{t+1} = \alpha + \beta x_t + u_{t+1}. $$

Further, it is assumed that the regressor follows an AR(1) process,

$$ x_{t+1} = \gamma + \rho x_t + v_{t+1}, $$

where $\rho = 1 + c/T,$ $t = 1, \ldots, T,$ and $T$ is the sample size. The autoregressive root of the regressor is parameterized as being local-to-unity, which captures the near unit-root, or highly persistent, behavior of many predictor variables, but is less restrictive than a pure unit-root assumption.\(^{2}\)

---

1 A predictive regressor is referred to as endogenous if the innovations to the dependent variable are contemporaneously correlated with the innovations to the regressor. For persistent regressors, this gives rise to the Stambaugh (1999) bias.

2 The near unit-root construction, where the autoregressive root drifts closer to unity as the sample size increases, is used as a tool to enable an asymptotic analysis where the persistence in the data remains large relative to the sample size. That is, if $\rho$ is treated as fixed and strictly less than unity, then as the sample size grows, the process $x_t$ will behave as a strictly stationary process asymptotically, and the standard first order asymptotic results will not provide a good guide to the actual small sample properties of the model. For $\rho = 1,$ the usual unit-root asymptotics apply to the model, but this is clearly a restrictive assumption for most potential predictor variables. Instead, by letting $\rho = 1 + c/T,$ the effects from the high persistence in the regressor will appear also in the asymptotic results, but without imposing the strict assumption of a unit root.
To complete the model specification, the joint error process, \( w_t = (u_t, v_t)' \), is assumed to satisfy a martingale difference sequence (mds) with finite fourth order moments. That is, let \( F_t = \{ w_s \mid s \leq t \} \) be the filtration generated by \( w_t \). Then \( E[w_t \mid F_{t-1}] = 0 \), \( E[w_t w_s'] = \Omega = [(\omega_{11}, \omega_{12}), (\omega_{12}, \omega_{22})] \), \( \sup_t E[u_t^4] < \infty \), \( \sup_t E[v_t^4] < \infty \), and \( E[x_t^2] < \infty \). In typical stock-return predictability regressions, the error terms \( u_t \) and \( v_t \) are often highly correlated and the regressor will be referred to as endogenous whenever this correlation, which will be labelled \( \delta = \omega_{12}/\sqrt{\omega_{11}\omega_{22}} \), is non-zero.

### 2.2. Test procedures

In order to evaluate the power properties of long-horizon tests, I rely on recent test statistics developed by Valkanov (2003) and Hjalmarsson (2011), which have been shown to control size well. Use of overlapping observations induces strong serial correlation in the regression residuals and standard errors that fail to account for this fact lead to biased inference. The traditional solution to this problem has been to use auto-correlation robust standard errors, such as those proposed by Newey and West (1987). However, the strong serial correlation tends to result in poor finite sample performance for such robust estimators. Valkanov (2003) and Hjalmarsson (2011) both overcome these issues by explicitly deriving the asymptotic properties of the long-run estimators and showing that a simple scaling of the standard \( t \)-statistic will result in a correctly sized test. Both of these studies use a framework where the predictor variable follows an AR(1) process with a root that is local to unity.

Valkanov (2003) derives his asymptotic results under the assumption that \( q/T \rightarrow \lambda \in (0, 1) \) as \( q, T \rightarrow \infty \), and shows that under this assumption, \( t_q/\sqrt{T} \) will have a well-defined distribution; again, \( t_q \) is the standard \( t \)-statistic corresponding to the OLS estimator \( \hat{\beta}_q \). The scaled \( t \)-statistic in Valkanov’s analysis is not normally distributed. Its asymptotic distribution is a function of the parameters \( \lambda \) (the degree of overlap), the local-to-unity parameter \( c \), and the degree of endogeneity \( \delta \); critical values must be obtained by simulation for a given combination of these three parameters. Since the critical values are a function of \( c \), Valkanov’s scaled \( t \)-test is generally infeasible since this parameter is unknown and not consistently estimable.\(^3\) He therefore proposes a so-called sup-bound test, where the test is evaluated at some bound for \( c \), outside of which it is assumed that \( c \) will not lie. Ruling out explosive processes, he suggests using \( c = 0 \) in the sup-bound test, which results in a conservative one-sided test against a positive alternative when \( \delta < 0 \).

Hjalmarsson (2011) derives results under the assumption that \( q \) is fixed. Under this assumption, he shows that \( t_q/\sqrt{q} \) will be normally distributed when the regressor is exogenous (\( \delta = 0 \)). For endogenous regressors (\( \delta \neq 0 \)), Hjalmarsson proposes a modified test-statistic based on an augmented regression equation along the lines of Phillips (1991). In particular, Hjalmarsson shows that for a known value of \( c \), the standard \( t \)-statistic from the long-run augmented regression equation (labeled \( t_q^+ \)) will again satisfy a scaling result such that \( t_q^+ / \sqrt{q} \) is asymptotically normally distributed. Since \( c \) is generally unknown, Hjalmarsson recommends a Bonferroni method to make the modified test statistic feasible, along similar lines to the short-run Bonferroni test proposed by Campbell and Yogo (2006). In effect, the Bonferroni method uses a confidence interval for \( c \) and finds the most conservative value of the test statistic for all values of \( c \) in that confidence interval.

In summary, I primarily rely on three different tests in the simulations below: (i) Valkanov’s sup-bound test, (ii) Hjalmarsson’s scaled OLS \( t \)-test, and (iii) Hjalmarsson’s scaled Bonferroni test. To make sure that none of the results are driven by the scaling approaches of Valkanov (2003) and Hjalmarsson (2011), I also briefly report results for (non-scaled) \( t \)-statistics based on Newey and West (1987) standard errors. Both Valkanov’s sup-bound test and Hjalmarsson’s scaled OLS \( t \)-test are based on the standard \( t \)-statistic corresponding to the long-run OLS estimator of Eq. (1). In Hjalmarsson’s OLS \( t \)-test, the scaled test statistic is asymptotically normally distributed as long as the regressor is exogenous (\( \delta = 0 \)), and using critical values based on the standard normal distribution will therefore result in correctly sized tests. When the regressor is endogenous, the test will be biased. In Valkanov’s test, the critical values are obtained through simulations for specific combinations of \( \lambda \), \( \delta \), and \( c \) (using \( c = 0 \) in the feasible sup-bound test). Valkanov’s test is therefore robust to endogenous regressors. However, it is only the critical values that reflect the endogeneity effects, the actual test statistic always remain

\(^3\) That is, \( \rho \) can be estimated consistently, but not with enough precision to identify \( c = T(\rho - 1) \).
the same. This is in contrast to Hjalmarsson's Bonferroni test, where the actual test statistic changes as a reflection of the endogeneity effects, but the critical values are always based on the standard normal distribution. As is seen in the simulation results below, this distinction has an important impact on the power properties of the two tests.\footnote{In robustness checks that are not reported in the paper, qualitatively identical results are found using the infeasible versions of Valkanov's sup-bound test and Hjalmarsson's Bonferroni test. That is, the infeasible tests, which use knowledge of the true value of \(c\), always achieve greater power than the corresponding feasible ones, but the overall shapes of the power curves remain the same. For \(c = 0\) (\(\rho = 1\)), which is one of the cases in the simulations below, Valkanov's sup-bound test coincides with the infeasible test since the sup-bound test uses \(c = 0\) to calculate the conservative critical values.}

### 3. Monte Carlo results

I begin the analysis with a set of simulations that highlight several interesting features of the power properties of long-horizon tests. The following section provides analytical results that help understand the simulation results.

Eqs. (2) and (3) are simulated, with \(u_t\) and \(v_t\) drawn from an iid bivariate normal distribution with mean zero, unit variance and correlations \(\delta = 0\) or \(\delta = -0.9\). In a stock return predictability context, \(\delta = 0\) proxies for nearly exogenous predictors, such as interest rate variables, whereas \(\delta = -0.9\) proxies for highly endogenous predictors, such as valuation ratios (see Campbell and Yogo, 2006). The sample size is either equal to \(T = 100\) or \(T = 500\). The intercept \(\alpha\) is set to zero and the local-to-unity parameter \(c\) is set to either 0 or \(-10\), which corresponds to \(\rho = 1\) and \(\rho = 0.9\) for \(T = 100\) and \(\rho = 1\) and \(\rho = 0.98\) for \(T = 500\). In order to assess the power of the tests, the slope coefficient \(\beta\) in Eq. (2) varies between 0 and 0.5. The power of each test is calculated as the average rejection rate for a 5 percent test against a positive alternative. That is, Hjalmarsson's tests reject when the test statistics are greater than 1.65 (the 95th percentile of the standard normal distribution), whereas the critical values for Valkanov's test depend upon the exact parameter combination and are obtained through simulations for each case. All results are based on 10,000 repetitions.

Fig. 1 shows the power curves for the case with \(q = 10\) and \(T = 100\). The two top panels show results for the scaled OLS \(t\)-test proposed by Hjalmarsson (2011) and the scaled sup-bound test suggested by Valkanov (2003), when the regressor is exogenous (\(\delta = 0\)). The power curves look fairly standard, although power seems to increase very slowly as \(\beta\) grows large. The two bottom panels show the results for endogenous regressors with \(\delta = -0.9\). Since Hjalmarsson's scaled \(t\)-test based on the OLS estimator is known to be biased in this case, I only show the results for Hjalmarsson's scaled Bonferroni test along with Valkanov's sup-bound test. The results are qualitatively similar to those for exogenous regressors with \(\delta = 0\), but there are some signs that the power of Valkanov's sup-bound test decreases for large values of \(\beta\).

Fig. 2 shows the results for \(T = 100\) and \(q = 20\). With exogenous regressors (\(\delta = 0\), top panels), the power curves are increasing in \(\beta\), but only weakly so as \(\beta\) grows larger. For endogenous regressors (\(\delta = -0.9\), bottom panels), the declining pattern in the power curves for Valkanov's sup-bound test that was hinted at in Fig. 1 is now evident; as \(\beta\) becomes larger, the power of the test declines. Note that the turning points of the power curves are not outside the relevant parameter region. For \(c = 0\), the power is already declining for \(\beta = 0.1\); the results in Campbell and Yogo (2006) show that in annual data the estimates of \(\beta\) are between 0.1 and 0.3 for the dividend and earnings-price ratios.\footnote{Valkanov (2003) also performs a Monte Carlo experiment of the power properties of his proposed test-statistics, without finding the sometimes decreasing patterns in the power curves reported here. However, Valkanov (2003) only considers the case with \(\delta = -0.9\), \(q = 10\), and \(T = 100\), for values of \(\beta\) between 0 and 0.1. As seen in Fig. 1 here, the power curves of all the tests are strictly increasing in \(\beta\) for these parameter values.}

Fig. 3 shows results for \(T = 500\) and \(q = 100\), which confirm and elaborate on the findings in Figs. 1 and 2. For endogenous regressors (bottom panels), the same pattern as in Fig. 2 emerges for Valkanov's test: after an initial increase in power as \(\beta\) becomes larger, the power starts to decrease. Further results for larger values of \(\beta\), which are not shown, indicate that the power curves do not converge to zero as \(\beta\) grows large; rather, they seem to level out after the initial decrease. The power curves for Hjalmarsson's scaled Bonferroni test do not seem to converge to one as \(\beta\) increases, although they do not decrease either, and stabilize at a much higher level than the power curves for Valkanov's test. The
power curves for Hjalmarsson’s scaled OLS $t$-test are also shown in the bottom panels. This test is biased for $\delta \neq 0$, but provides an interesting comparison to the power curves of Valkanov’s test. As is seen, the power curves for Hjalmarsson’s scaled OLS $t$-test have the same general shape as those for Valkanov’s sup-bound test. The difference in behavior between Hjalmarsson’s Bonferroni test and Valkanov’s test thus appears to stem from the endogeneity correction and not the manner in which they are scaled. Finally, the bottom panels in Fig. 3 also show that the patterns established for the power curves of the tests proposed in Valkanov (2003) and Hjalmarsson (2011) are not a result of scaling the test statistics. If one uses Newey–West standard errors to calculate the (non scaled) $t$-statistic from the long-run OLS regression, a similar pattern emerges; note that the test based on Newey–West standard errors will be biased both because the standard errors do not properly control for the overlap in the data, and because the $t$-statistic from the long-run OLS regression does not control for the endogeneity in the regressors. The Newey–West standard errors were calculated using $q$ lags. For exogenous regressors ($\delta = 0$), the top panels in Fig. 3 show that the power of the scaled OLS $t$-statistic and Valkanov’s test statistic are almost identical and again seem to converge to some fixed level less than one. The $t$-statistic based on Newey–West standard errors also exhibits the same pattern;
here, the bias in this test resulting from the overlap in the data alone is evident, with a rejection rate around 20 percent under the null.

4. Asymptotic power properties of long-run tests

The simulation evidence in the previous section raises questions about the properties of long-run tests under the alternative of predictability. In particular, the power of the tests does not seem to converge to one as the slope coefficient increases and for endogenous regressors the power curves appear to sometimes decrease as the slope coefficient drifts away from the null hypothesis. In this section, I therefore derive some analytical results for the power properties of long-run tests. I focus on the standard (scaled) OLS $t$-statistics, derived by Valkanov (2003) and Hjalmarsson (2011), since the behavior of Hjalmarsson’s scaled Bonferroni test statistic is similar to the scaled OLS $t$-test with exogenous regressors.

**Theorem 1.** Suppose the data is generated by Eqs. (2) and (3). Under the alternative hypothesis that $\beta \neq 0$: 
For a fixed $q$ as $T \to \infty$,

$$
\frac{q}{T} \frac{t_q}{\sqrt{q}} \Rightarrow \frac{\beta}{\sqrt{\left( \frac{\alpha q}{q^2} + \beta \frac{1}{q} \left( 1 - \frac{1}{q} \right) \omega_{12} + \beta^2 g(q) \omega_{22} \right) \left( \int_0^1 f_x^2 \right)^{-1}}},
$$

where $g(q) = \frac{1}{6\sqrt{q}} - \frac{1}{2q} + \frac{1}{T}$.

(ii) As $q, T \to \infty$, such that $q/T \to \lambda$,

$$
\frac{t_q}{\sqrt{T}} \Rightarrow \sqrt{\left( \int_0^{1-\lambda} f_x^2 - \int_0^{1-\lambda} f_x^2 \right) \left( \int_0^{1-\lambda} f_x^2 \right)^{-1}} = O_p(1).
$$

Here $J_+ = \int_0^r e^{t\xi}d\mathcal{B}_2(s)$, $J_- = J_+ - \int_0^1 J_+$. $J_+(r; \lambda) = \int_r^{\lambda r} J_+(r)$, and $B(\cdot) = (B_1(\cdot), B_2(\cdot))'$ denotes a two dimensional Brownian motion with variance–covariance matrix $\Omega$. 

Fig. 3. Power curves for $T = 500$, and $q = 100$. The graphs show the average rejection rates for a one-sided 5 percent test of the null hypothesis of no predictability against a positive alternative. The $x$-axis shows the true value of the parameter $\beta$, and the $y$-axis indicates the average rejection rate. The left-hand graphs give the results for the case of $c = 0 (\rho = 1)$, and the right-hand graphs give the results for $c = -10 (\rho = 0.98)$. The two top graphs correspond to exogenous regressors ($\delta = 0$) and the two bottom graphs correspond to endogenous regressors ($\delta = -0.9$). The results for Hjalmarsson’s scaled OLS $t$-test are given by the solid lines, the results for Valkanov’s sup-bound test are given by the dotted lines, the results for the (non-scaled) $t$-test using Newey–West standard errors are given by the dotted and dashed lines, and the results for Hjalmarsson’s Bonferroni test are given by the dashed lines. The results are based on the Monte Carlo simulations described in the main text, and the power is calculated as the average rejection rate over 10,000 repetitions.
Part (i) of the theorem provides the limiting distribution of the scaled OLS $t$-test analyzed by Hjalmarsson (2011), and part (ii) provides results for Valkanov’s (2003) scaled test. The results in part (i) can be simplified through some approximations. In particular, for typical values of $q,g(q) \approx \frac{1}{q}$ and $\frac{1}{q}(1-\frac{1}{q}) \approx \frac{1}{q}$. Therefore, the following result holds approximately,

$$\frac{q}{T} \frac{t_q}{\sqrt{q}} \Rightarrow \frac{\beta}{\sqrt{\left(\frac{\omega_1}{q} + \beta \frac{\omega_2}{q} + \beta^2 \frac{\omega_3}{q}\right) \left(\frac{T}{q}\right)^{1/2}}}.$$

In addition, if $\beta$ is large, the third term in the denominator of the limiting distribution in Eq. (6) will dominate, and the following simplification is approximately true,

$$\frac{q}{T} \frac{t_q}{\sqrt{q}} \Rightarrow \frac{\beta}{\sqrt{\frac{3}{\omega_2} \left(\int_0^1 I^2\right)^{1/2}}}.$$

The analysis in part (i) yields more detailed and easily interpretable results than in part (ii), and I focus the discussion below on the results in part (i). The simulations presented above, especially in Fig. 3, show qualitatively identical results for Hjalmarsson’s (2011) scaled OLS $t$-test and Valkanov’s (2003) scaled sup-bound test, so the conclusions drawn from Hjalmarsson’s OLS $t$-test should extend to Valkanov’s test, as well as to general “robust” $t$-statistics, such as those formed from Newey–West standard errors shown in Fig. 3.

Part (i) of the theorem shows that the power of the scaled OLS $t$-test analyzed by Hjalmarsson (2011) will increase with the relative size of the sample to the forecasting horizon; thus, as long as the ratio between $q$ and $T$ is fixed, there are no asymptotic power gains. That is, $t_q/\sqrt{q} = O_p(T/q)$.\(^6\)

For large values of the slope coefficient $\beta$, the power is also independent of the value of $\beta$, as shown in Eq. (7). This explains the leveling out of the power curves as $\beta$ grows large, and their failure to converge to one for large values of $\beta$. The intuition behind the independence of $\beta$ in the limiting distribution is best understood by explicitly writing out the equation for the long-run returns under the alternative of predictability. That is, since the true model is given by Eqs. (2) and (3), the long-run regression equation is a fitted regression, rather than the data generating process. As shown in Appendix A, under the alternative of predictability, the long-run returns $r_{t+q}(q)$ actually satisfy the following relationship when ignoring the constant, derived from Eqs. (2) and (3),

$$r_{t+q}(q) = \beta_q x_t + \beta \sum_{h=1}^{q-1} \left(\sum_{p=h}^{q-1} \rho^{p-h}\right) v_{t+h} + u_{t+q}(q),$$

where $\beta_q = \beta(1 + \rho + \ldots + \rho^{q-1})$ and $u_{t+q}(q) = \sum_{i=1}^{q} u_{t+i}$. There are now, in effect, two error terms, the usual $u_{t+q}(q)$ plus the additional term $\beta \sum_{h=1}^{q-1} \left(\sum_{p=h}^{q-1} \rho^{p-h}\right) v_{t+h}$, which stems from the fact that at time $t$ there is uncertainty regarding the path of $x_{t+j}$ for $j = 1, \ldots, q - 1$. That is, when forming $q$-period ahead forecasts, there is uncertainty regarding both the future realizations of the returns and the predictor variable. The error term $\beta \sum_{h=1}^{q-1} \left(\sum_{p=h}^{q-1} \rho^{p-h}\right) v_{t+h}$ grows with the size of the slope coefficient, and will dominate the behavior of the joint error term, $\beta \sum_{h=1}^{q-1} \left(\sum_{p=h}^{q-1} \rho^{p-h}\right) v_{t+h} + u_{t+q}(q)$, for large values of $\beta$. The multiplication by $\beta$ in this error term cancels out the power gains that would otherwise occur as $\beta$ drifts further away from zero.

As seen in the power curves in the figures above, it is clear that for small values of $\beta$, the power of the long-run tests is a function of $\beta$. And, in particular, there appears to be regions of the parameter space where the power of the tests are decreasing in the magnitude of the slope coefficient. Eq. (6)\(^6\) Heuristically, this result also suggests that for a fixed sample size, more powerful tests of predictability are achieved by setting the forecasting horizon as small as possible, which is also in line with the simulation results presented above and in Hjalmarsson (2011). The dgp specified by Eq. (2) coincides with the one-period ($q = 1$) fitted regression, and it may therefore not seem surprising that tests based on shorter horizons appear to have better properties (see also discussion in Hjalmarsson, 2011). As pointed out in the introduction, the goal of this study is not to analyze power across horizons as such, but rather to try to glean some general understanding of the power properties of long-horizon tests.
highlights these effects. To form an intuition, consider again the representation of the long-run regression in Eq. (8). For small values \( \beta \), the usual error term \( u_{t+q}(q) \) and the additional term \( \beta^{q} \sum_{h=1}^{q-1} \left( \sum_{p=1}^{h-1} \rho^{p-h} \right) u_{t+h} \) will both be of a similar order of magnitude. When calculating the variance of the fitted residuals, which enters into the denominator of the \( t \)-statistic, the variance of both of these terms as well as their covariance will enter, as seen in Eq. (6). The covariance \( \omega_{12} \), when it is negative, will induce the non-monotonicity in the \( t \)-statistic as a function of \( \beta \). Initially, as the slope coefficient drifts away from zero, the first term in the denominator will dominate and the power of the test is increasing in \( \beta \), since the variance of \( u_{t+q}(q) \) is independent of \( \beta \). In a middle stage, the covariance term becomes important as well and the \( t \)-statistic decreases with the slope coefficient. Finally, as \( \beta \) grows large, the last term dominates and will exactly cancel out the dependence on \( \beta \) in the numerator and denominator.

Part (ii) of the theorem states that Valkanov’s scaled \( t \)-statistic converges to a well defined limiting distribution that is independent of \( \beta \) and \( T \), although it is a function of \( \lambda = q/T \). Thus, under the assumptions on \( q \) and \( T \) maintained by Valkanov, the \( t \)-statistic scaled by \( \sqrt{T} \) does not diverge and hence the power of the test does not converge to one. Of course, for a fixed \( q/T \), the same heuristic result follows from part (i), since as long as \( q/T \) does not change, there are no power gains. The independence of \( \beta \) for the power of Valkanov’s test follows the same intuition as in part (i) of the Theorem.

5. Conclusion and discussion

Analyzing power in long-horizon regressions is difficult because the long-horizon regression should be interpreted as a fitted regression and not the data generating process (dgp). The linear framework defined by Eqs. (2) and (3) is probably the only widely-agreed-upon predictive model for stock returns, the leading application of long-horizon regressions. But how sensitive are the power results to this exact model specification? With sufficiently strong non-linearities in the data generating process, it seems likely that any of the results could be overturned. For instance, Rapach and Wohar (2005) allow non-linearities in the predictor variable, and show that this can qualitatively change the relative power across different forecasting horizons. For moderate deviations from the exact model specified here, there are, however, reasons to believe that the results in this paper would continue to hold.

The lack of convergence to unity of the power curves as the slope coefficient drifts further away from zero effectively results from uncertainty regarding the future path of the predictor variable. When this uncertainty dominates the uncertainty coming from the non-predictive part of the future returns, there will be a cancelling out of the power gains from greater predictability through a larger slope coefficient and the power loss from a greater residual “error” term in the regression. How powerful this effect is will likely depend on the persistence of the regressors, with a stronger effect the more persistent the regressor. The non-monotonic power curves for endogenous regressors stem from a correlation between the forecast error due to uncertainty regarding the future path of the predictor and the forecast error due to the idiosyncratic innovations in the returns. Again, this seems like a fairly general result. Hjalmarsson (2011) shows that, under the null hypothesis of no predictability, the rate of convergence for the OLS estimator of the slope coefficient in a long-horizon regression is given by the ratio of the forecasting horizon to the sample size. It therefore seems likely that for local deviations from this null hypothesis, the rate of convergence remains similar, and that, asymptotically, power will only increase if the sample size increases relative to the forecasting horizon. These conjectures implicitly rely on linear relationships in the data, but suggest that the results derived here should be of relevance also for moderate deviations from the baseline model.

Appendix A. Proof of Theorem 1

For ease of notation the case with no intercept is treated. The results generalize immediately to regressions with fitted intercepts by replacing all variables by their demeaned versions. Note that, by standard arguments, \( T^{-1/2} \sum_{t=1}^{T} w_{t} \Rightarrow B(r) = BM(\Omega)(r) \), where \( B(\cdot) = (B_{1}(\cdot), B_{2}(\cdot))' \) denotes a two dimensional Brownian motion and \( \Rightarrow \) denotes weak convergence of the associated probability mea-
sures. Further, as $T \to \infty$, $T^{-1/2} \mathbf{x}_{t|T} \Rightarrow J_e(r) = \int_0^r e^{(r-s)c} dB_2(s)$ and an analogous result holds for the demeaned variables $\mathbf{x}_t = \mathbf{x}_t - T^{-1} \sum_{t=1}^n \mathbf{x}_t$, with the limiting process $J_e$ replaced by $J_e = J_e - \int_0^t J_e$ (Phillips, 1987).

(i) Consider first the case when $q$ is fixed. Under the alternative of predictability, by summing up on both sides in Eq. (2), it follows that

$$
\begin{align*}
    r_{t+q}(q) &= \beta (x_t + x_{t+1} + \cdots + x_{t+q-1}) + u_{t+q}(q) \\
               &= \beta \left( x_t + \rho x_t + \cdots + \rho^{q-1} x_t + v_{t+1} + (\rho v_{t+1} + v_{t+2}) + \cdots + \sum_{p=2}^q \rho^{q-p} v_{t+p-1} \right) + u_{t+q}(q) \\
               &= \beta g x_t + \beta \sum_{h=1}^{q-1} \left( \sum_{p=h}^{q-1} \rho^{p-h} \right) v_{t+h} + u_{t+q}(q),
\end{align*}
$$

where $\beta_g = \beta (1 + \rho + \cdots + \rho^{q-1})$. Using the expression above, and the results in Hjalmarsson (2008), it follows that $\hat{u}_{t+q}(q) = r_{t+q}(q) - \hat{\beta}_g x_t = \beta \sum_{h=1}^{q-1} \left( \sum_{p=h}^{q-1} \rho^{p-h} \right) v_{t+h} + u_{t+q}(q) + o_p(1)$. Now, consider

$$
\begin{align*}
    \frac{1}{qT} \sum_{t=1}^{T-q} \left( \beta \sum_{h=1}^{q-1} \left( \sum_{p=h}^{q-1} \rho^{p-h} \right) v_{t+h} + u_{t+q}(q) \right)^2 &= \frac{\beta^2}{qT} \sum_{t=1}^{T-q} \left( \sum_{h=1}^{q-1} \left( \sum_{p=h}^{q-1} \rho^{p-h} \right) v_{t+h}^2 + \frac{1}{qT} \sum_{t=1}^{T-q} u_{t+q}(q)^2 \right)^2 \\
    &+ 2\beta \frac{1}{qT} \sum_{t=1}^{T-q} \sum_{h=1}^{q-1} \sum_{k=1}^{q-1} (q-h)(q-k) v_{t+h} v_{t+k} + o_p(1)
\end{align*}
$$

Since $\rho = 1 + c/T, \sum_{p=h}^{q-1} \rho^{p-h} = \sum_{p=h}^{q-1} (1 + c/T)^{p-h} = (q-h) + O(T^{-1})$. Thus, for a fixed $q$ as $T \to \infty$,

$$
\begin{align*}
    I &= \beta^2 \frac{1}{qT} \sum_{t=1}^{T-q} \sum_{h=1}^{q-1} \sum_{k=1}^{q-1} (q-h)(q-k) v_{t+h} v_{t+k} + o_p(1) \\
    &= \beta^2 \frac{1}{qT} \sum_{t=1}^{T-q} \sum_{h=1}^{q-1} (q-h)^2 v_{t+h}^2 + o_p(1) \to B^2 q^2 g(q) \omega_{12},
\end{align*}
$$

where $g(q) \equiv \frac{1}{q(q-1)} - \frac{1}{q^2} + \frac{1}{q}$. The second equality follows from the mds assumption on $v_t$, and the asymptotic limit follows from the law of large numbers (LLN) and $\frac{1}{q} \sum_{h=1}^{q-1} (q-h)^2 = \frac{1}{6} - \frac{1}{2} q + \frac{1}{4} q^2 = \frac{1}{4} g(q)$. Similarly, using $u_{t+q}(q) = \sum_{j=t+q}^q u_{t+j}$, as $T \to \infty$ with $q$ fixed,

$$
\begin{align*}
    II &= 2\beta \frac{1}{T} \sum_{t=1}^{T-q} \sum_{h=1}^{q-1} \sum_{j=1}^{q-1} (q-h) v_{t+h} u_{t+j} + o_p(1) \\
    &= 2\beta \frac{1}{T} \sum_{t=1}^{T-q} \frac{q}{q-1} (q-h) v_{t+h} u_{t+h} + o_p(1) \to 2 \beta \left( \frac{1 - 1}{q} \right) \omega_{12},
\end{align*}
$$

where the second equality follows from the mds assumption, and the asymptotic limit again follows from the LLN and $\frac{1}{T} \sum_{h=1}^{q-1} (q-h) = q \left( \frac{1}{2} - \frac{1}{q-1} \right)$. From Hjalmarsson (2011), $II = \frac{1}{qT} \sum_{t=1}^{T-q} u_{t+q}(q)^2 \to \beta \omega_{11}$ as $T \to \infty$ for a fixed $q$. The scaled $t$-statistic, for $H_0 : \beta = 0$, thus satisfies

$$
\frac{1}{T} \frac{t_q}{\sqrt{q}} \to \begin{pmatrix} \frac{1}{q} \beta_g \end{pmatrix} \begin{pmatrix} \frac{1}{q} \sum_{t=1}^{T-q} u_t^2(q) \\
\frac{1}{T} \sum_{t=1}^{T-q} x_t^2 \end{pmatrix}^{-\frac{1}{2}} \to \sqrt{\left( \omega_{11} + \beta \left( 1 - \frac{1}{q} \right) \omega_{12} + \beta^2 q^2 g(q) \omega_{22} \right) \left( \frac{q}{q} \right)}.
$$

(ii) Consider next the case when $q/T = \lambda$ as $q, T \to \infty$. By summing up on both sides in Eq. (2),

$$
\begin{align*}
    r_{t+q}(q) &= \beta (x_t + x_{t+1} + \cdots + x_{t+q-1}) + u_{t+q}(q) \\
               &= \beta x_t + \beta \sum_{h=1}^{q-1} \left( \sum_{p=h}^{q-1} \rho^{p-h} \right) v_{t+h} + u_{t+q}(q).
\end{align*}
$$
\[ \hat{\beta}_q = \left( \sum_{t=1}^{T-q} r_{t+q}(q)x_t \right) \left( \sum_{t=1}^{T-q} x_t^2 \right)^{-1} = \beta \left( \sum_{t=1}^{T-q} x_{t+q-1}(q)x_t \right) \left( \sum_{t=1}^{T-q} x_t^2 \right)^{-1} + \left( \sum_{t=1}^{T-q} u_{t+q}(q)x_t \right) \left( \sum_{t=1}^{T-q} x_t^2 \right)^{-1}. \]

With \( q/T = \lambda \) as \( q, T \to \infty \), \( \frac{u_t(q)}{\sqrt{T}} = \frac{1}{\sqrt{T}} \sum_{i=1}^q u_{t+i} \to B_1(r + \lambda) - B_1(r) \equiv B_1(r; \lambda) \), and
\[
\left( \frac{1}{T} \sum_{t=1}^{T-q} u_{t+q}(q)x_t \right) \left( \frac{1}{T} \sum_{t=1}^{T-q} x_t^2 \right)^{-1} \Rightarrow \left( \int_0^{1-\lambda} B_1(r; \lambda)J(r)dr \right) \left( \int_0^{1-\lambda} f_r^2 \right)^{-1} = O_p(1).
\]
Similarly, \( \frac{1}{T} \sum_{t=1}^{T-q} x_{t+q-1}(q)x_t \Rightarrow J_t^* \equiv J_t(r; \lambda) \), and
\[
\left( \frac{1}{T} \sum_{t=1}^{T-q} x_{t+q-1}(q)x_t \right) \left( \frac{1}{T} \sum_{t=1}^{T-q} x_t^2 \right)^{-1} \Rightarrow \left( \int_0^{1-\lambda} J_t(r; \lambda)J_t(r) \right) \left( \int_0^{1-\lambda} f_r^2 \right)^{-1}.
\]

It follows that \( \frac{\hat{\beta}_q}{\sqrt{T}} \Rightarrow \beta \left( \int_0^{1-\lambda} J_t(r; \lambda)J_t(r) \right) \left( \int_0^{1-\lambda} f_r^2 \right)^{-1}. \) The fitted residuals satisfy
\[
\frac{1}{T} \sum_{t=1}^{T-q} \hat{u}_t(q)^2 = \frac{1}{T} \sum_{t=1}^{T-q} \left( \beta x_{t+q-1}(q) + u_{t+q}(q) - \hat{\beta}_q x_t \right)^2 = \beta^2 \left( \frac{1}{T} \sum_{t=1}^{T-q} \left( x_{t+q-1}(q) \right)^2 \right) + O_p(T^{-1})
\]
\[
\Rightarrow \beta^2 \int_0^{1-\lambda} J_t(r; \lambda)^2.
\]

Putting these results together, as \( q, T \to \infty \), with \( q/T = \lambda \),
\[
\frac{t_q}{\sqrt{T}} = \frac{\hat{\beta}_q \sqrt{T} \left( \frac{1}{T} \sum_{t=1}^{T-q} x_t^2 \right)^{-1}}{\sqrt{\beta^2 \int_0^{1-\lambda} J_t(r; \lambda)^2}} \Rightarrow \beta \left( \int_0^{1-\lambda} J_t(r; \lambda)J_t(r) \right) \left( \int_0^{1-\lambda} f_r^2 \right)^{-1/2} \rightarrow \beta \left( \int_0^{1-\lambda} J_t(r; \lambda)J_t(r) \right) \left( \int_0^{1-\lambda} f_r^2 \right)^{-1/2}.
\]

\section*{References}


