

Properties of the power envelope for tests against both stationary and explosive alternatives: the effect of trends

Patrick Marsh*

School of Economics and Granger Centre for Time Series Econometrics

University of Nottingham

University Park

NG7 2RD

email: patrick.marsh@nottingham.ac.uk

February 11th 2019

Abstract

This paper details a precise analytic effect that inclusion of a linear trend has on the power of Neyman-Pearson point optimal unit root tests and thence the power envelope. Both stationary and explosive alternatives are considered. The envelope can be characterized by probabilities for two, related, sums of chi-square random variables. A stochastic expansion, in powers of the local-to-unity parameter, of the difference between these loses its leading term when a linear trend is included. This implies that the power envelope converges to size at a faster rate, which can then be exploited to prove that the power envelope must necessarily be lower. This effect is shown to be, analytically, greater asymptotically than in small samples and numerically far greater for explosive than for stationary alternatives. Only a linear trend has a specific rate effect on the power envelope, however other deterministic variables will have some effect. The methods of the paper lead to a simple direct measure of this effect which is then informative about power, in practice.

*Thanks are due to Francesco Bravo, Giovanni Forchini, David Harvey, Steve Leybourne, Peter Phillips, Robert Taylor and participants at seminars at the Universities of Birmingham, Manchester, Monash and York.

1 Introduction

The power envelope is a fundamental measure of how effectively we can discriminate between false null hypotheses and specified alternatives. Every new unit root test, whether testing against stationary or explosive/bubble alternatives, must have its power characteristics compared with this envelope. Despite this, the analytic properties of the unit root power envelope are generally unknown. The focus has instead been on the stochastic properties of tests and estimators, capitalizing on the pioneering methods of Phillips (1987a, 1987b) and Chan and Wei (1987).

This paper seeks to capture the precise effect, on the power envelope, of the inclusion of a linear trend. For Economic data the unit root remains one of the most tested hypotheses. And the inclusion or otherwise of a linear trend has a profound effect on both the theoretical and observed properties of unit root tests, see both Elliott, Rothenberg and Stock (1996) and Nielsen (2008). To emphasize the importance of this, as measured via numerical resolution of the asymptotic power envelope of the former paper and in the context of an autoregressive parameter $\rho_T = 1 + c/T$, tests can have 50% power against a local alternative value $c = -7$ with no linear trend, but not until $c = -13.5$ if there is. The net effect of a linear trend on power is equivalent to a practitioner discarding 48% of their data. There is no other context in Econometrics where the effect of a single regressor is so profound.

Since there is no uniformly best invariant (UBI) test against either stationary or explosive alternatives the power envelope is constructed via the union of the powers of the continuum of point optimal tests. For each, a critical value is first required to fix size under the null, before its power is evaluated under the alternative. Therefore, for every value under the alternative two probabilities must be considered. In this paper these are characterized via probabilities for two, related, weighted sums of chi-squared random variables, similar to the original representations in Dickey and Fuller (1979). These two weighted sums generally have a stochastic difference, near $c = 0$, of order $O_p(c^2)$. When there is a linear trend, this falls to $O_p(c^4)$, asymptotically.

This induces a change in the rate of convergence of the power envelope itself to the chosen size. Specifically, for arbitrarily small positive ϖ_1 and ϖ_2 it is $O(c^{2-\varpi_1})$ when there is no linear trend and $O(c^{4-\varpi_2})$, when there is. This step change can be exploited to formally prove that powers of linear trend invariant tests are necessarily

lower. Intuitively this arises because the covariance and its derivative are proportional, when there is a unit root. The algebraic mechanism by which this occurs can also be used to construct a simple measure of the impact of regressor invariance on any hypothesis on the covariance structure of data. In the current context this measure correlates very well with power.

The next Section presents the main results, two Lemmas (proved in the online supplementary material to this paper) and a Theorem detailing the analytic effects of a trend which is proved in the Appendix. Section 3 discusses the implications of these results utilizing numerical results also presented in tables in the supplementary material.

2 Characterization of the Power Envelope and its Properties

The Gaussian power envelope is constructed from the powers of each point optimal test, e.g. see King (1980) and King and Srikanthakumar (2015). However, as is clear from Elliott, Rothenberg and Stock (1996) and Marsh (2011), the asymptotic distribution of these tests is the same under far more general assumptions. Let $(y_t)_{t=1}^T$ be generated from,

$$y_t = d_t + u_t \quad : \quad d_t = x_t' \beta \quad \text{and} \quad u_t = \rho_T u_{t-1} + \varepsilon_t, \quad (1)$$

where x_t is a $k \times 1$ deterministic regressor, β a $k \times 1$ unknown parameter, ε_t is a zero mean error process and we put $\rho_T = 1 + c/T$.

We will consider tests of $H_0 : c = 0$ against both stationary (S) and explosive (E) alternatives, as in

$$H_1^S : c < 0 \quad \text{and} \quad H_1^E : c > 0. \quad (2)$$

Elliott, Rothenberg and Stock (1996), under their Condition A, provide representations of the power envelope against H_1^S in two cases. First when $d_T = o(T^{1/2})$ (their equation (4)) and second when $d_t = \beta_1 + \beta_2 t$ (their equation (8)). Here we will denote those two, size α , envelopes by $\Pi_\alpha^\mu(c)$ and $\Pi_\alpha^\tau(c)$, respectively. Full expressions for these are also provided in the supplementary material. Although originally provided only for tests against H_1^S , power envelopes for H_1^E can also be generated using the

results of Phillips (1987b) and Chan and Wei (1987), see for example Harvey and Leybourne (2014).

Let $y = (y_1, \dots, y_T)'$, $X = (x_1, \dots, x_T)'$, $\beta = (\beta_1, \dots, \beta_k)'$, L be a lower triangular matrix with 1's on the first lower diagonal and 0's elsewhere, $\Delta_\rho = I - \rho L$ and let $W = \Delta_1 X$. Put $n = T - k$, and define

$$A = A_c = C' \Delta_1 (\Delta'_{1+c/T} \Delta_{1+c/T})^{-1} \Delta'_1 C, \quad (3)$$

where $CC' = M_W = I - W(W'W^{-1})W'$ and $C'C = I_n$, and let λ_i , $i = 1, \dots, n$, be the ordered eigenvalues of A . Finally let $z = (z_1, \dots, z_n) = C' \Delta_1 y$ and define the following two statistics;

$$Q_{0,n}(c) = \sum_{i=1}^n \lambda_i^{-1} \bar{z}_i^2 \quad \text{and} \quad Q_{1,n}(c) = \left(\sum_{i=1}^n \lambda_i \bar{z}_i^2 \right)^{-1}, \quad (4)$$

where $\bar{z}_i = z_i / \sqrt{\sum_{i=1}^n z_i^2}$.

The following Lemma provides alternative characterizations of the asymptotic power envelopes as well as a stochastic expansion of the limits of the two statistics defined in (4). Both the general assumptions under which it applies and its algebraic demonstration is given in the supplementary material.

Lemma 1 (i) Let $\{\tilde{z}_i\}_{i \in \mathbb{Z}}$ denote a sequence of IID $N(0, 1)$ random variables, then the asymptotic power envelope at size α for testing either H_1^S and H_1^E , for any set of explanatory variables X , can be characterized by

$$\Pi_\alpha(c) = \Pr[Q_1(c) < \kappa_\alpha], \quad Q_1(c) = \lim_{n \rightarrow \infty} \left(\sum_{i=1}^n \lambda_i \tilde{z}_i^2 \right)^{-1}, \quad (5)$$

where the critical value κ_α is defined by

$$\alpha = \Pr[Q_0(c) < \kappa_\alpha], \quad Q_0(c) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \lambda_i^{-1} \tilde{z}_i^2. \quad (6)$$

(ii) Denote the j^{th} derivative of A with respect to c , evaluated at 0, by D_j , and let $\{\delta_{1,i}\}_1^\infty$, $\{\delta_{2,i}\}_1^\infty$ and $\{\delta_{12,i}\}_1^\infty$ be the eigenvalues of the matrices D_1 , D_2 and $D_1 D_2$, and let $\{\xi_{1,i}^2\}_1^\infty$, $\{\xi_{2,i}^2\}_1^\infty$ and $\{\xi_{12,i}^2\}_1^\infty$ be sequences of independent chi-square variables, then in a neighbourhood of $c = 0$,

$$Q_1(c) - Q_0(c) = c^2 \gamma_1 + \frac{c^3}{6} \gamma_2 + O_p(c^4),$$

where γ_1 and γ_2 are defined by

$$\gamma_1 = \lim_{n \rightarrow \infty} \left\{ \left(\frac{1}{n} \sum_{i=1}^n \delta_{1,i} \xi_{1,i}^2 \right)^2 - \frac{1}{n} \sum_{i=1}^n \delta_{1,i}^2 \xi_{1,i}^2 \right\},$$

and

$$\begin{aligned} \gamma_2 = & \lim_{n \rightarrow \infty} \left\{ 3 \left(\frac{1}{n} \sum_{i=1}^n \delta_{1,i} \xi_{1,i}^2 \right) \left(\frac{1}{n} \sum_{i=1}^n \delta_{2,i} \xi_{2,i}^2 \right) \right. \\ & \left. - 6 \left(\left(\frac{1}{n} \sum_{i=1}^n \delta_{1,i} \xi_{1,i}^2 \right)^3 - \frac{1}{n} \sum_{i=1}^n \delta_{12,i} \xi_{12,i}^2 + \frac{1}{n} \sum_{i=1}^n \delta_{1,i}^3 \xi_{1,i}^2 \right) \right\}. \quad \blacksquare \end{aligned}$$

Lemma 1 provides a representation for the asymptotic power envelope in terms of an (infinite) weighted sum of chi-square random variables. Given that there is no UBI test, the properties of the power envelope can only be explored by directly comparing $Q_1(c)$ with $Q_0(c)$. Via the stochastic expansions presented in Lemma 1 (ii) we can establish the rate of convergence of the asymptotic power envelope to the chosen size, as in the following Lemma.

Lemma 2 *Let $\bar{Q} = c^{-2k} (Q_0(c) - Q_1(c))$, where k is such that \bar{Q} is $O_p(1)$, then the power envelope satisfies*

$$\Pi_\alpha(c) = \alpha + O(c^{2k-\varpi_k}),$$

where $\varpi_k > 0$ is an arbitrarily small constant. \blacksquare

Generally $Q_1(c) - Q_0(c) = O_p(c^2)$ and so the envelope is locally $\alpha + O(c^{2-\varpi_1})$. Now denote the column space of X by $\mathcal{M}(X)$ and the linear trend by $\tau = (t)_{t=1}^T$. Suppose now a linear trend is included in the regressors, i.e. $\tau \in \mathcal{M}(X)$, then the following Theorem, proved in the Appendix, demonstrates that in this case $Q_1(c) - Q_0(c) = O_p(c^4)$, and thus the envelope is $\alpha + O(c^{4-\varpi_2})$. As with Moon, Perron and Phillips (2007) the effect manifests itself as an order of magnitude step change in the order of convergence, although here in the parameter itself. These results hold in a neighbourhood of $c = 0$, however by exploiting the analytic properties of $\Pi_\alpha(c)$ these findings can be continued to demonstrate that inclusion of a linear trend necessarily implies the power envelope is strictly lower for any finite value of c .

Theorem 1 *i) Suppose that $\tau \in M(X)$, then*

$$Q_1(c) - Q_0(c) = O_p(c^4).$$

ii) At the α significance level,

$$\text{if } \tau \notin \mathcal{M}(X) \text{ then } \Pi_\alpha(c) - \alpha = O(c^{2-\varpi_1}), \text{ while}$$

$$\text{if } \tau \in \mathcal{M}(X) \text{ then } \Pi_\alpha(c) - \alpha = O(c^{4-\varpi_2}).$$

iii) Let the set of regressors X satisfy $x'_T\beta = o(T^{1/2})$, so that the power envelope is $\Pi_\alpha^\mu(c)$. If we add the column τ to X , then we obtain power envelope $\Pi_\alpha^\tau(c)$, which satisfies

$$\Pi_\alpha^\tau(c) < \Pi_\alpha^\mu(c),$$

for all finite c . ■

3 Analysis and Conclusions

i) Parts (i) and (ii) of Theorem 1 apply only in a neighbourhood of $c = 0$. However, the power envelope (via (5) and (6)) is a function of both regressor set X and local parameter c through the eigenvalues of the matrix A_c . Since these eigenvalues are analytic in c , then so is the power envelope. That $\Pi_\alpha^\tau(c)$ is smaller than $\Pi_\alpha^\mu(c)$ at some, local, value of c can therefore be analytically continued to all finite values of c . This finding links directly to findings in Nabeya and Tanaka (1990), which shows that there is no Locally Best Invariant test of a unit root when there is a linear trend although that paper contains no explicit results for the power envelope, itself. Equally Theorem 1 explains how the precise finding of Marsh (2007a) for the null $c = 0$ has, in fact, an impact for any finite value of c under the alternative. In the absence of any formal distribution theory, asymptotic or otherwise, for trend invariant estimators or tests, Theorem 1 offers the only analytic demonstration of the power loss of such trends, hitherto observed only experimentally.

ii) The local expansions remain informative about the power envelope globally. Table 1a in the supplementary material presents outcomes of the power envelope for

a variety of simple choices of d_t (a constant, a linear trend and trends involving the logarithm, square root, square and exponent of time) in (1) for $T = 250$. It is worth noting that not all trends are associated with low power, exponential trends imply powers similar to those of the constant case. In Table 1b the power envelopes are approximated using stochastic expansions of $Q_{0,n}(c)$ and $Q_{1,n}(c)$ to order $O_p(c^3)$. As is evident comparing across entries in Tables 1a and 1b simulation of just the leading terms of these statistics capture the envelope almost entirely.

iii) Although the results in Theorem 1 are asymptotic, their proof yields the insight that the effect of a linear trend can be greater asymptotically, than in finite samples. From the proof of Theorem 1(i), when there is a trend, the $O_p(c^3)$ term in the stochastic difference between $Q_{0,n}(c)$ and $Q_{1,n}(c)$ is

$$\frac{c^3}{2T} \left[\left(\sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{2,i}^2 - 2 \sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{12,i}^2 \right) \right]$$

which converges in probability to 0 only as $T \rightarrow \infty$. This indicates a differential relative effect that a linear trend has on the finite sample and asymptotic envelopes. To illustrate, Table 2 contains the ratios of the power envelopes evaluated for $d_t = \beta_1 + \beta_2 t$ and $d_t = \beta_1$ for values of c from 1.25 to -5.0 and for different significance levels, $\alpha = .01, .05, .10$ evaluated for sample sizes of $T = 50, 250, 500$. The effects are clear and significant, particularly when c is small. This difference in the behaviour of the asymptotic and finite sample envelopes has significance for the choice of unit root tests in practice. As Francke and de Vos (2007) note, tests designed to have power close to the asymptotic power envelope may not have power functions close to the finite sample one, in the presence of trends. This can only be explained via the quantitative difference between them found in this paper. It is also suggestive that new tests ought to be compared to *both* finite sample *and* asymptotic envelopes to justify their properties.

iv) The mechanism by which the power envelope is reduced on inclusion of a linear trend is algebraic. Specifically, as in the proof of Theorem 1(i), letting $\Sigma_{1+c/T} = \left(\Delta'_{1+c/T} \Delta_{1+c/T} \right)^{-1}$ be the covariance of a pure ‘near unit root’ process, then

$$\left. \frac{d\Sigma_{1+c/T}}{dc} \right|_{c=0} = - \left(\Delta_1^{-1} (\Delta_1^{-1})' - \tau \tau' \right).$$

To construct invariant tests we first let $w = C' \Delta_1 y$, which removes dependence on β . Let $V[w] = \sigma^2 A_c$ so that the null implies $H_0 : V[w] = \sigma^2 A_0 \propto I_n$. If we take a series

expansion of A_c around 0, $A_c = A_0 + \sigma^2 \sum_{j=1}^{\infty} \frac{c^j}{j!} D_j$, where $D_j = d^j A_c / dc^j|_{c=0}$ then when X contains a linear trend we find $D_1 = -T^{-1}I_n = -T^{-1}A_0$ and hence

$$\sigma^2 A_c = \sigma^2 (1 - cT^{-1}) A_0 + \sigma^2 \sum_{j=2}^{\infty} \frac{c^j}{j!} D_j.$$

That is $\sigma^2 A_c$ is proportional to $\sigma^2 A_0$ up to and including the $O(c)$ term when there is a linear trend. Since we also require scale invariance this, heuristically, captures the effective cause of the dramatic loss of power. Algebraically this proportionality is exact in the unit root/linear trend problem. Generically, suppose we wish to test $H_0 : V[w] = \sigma^2 A_0$ vs. $H_1 : V[w] = \sigma^2 A_c$. If the derivative of A_c at $c = 0$ is D_1 then we would expect low power if A_0 and D_1 are proportional. A simple measure of the proportionality of two matrices is the variation in the ratio of their respective ordered eigenvalues, λ_i and $\delta_{1,i}$. To proceed, let $\bar{\lambda}_X = n^{-1} \sum_{i=1}^n \lambda_i / \delta_{1,i}$ and $S_X^2 = \sum_{i=1}^n \left(\frac{\lambda_i}{\delta_{1,i}} - \bar{\lambda}_X \right)^2$, and define

$$\Lambda_X^2 = S_X^2 / S_{\emptyset}^2, \quad (7)$$

where \emptyset denotes the null set, i.e. X is empty and no invariance is required in the construction of w . Λ_X^2 thus measures the relative variation in eigenvalues for a given choice of X compared to the case of no regressors, i.e. only scale invariance is required. In the linear trend case $\bar{\Lambda}_X^2 = 0$. For example, in the cases enumerated in the supplementary material we find, with $T = 250$,

d_t	β_1	$\beta_1 + \beta_2 t$	$\beta_1 + \beta_2 \ln(t)$	$\beta_1 + \beta_2 \sqrt{t}$	$\beta_1 + \beta_2 t^2$	$\beta_1 + \beta_1 e^t$
$\bar{\Lambda}_X^2$	0.996	0.000	0.759	0.126	0.058	0.979

In terms of ranking these outcomes match perfectly the power envelopes given in Table 1a. This measure could be adapted for any (simple) hypothesis test on a covariance matrix, when invariance with respect to the mean is required. It provides a simple measure of the sensitivity of power to the choice of deterministics, similar in spirit to the analysis of Leamer (1985).

v) Bykhovskaya and Phillips (2018) explore tests involving functional local alternatives where the local parameter depends on time. e.g. $H_0 : c_t = 0$ vs. $H_1^F : c_t = c(t/T)/T$, so that only the initial value has a unit root. Although invariance with respect to neither the mean nor scale is pursued in that paper, it is trivial to apply

the framework here to such cases, as well as to functional stationary alternatives, where $c < 0$. Define $\tilde{\Delta}_c = I_T - L \left(I_T + c/T \text{diag} \{t/T\}_{t=1}^T \right)$ then the error covariance matrix under H_1^F is $\tilde{\Sigma}_c = \left(\tilde{\Delta}_c \tilde{\Delta}_c' \right)^{-1}$ and its derivative at $c = 0$ is the $T \times T$ matrix, $\tilde{\Sigma}_0^1$ having $(i, j)^{th}$ element $\max[i, j] - i\mathbb{I}(i = j)$, where $\mathbb{I}(\cdot)$ is the indicator function. In this case the covariance of $w = C' \Delta_1 y$ is $\tilde{A}_c = C' \Delta_1 \tilde{\Sigma}_c \Delta_1' C$ and has slope $\tilde{D}_1 = C' \Delta_1 \tilde{\Sigma}_0^1 \Delta_1' C$. Even in the case that X contains a linear trend \tilde{A}_0 is not proportional to D_1 .

Calculating the eigenvalue variation defined in (7) we find $\bar{\Lambda}_X^2 = 0.997$ for the constant case and $\bar{\Lambda}_X^2 = 0.993$ for the linear trend case, so the relative impact of a linear trend is extremely small when testing against functional alternatives. This is borne out in the outcomes for the ratios of the power envelopes presented in Table 3 which repeat the experiments reported in Table 2, but for the functional alternatives, H_1^F , in both stationary and explosive directions.

vi) The focus thus far has been on the theoretical implications on the testing problem of the inclusion of a linear trend. Harvey *et al* (2009) detail practical procedures which account for uncertainty over whether or not a trend is required. Numerically, their tests are shown to have power curves close to $\Pi_\alpha^\mu(c)$ when there is no trend, and close to $\Pi_\alpha^\tau(c)$ when there is. Marsh (2009) characterizes this uncertainty in terms of a Bernoulli mixture of the trend and no-trend cases. The results of this paper demonstrate, unequivocally, the necessity of the Harvey *et al* (2009) pre-test or union of rejections based tests. Specifically, under such uncertainty, any other test must either be inefficient (its power will be bounded by $\Pi_\alpha^\tau(c) < \Pi_\alpha^\mu(c)$) when no trend is present, or inconsistent when it is.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

References

- Bykhovskaya, A. and P.C.B. Phillips 2018. Boundary Limit Theory for Functional Local to Unity Regression. *J.T.S.A.* **39**: 523-562.
- Chan, N.H. and C.Z. Wei 1987. Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* **15**: 1050-1063.

- Dickey, D.A. and W.A. Fuller 1979. Distribution of the estimators for autoregressive time series with a unit root. *J.A.S.A.* **366**: 427-431.
- Elliott, G., Rothenberg, T.J. and J.H. Stock 1996. Efficient tests for an autoregressive unit root. *Econometrica* **64**: 813-836.
- Francke, M.K. and A.F. de Vos 2007. Marginal likelihood and unit roots. *Journal of Econometrics* **137**: 708-728.
- Harvey, D. I., Leybourne, S.J. and A.M.R. Taylor 2009. Unit Root Testing in Practice: Dealing with Uncertainty over the Trend and Initial Condition. *Econometric Theory*, **25**: 587-636.
- Harvey, D. I. and S.J. Leybourne 2014. Asymptotic behaviour of tests for a unit root against an explosive alternative. *Economics Letters* **122**: 64-68.
- King, M. L. 1980. Robust tests for spherical symmetry and their application to least squares regression. *Annals of Statistics* **8**: 1265–1271.
- King, M.L. and S. Srikanthakumar 2015. Point optimal testing: a survey of the post 1987 literature. *Model Assisted Statistics and Algorithms* **10**: 79-196.
- Leamer E.E. 1985. Sensitivity Analyses Would Help. *The American Economic Review* **75**: 308-313.
- Marsh, P. 2007a. The available information for invariant tests of a unit root. *Econometric Theory* **23**: 686-710.
- Marsh, P. 2009. Commentaries on ‘Unit Root Testing in Practice: Dealing with Uncertainty over the Trend and Initial Condition’. *Econometric Theory*, **25**: 637-657.
- Marsh, P. 2011. Saddlepoint and estimated saddlepoint approximations for optimal unit root tests. *Econometric Theory* **27**: 1026-1047.
- Moon, H.R., B. Perron and P.C.B. Phillips 2007. Incidental trends and the power of panel unit root tests. *Journal of Econometrics* **141**: 416-459.
- Nabeya, S. and K. Tanaka 1990. Limiting power of unit-root tests in time-series regression. *Journal of Econometrics* **46**: 247–271.
- Nielsen, B. 2008. Power of tests for unit roots in the presence of a linear trend. *Oxford Bulletin of Economics and Statistics* **70**: 619-644.
- Phillips, P.C.B. 1987a. Time Series Regression with a Unit Root. *Econometrica* **55**: 277-301.

Phillips, P.C.B. 1987b. Towards a unified asymptotic theory for autoregression. *Biometrika* **74**: 535-547.

Appendix

Proof of Theorem 1

Part (i). Suppose that X contains a linear trend. Let $e = (1, 1, 1, \dots, 1)'$ be the constant vector, so that a linear trend is defined by $\tau = \Delta_1^{-1}e$, where $\Delta_1 = I - L$. If $\mathcal{M}(X)$ is the column space of X , then

$$\tau \in \mathcal{M}(X) \rightarrow \Delta_1 \tau \in \mathcal{M}(\Delta_1 X) \equiv \mathcal{M}(W) \iff C' \Delta_1 \tau = 0,$$

where C is defined above. The first derivative of A_c at $c = 0$ is

$$\begin{aligned} D_1 &= \left. \frac{dA_c}{dc} \right|_{c=0} = C' \Delta_1 \left[\frac{d \left(\Delta_{1+c/T}^{-1} \left(\Delta_{1+c/T}^{-1} \right)' \right)}{dc} \right]_{c=0} \Delta_1' C \\ &= C' \Delta_1 \left[\frac{d \Delta_{1+c/T}^{-1}}{dc} \left(\Delta_{1+c/T}^{-1} \right)' + \Delta_{1+c/T}^{-1} \frac{d \left(\Delta_{1+c/T}^{-1} \right)'}{dc} \right]_{c=0} \Delta_1' C \\ &= -T^{-1} C' \Delta_1 \left[\Delta_1^{-1} L \Delta_1^{-1} \left(\Delta_1^{-1} \right)' + \Delta_1^{-1} \left(\Delta_1^{-1} L \Delta_1^{-1} \right)' \right] \Delta_1' C \\ &= -T^{-1} C' \Delta_1 \left[\Delta_1^{-1} \left(\Delta_1^{-1} \right)' - \tau \tau' \right] \Delta_1' C = -T^{-1} I_n, \end{aligned}$$

see also the proof of Theorem 1 in Marsh (2007a). Consequently and only when $\tau \in \mathcal{M}(X)$, $\delta_{1,i} = T^{-1}$ and $\delta_{12,i} = T^{-1} \delta_{2,i}$. Substituting these into the definitions of γ_1 and γ_2 in the statement of Lemma 1, we find

$$Q_{1,n}(c) - Q_{0,n}(c) = 0 + \frac{c^3}{T} \left[\left(2 \sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{2,i}^2 - \sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{12,i}^2 \right) \right] + O_p(c^4).$$

Since also both $\sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{2,i}^2$ and $\sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{12,i}^2$ are $O_p(1)$ as $n \rightarrow \infty$, then we find $Q_1(c) - Q_0(c) = \lim_{n \rightarrow \infty} (Q_{1,n}(c) - Q_{0,n}(c)) = O_p(c^4)$, as required.

Part (ii). When $\tau \notin \mathcal{M}(X)$, put $k = 1$ so $\bar{Q} = c^{-2} (Q_1(c) - Q_0(c))$ is $O_p(1)$ and hence, immediately, $\Pi_\alpha(c) - \alpha = O(c^{2-\varpi_1})$. However, when $\tau \in \mathcal{M}(X)$ put $k = 2$ and $\bar{Q} = c^{-4} (Q_1(c) - Q_0(c))$ is $O_p(1)$ so $\Pi_\alpha(c) - \alpha = O(c^{4-\varpi_2})$ instead.

Part (iii). Since point optimal tests are unbiased and the power envelope is monotone, then for any $c \neq 0$ both $\Pi_\alpha^\mu(c) > \alpha$ and $\Pi_\alpha^\tau(c) > \alpha$. Suppose first

that $c < 0$, then the difference in rates implied by part (ii) implies there exists some value $c^* < 0$ such that $\Pi_\alpha^\tau(c^*) < \Pi_\alpha^\mu(c^*)$.

Similar to the proof of Theorem 1 in Marsh (2011), let

$$\begin{aligned} F_{1,n}(\kappa_\alpha) &= \Pr[Q_{1,n}(c) < \kappa_\alpha] = \Pr\left[\left(\sum_{i=1}^n \lambda_i \tilde{z}_i^2\right)^{-1} < \kappa_\alpha\right] = \Pr\left[\sum_{i=1}^n \lambda_i \tilde{z}_i^2 > \kappa_\alpha\right] \\ &= 1 - \frac{1}{2\pi i} \int_{\tau-i\infty}^{\tau+i\infty} \frac{\exp\{n(R_1(\theta) - x\theta)\}}{\theta} d\theta, \end{aligned}$$

where $R_1(\theta) = -\frac{1}{2n} \sum_{i=1}^n \log(1 - 2\theta\lambda_i)$ is the mean cumulant generating function of $\sum_{i=1}^n \lambda_i \tilde{z}_i^2$. Note that $\Pi_\alpha(c) = \lim_{n \rightarrow \infty} F_{1,n}(\kappa_\alpha)$, i.e. the asymptotic distributions are defined as the limit of the finite sample, consistent with the set-up of Lemma 1.

Since A is an analytic function of c then so are the λ_i and hence so is $F_{1,n}(\kappa_\alpha)$, and its limit, through $R_1(\theta)$. Consequently both $\Pi_\alpha^\mu(c)$ and $\Pi_\alpha^\tau(c)$ are analytic in c , since they are functions only of the eigenvalues of A . Thus $\bar{\Pi}(c) = \Pi_\alpha^\mu(c) - \Pi_\alpha^\tau(c)$ is analytic in c . Note that $\bar{\Pi}(c) \geq 0$, since adding an additional invariance requirement - in this case for a linear trend - cannot increase power. Let \mathbb{R}^- denote the set of negative real numbers and let U be any closed subset of \mathbb{R}^- . A fundamental property of bounded analytic functions is that if $\bar{\Pi}(c) = 0$ for some $c \in U$ then $\bar{\Pi}(c) = 0$ for all $c \in U$. Since this is not true for $c = c^*$ then it cannot be true for any c satisfying $0 > c > -\infty$, consequently it must be that $\bar{\Pi}(c) > 0$ for all finite c . The proof for the case for $c > 0$ is identical. ■

Supplementary Material

The asymptotic power envelope of Elliott, Rothenberg and Stock (1996) was formally derived under their Condition A, although can be readily generalized, as in the following Assumption taken from Chang and Park (2002). Let $(y_t)_{t=1}^T$ be generated from

$$y_t = d_t + u_t \quad : \quad d_t = x_t' \beta \quad \text{and} \quad u_t = \rho_T u_{t-1} + \varepsilon_t, \quad (\text{S1})$$

and assume:

Assumption 1 (i) (η_t, \mathcal{F}_t) is a martingale difference sequence with filtration \mathcal{F}_t , and such that a) $E[\eta_t^2] = \sigma^2$, b) $\text{plim } T^{-1} \sum_{t=1}^T \eta_t^2 = \sigma^2$, c) $E[|\eta_t|^s] \leq m_\varepsilon < \infty$ for $s \geq 4$,
(ii) in (S1) $\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \eta_{t-j}$, where $\sum_{j=0}^{\infty} j |\psi_j| \leq m_\psi < \infty$, $\psi_0 = 1$, and
(iii) $u_0 = o_p(T^{1/2})$. ■

Denoting the power envelope by $\Pi_\mu(c)$ when $d_t = o(t^{1/2})$, then

$$\Pi_\mu(c) = \Pr \left[c^2 \int_0^1 W_c^2(t) dt - c W_c^2(1) < b_\alpha^\mu(c) \right], \quad (\text{S2})$$

while when d_t also includes a linear trend:

$$\Pi_\tau(c) = \Pr \left[c^2 \int_0^1 V_c^2(t, c) dt - (1-c) V_c^2(1, c) < b_\alpha^\tau(c) \right]. \quad (\text{S3})$$

In (S2) and (S3) $V_c(t, c)$ is the Gaussian process,

$$V_c(t, c) = W_c(t) - t \left(\lambda W_c(1) - 3(1-\lambda) \int_0^t s W_c(s) ds \right),$$

with $\lambda = (1-c)/(1-c+c^2/3)$, $W_c(s) = \int_0^s e^{(r-s)c} dW(r)$, $W(r)$ is standard Brownian motion and $b_\alpha^\mu(c)$ and $b_\alpha^\tau(c)$ are critical values chosen so that the point optimal tests have size α . It is not apparent that $\Pi_\tau(c) < \Pi_\mu(c)$, unless these probabilities are numerically resolved.

Proof of Lemma 1:

(i) The asymptotic power envelope is the same for any process $\{y_t\}_{t \in \mathbb{Z}}$ constructed as in (S1) and satisfying Assumption 1. Therefore, without loss of any generality, we may assume that $\varepsilon_t \sim iidN(0, \sigma^2)$, $0 < \sigma < \infty$ and $u_0 = 0$. Let y, X, β and Δ_ρ be as

defined in the main text, and let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)$ then the equations in (S1) define the following generalized linear regression model, $y = X\beta + \Delta_{1+c/T}^{-1}\varepsilon$, or

$$\Delta_1 y = \Delta_1 X\beta + \Delta_1 \Delta_{1+c/T}^{-1} \varepsilon. \quad (\text{S4})$$

Following the analysis of Marsh (2007a), put $W = \Delta_1 X$ and apply the results of King (1980), then the maximal invariant for testing $H_0 : c = 0$ in (S4), is

$$v = \frac{w}{|w|} = \frac{C' \Delta_1 y}{\sqrt{y' \Delta_1' M \Delta_1 y}}, \quad M = I - W(W'W)^{-1}W',$$

where $CC' = M$ and $C'C = I_n$. Also from King (1980), the density of v (with respect to normalized Haar measure on the surface of the unit sphere in $n = T - k$ dimensions, \mathbb{S}^n) is

$$pdf(v) = |A|^{-1/2} (v' A^{-1} v)^{-(T-k)/2},$$

where $A = C' \Delta_1' \Delta_\rho^{-1} (\Delta_\rho^{-1})' \Delta_1 C$.

Immediately, and see also Podivinsky and King (2000), the Neyman-Pearson point optimal test of $H_0 : \rho = 1$ vs. $H_1 : \rho \neq 1$, is

$$\text{reject } H_0 \text{ if } P_v = v' A^{-1} v < \kappa_\alpha, \quad (\text{S5})$$

where the critical value κ_α in (S5) is chosen so that the size of the test is α .

Here the point optimal tests take the form of a quadratic form distributed on the sphere, rather than the difference in quadratic forms presented in Elliott, Rothenberg and Stock (1996). Marsh (2011) demonstrates the equivalency of these formulations using the marginal-likelihood methods developed in Francke and de Vos (2007), while Marsh (2007b) extends to the case of testing for a unit root in a lagged dependent variable.

Under $H_0 : c = 0$, $w = C' \Delta_1 y \sim N(0, \sigma^2 I_n)$, so the critical value k_α of the the point optimal test (S5) for testing H_0 against $H_1 : c \neq 0$, at size α , is defined by

$$\begin{aligned} \alpha &= \Pr[v' A^{-1} v < \kappa_\alpha | H_0] = \Pr\left[\frac{w' A^{-1} w}{w' w} < \kappa_\alpha | w \sim N(0, \sigma^2 I_n)\right] \\ &= \Pr\left[\frac{\sum_{i=1}^n \lambda_i^{-1} z_i^2}{\sum_{i=1}^n z_i^2} < \kappa_\alpha\right], \end{aligned}$$

where $z_i \sim iidN(0, 1)$ and the $\{\lambda_i\}_{i=1}^n$ are the ordered eigenvalues of A . The power of (S5) against $H_1 : c \neq 0$ is

$$\Pi_\alpha(c) = \Pr[v' A^{-1} v < \kappa_\alpha | H_1] = \Pr\left[\frac{w' A^{-1} w}{w' w} < \kappa_\alpha | w \sim N(0, \sigma^2 A)\right].$$

Letting $z = (z_1, \dots, z_n)' = \sigma^{-1} A^{-1/2} w \sim N(0, I_n)$, then

$$\Pi_\alpha(c) = \Pr \left[\frac{z' z}{z' A z} < \kappa_\alpha \right] = \Pr \left[\frac{\sum_{i=1}^n \lambda_i^{-1} z_i^2}{\sum_{i=1}^n z_i^2} < \kappa_\alpha \right],$$

again with $z_i \sim iid N(0, 1)$. Defining $\bar{z}_i = z_i / \sqrt{\sum_{i=1}^n z_i^2}$ and then

$$Q_{0,n}(c) = \sum_{i=1}^n \lambda_i^{-1} \bar{z}_i^2 \quad \text{and} \quad Q_{0,n}(c) = \left(\sum_{i=1}^n \lambda_i \bar{z}_i^2 \right)^{-1}$$

establishes part (i) of the Lemma via the limit $n \rightarrow \infty$.

(ii) Define the two infinite sums,

$$Q_0(c) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \lambda_i^{-1} \bar{z}_i^2 \quad \text{and} \quad Q_1(c) = \lim_{n \rightarrow \infty} \frac{1}{\sum_{i=1}^n \lambda_i \bar{z}_i^2},$$

then since $0 < \lambda_i < \infty$ both $Q_0(c)$ and $Q_1(c)$ are $O_p(1)$, and satisfy

$$\min_i (\lambda_i^{-1}) < Q_j(c) < \max_i (\lambda_i^{-1}), \quad j = 1, 2.$$

Because the power envelope is constructed, at each point, via probabilities involving both statistics a stochastic expansion the difference between $Q_1(c)$ (the random variable defining power) and $Q_0(c)$ (the random variable defining size) is required. When $c = 0$, we have that $\lambda_i = 1$ for all i , and so the problem is degenerate, with $Q_1(0) - Q_0(0) = 0$, or local to the origin in c ,

$$Q_1(c) = Q_0(c) + o_p(1).$$

For any finite c , A and A^{-1} are positive definite, symmetric and have eigenvalues and vectors satisfying, respectively,

$$\lambda_i A = \lambda_i v_i \quad \text{and} \quad \lambda_i^{-1} A^{-1} = \lambda_i^{-1} v_i,$$

with $v_i' v_i = 1$. A is analytic in ρ , and hence c , with derivatives

$$\frac{d^j A}{dc^j} = C' \Delta_1' \frac{d^j \Sigma_{1+c/T}}{dc^j} \Delta_1 C,$$

where $\Sigma_\rho = (T_\rho^{-1})' T_\rho$.

We can write down Taylor Series expansions around $c = 0$ of both $Q_{0,n}(c)$ and $Q_{1,n}(c)$. First we have,

$$Q_{0,n}(c) = \sum_{i=1}^n \lambda_i^{-1} \bar{z}_i^2 = 1 + \sum_{i=1}^n \left(\sum_{j=1}^{\infty} \frac{c^j}{j!} \frac{d^j \lambda_i^{-1}}{dc^j} \Big|_{c=0} \right) \bar{z}_i^2,$$

while for $Q_{1,n}(c)$ we instead have,

$$\begin{aligned} Q_{1,n}(c) &= \frac{1}{\sum_{i=1}^n \lambda_i \tilde{z}_i^2} = \frac{1}{\left(1 + \sum_{i=1}^n \left(\sum_{j=1}^{\infty} \frac{c^j}{j!} \frac{d^j \lambda_i}{dc^j} \Big|_{c=0}\right) \tilde{z}_i^2\right)} \\ &= 1 + \sum_{i=1}^n \left(\sum_{k=1}^{\infty} (-1)^k \left[\sum_{j=1}^{\infty} \frac{c^j}{j!} \frac{d^j \lambda_i}{dc^j} \Big|_{c=0} \tilde{z}_i^2\right]^k\right). \end{aligned}$$

Including up to $O_p(c^3)$ terms, we then obtain

$$Q_{0,n}(c) = 1 + c \sum_{i=1}^n \frac{d\lambda_i^{-1}}{dc} \Big|_{c=0} \tilde{z}_i^2 + \frac{c^2}{2} \sum_{i=1}^n \frac{d^2 \lambda_i^{-1}}{dc^2} \Big|_{c=0} \tilde{z}_i^2 + \frac{c^3}{6} \sum_{i=1}^n \frac{d^3 \lambda_i^{-1}}{dc^3} \Big|_{c=0} \tilde{z}_i^2 + O_p(c^4),$$

and

$$\begin{aligned} Q_{1,n}(c) &= 1 - c \sum_{i=1}^n \frac{d\lambda_i}{dc} \Big|_{c=0} \tilde{z}_i^2 - \frac{c^2}{2} \left[\sum_{i=1}^n \frac{d^2 \lambda_i}{dc^2} \Big|_{c=0} \tilde{z}_i^2 - 2 \left(\sum_{i=1}^n \frac{d\lambda_i}{dc} \Big|_{c=0} \tilde{z}_i^2 \right)^2 \right] \\ &\quad - \frac{c^3}{6} \left[6 \left(\sum_{i=1}^n \frac{d\lambda_i}{dc} \Big|_{c=0} \tilde{z}_i^2 \right)^3 - 3 \left(\sum_{i=1}^n \frac{d\lambda_i}{dc} \Big|_{c=0} \tilde{z}_i^2 \right) \left(\sum_{i=1}^n \frac{d^2 \lambda_i}{dc^2} \Big|_{c=0} \tilde{z}_i^2 \right) \right. \\ &\quad \left. + \sum_{i=1}^n \frac{d^3 \lambda_i}{dc^3} \Big|_{c=0} \tilde{z}_i^2 \right] + O_p(c^4). \end{aligned}$$

Considering first the $O_p(c)$ terms, notice that for all i ,

$$\frac{d\lambda_i}{dc} = v_i' \frac{dA}{dc} v_i \quad \text{and} \quad \frac{d\lambda_i^{-1}}{dc} = v_i' \frac{dA^{-1}}{dc} v_i$$

see Section 8.7 of Magnus and Neudecker (1988). We then have

$$\frac{d\lambda_i^{-1}}{dc} \Big|_{c=0} = v_i' \frac{dA^{-1}}{dc} v_i \Big|_{c=0} = -v_i' A^{-1} \frac{dA}{dc} A^{-1} v_i \Big|_{c=0} = -\frac{d\lambda_i}{dc} \Big|_{c=0},$$

since $A|_{c=0} = I_n$, and $\frac{dA^{-1}}{dc} = -A^{-1} \frac{dA}{dc} A^{-1}$. Immediately we then find that $Q_{1,n}(c) - Q_{0,n}(c) = O_p(c^2)$, for all cases.

For the $O_p(c^2)$ and $O_p(c^3)$ terms, although in principle expressions for the higher-order derivatives of the eigenvalues and their inverse can be found resolving them can prove prohibitive since they also involve the derivatives of the eigenvectors. Instead, for $Q_{1,n}(c)$, write

$$\begin{aligned} \sum_{i=1}^n \frac{d^j \lambda_i}{dc^j} \Big|_{c=0} \tilde{z}_i^2 &= \frac{d^j \sum_{i=1}^n \lambda_i \tilde{z}_i^2}{dc^j} \Big|_{c=0} = \frac{d^j \tilde{z}' A \tilde{z}}{dc^j} \Big|_{c=0} \\ &= \tilde{z}' \frac{d^j A}{dc^j} \tilde{z} \Big|_{c=0} = \tilde{z}' D_j \tilde{z} \Big|_{c=0}, \end{aligned} \tag{S6}$$

where $\tilde{z} = z/\sqrt{\sum_{i=1}^n z_i^2}$ and D_j is defined in the statement of the theorem.

Now let $\{\delta_{j,i}, v_{j,i}\}_{i=1}^n$ be the eigenvalues and eigenvectors of D_j and define

$$E_j = \left(\tilde{\xi}_{j,1}, \dots, \tilde{\xi}_{j,n} \right)' = U_j \tilde{z},$$

where $U_j = \{v_{j,1}, \dots, v_{j,n}\}$, so that we can write

$$\begin{aligned} Q_{1,n}(c) = & 1 - c \sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 + \frac{c^2}{2} \left[2 \left(\sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 \right)^2 - \sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{2,i}^2 \right] \\ & - \frac{c^3}{6} \left[6 \left(\sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 \right)^3 - 3 \left(\sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 \right) \left(\sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{2,i}^2 \right) + \sum_{i=1}^n \delta_{3,i} \tilde{\xi}_{3,i}^2 \right] \\ & + O_p(c^4). \end{aligned} \quad (S7)$$

For $Q_{0,n}(c)$, and similar to (S6), we also have

$$\sum_{i=1}^n \frac{d^j \lambda_i^{-1}}{dc^j} \Big|_{c=0} \tilde{\xi}_i^2 = \tilde{z}' \frac{d^j A^{-1}}{dc^j} \Big|_{c=0} \tilde{z},$$

so that if we successively differentiate the identity $A^{-1}A = I$ (noting that all derivatives of A are symmetric) then we obtain,

$$\begin{aligned} \frac{dA^{-1}}{dc} A + \frac{dA}{dc} A^{-1} &= 0, \\ \frac{d^2 A^{-1}}{dc^2} A + 2 \frac{dA^{-1}}{dc} \frac{dA}{dc} + A^{-1} \frac{d^2 A}{dc^2} &= 0, \\ \frac{d^3 A^{-1}}{dc^3} A + 3 \frac{d^2 A^{-1}}{dc^2} \frac{dA}{dc} + 3 \frac{dA^{-1}}{dc} \frac{d^2 A}{dc^2} + A^{-1} \frac{d^3 A}{dc^3} &= 0. \end{aligned}$$

Evaluating these derivatives at $c = 0$ and noting the definition of D_j , then we find

$$\begin{aligned} \frac{dA^{-1}}{dc} \Big|_{c=0} &= -D_1 \\ \frac{d^2 A^{-1}}{dc^2} \Big|_{c=0} &= 2(D_1)^2 - D_2 \\ \frac{d^3 A^{-1}}{dc^3} \Big|_{c=0} &= -6(D_1)^3 + 6D_1 D_2 - D_3. \end{aligned}$$

As a consequence, and letting $\{\delta_{12,i}, v_{12,i}\}_{i=1}^n$ be the eigenvalues and eigenvectors of the matrix $D_1 D_2$, we then have the following stochastic expansion for $Q_{0,n}(c)$,

$$\begin{aligned} Q_{0,n}(c) = & 1 - c \sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 + \frac{c^2}{2} \sum_{i=1}^n \left(2\delta_{1,i}^2 \tilde{\xi}_{1,i}^2 - \delta_{2,i} \tilde{\xi}_{2,i}^2 \right) \\ & + \frac{c^3}{6} \sum_{i=1}^n \left(-6\delta_{1,i}^3 \tilde{\xi}_{1,i}^2 + 6\delta_{12,i} \tilde{\xi}_{12,i}^2 - \delta_{3,i} \tilde{\xi}_{3,i}^2 \right) + O_p(c^4), \end{aligned} \quad (S8)$$

where the $\tilde{\xi}_{12,i}$ are elements of the vector $E_{12} = U_{12}\tilde{z}$ and $U_{12} = \{v_{12,1}, \dots, v_{12,n}\}$.

Directly subtracting (S7) from (S8) we get

$$\begin{aligned}
Q_{1,n}(c) - Q_{0,n}(c) &= c^2 \left[\left(\sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 \right)^2 - \sum_{i=1}^n \delta_{1,i}^2 \tilde{\xi}_{1,i}^2 \right] \\
&\quad + c^3 \left[\sum_{i=1}^n \delta_{1,i}^3 \tilde{\xi}_{1,i} - \left(\sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 \right)^3 \right] + \\
&\quad \frac{1}{2} \left(\sum_{i=1}^n \delta_{1,i} \tilde{\xi}_{1,i}^2 \right) \left(\sum_{i=1}^n \delta_{2,i} \tilde{\xi}_{2,i}^2 \right) - \sum_{i=1}^n \delta_{12,i} \tilde{\xi}_{12,i}^2 \quad (S9) \\
&\quad + O_p(c^4).
\end{aligned}$$

Taking the limit as $n \rightarrow \infty$ of (S9) with

$$\sqrt{n} \tilde{\xi}_{1,i}^2 \rightarrow_p \xi_{1,i} \sim iidN(0, 1), \quad \text{for } i = 1, \dots, n,$$

since $n^{-1} \sum_{i=1}^n z_i^2 \rightarrow_p 1$, and analogous results for $\tilde{\xi}_{2,i}^2$ and $\tilde{\xi}_{12,i}^2$ then establishes part (ii) of the Lemma.

Proof of Lemma 2:

For brevity put $Q_0 = Q_0(c)$ and $Q_1 = Q_1(c)$ and let $\bar{Q} = c^{-2k} [Q_0 - Q_1]$. Using the characterization of the asymptotic power envelope derived in Lemma 1, we have

$$\begin{aligned}
\Pi_\alpha(c) &= \Pr[Q_1 < \kappa_\alpha] = \Pr[Q_0 + (Q_0 - Q_1) < \kappa_\alpha] = \Pr[Q_0 - c^{2k} \bar{Q} < \kappa_\alpha] \\
&= \Pr[\bar{Q} > c^{-2k} (Q_0 - \kappa_\alpha)].
\end{aligned}$$

Now let $f(q_0, \bar{q})$ denote the joint density of (Q_0, \bar{Q}) and let $f_{Q_0}(q_0)$ and $f_{\bar{Q}}(\bar{q})$ denote the marginal densities of Q_0 and \bar{Q} , so that

$$\begin{aligned}
\Pi_\alpha(c) &= \int_0^\infty \int_{c^{-2k}(q_0 - \kappa_\alpha)}^\infty f(q_0, \bar{q}) d\bar{q} dq_0 \\
&\quad \left(\int_0^{\kappa_\alpha} + \int_{\kappa_\alpha}^{\kappa_\alpha + \varepsilon} + \int_{\kappa_\alpha + \varepsilon}^\infty \right) \left(\int_{c^{-2k}(q_0 - \kappa_\alpha)}^\infty f(q_0, \bar{q}) d\bar{q} \right) dq_0. \quad (S10)
\end{aligned}$$

If we take the three (double) integrals in (S10) individually, then we first have

$$\int_0^{\kappa_\alpha} \int_{c^{-k}(q_0 - \kappa_\alpha)}^\infty f(q_0, \bar{q}) d\bar{q} dq_0 \leq \int_0^{\kappa_\alpha} \int_{-\infty}^\infty f(q_0, \bar{q}) d\bar{q} dq_0 = \int_0^{\kappa_\alpha} f_{Q_0}(q_0) dq_0 = \alpha. \quad (S11)$$

The second integral is,

$$\begin{aligned} \int_{\kappa_\alpha}^{\kappa_\alpha + \varepsilon} \int_{c^{-k}(q_0 - \kappa_\alpha)}^{\infty} f(q_0, \bar{q}) d\bar{q} dq_0 &\leq \int_{\kappa_\alpha}^{\kappa_\alpha + \varepsilon} f_{Q_0}(q_0) dq_0 \\ &\leq \varepsilon \left(\sup_{\kappa_\alpha \leq q_0 \leq \kappa_\alpha + \varepsilon} f_{Q_0}(q_0) \right) \leq \varepsilon M_1 < \infty, \end{aligned} \quad (\text{S12})$$

the latter following since Q_0 is (the limit of) a quadratic form on the unit sphere with finite density, e.g. see Hillier (2001). The final integral satisfies,

$$\begin{aligned} \int_{\kappa_\alpha + \varepsilon}^{\infty} \int_{c^{-2k}(q_0 - \kappa_\alpha)}^{\infty} f(q_0, \bar{q}) d\bar{q} dq_0 &\leq \int_{\kappa_\alpha + \varepsilon}^{\infty} \int_{c^{-2k}\varepsilon}^{\infty} f(q_0, \bar{q}) d\bar{q} dq_0 \leq \int_{c^{-2k}\varepsilon}^{\infty} f(q_0, \bar{q}) d\bar{q} \\ &= \Pr[\bar{Q} > c^{-2k}\varepsilon] \leq \Pr[|\bar{Q}| > c^{-2k}\varepsilon] \\ &= \Pr[e^{2\eta k}|\bar{Q}| \geq e^{2\eta k c^{-2k}\varepsilon}] \\ &\leq \frac{E[e^{2\eta k}|\bar{Q}|]}{e^{2\eta k c^{-2k}\varepsilon}} = \frac{M_\eta(2k|\bar{Q}|)}{e^{2\eta k c^{-2k}\varepsilon}}, \end{aligned} \quad (\text{S13})$$

where $M_\eta(2k|\bar{Q}|)$ is the moment generating function of $2k|\bar{Q}|$. Both Q_1 and Q_2 are bounded above by $\max_{i \in \mathbb{Z}} \lambda_i^{-1}$ and below by $\min_{i \in \mathbb{N}} \lambda_i^{-1}$. Since λ_i is continuous and differentiable on $[0, c]$ then so also are $\max_{i \in \mathbb{N}} \lambda_i^{-1}$ and $\min_{i \in \mathbb{N}} \lambda_i^{-1}$, both therefore permit a mean value expansion and are bounded on $[0, c]$. Consequently, $|\bar{Q}|$ is bounded, hence $M_\eta(2k|\bar{Q}|) \leq M_2 < \infty$. Adding (S11), (S12) and (S13) we have,

$$\Pi_\alpha(c) \leq \alpha + \varepsilon M_1 + \frac{M_2}{e^{2\eta k c^{-2k}\varepsilon}}.$$

and following the proof of Lemma 3 in Bykhovskaya and Phillips (2017), we can put $\varepsilon = -c^k(\ln c) > 0$, giving

$$\Pi_\alpha(c) \leq \alpha + O(c^{2k - \varpi_k}),$$

for some, arbitrarily small, $\varpi_k > 0$, as required.

Tables

The following tables report the outcomes of experiments detailed in the discussion and conclusion of the main paper. All reported values are based on simulations involving 100000 Monte Carlo replications.

Table 1a: Power Envelopes, with a constant and various trends, for testing $H_0 : c = 0$ vs. H_1^S and H_1^E in model (S1), $T = 250$.

d_t c	β_1	$\beta_1 + \beta_2 \ln t$	$\beta_1 + \beta_2 \sqrt{t}$	$\beta_1 + \beta_2 t$	$\beta_1 + \beta_2 t^2$	$\beta_1 + \beta_2 e^t$
1.25	0.361	0.351	0.332	0.057	0.076	0.356
1.00	0.274	0.262	0.246	0.052	0.074	0.269
0.75	0.197	0.188	0.181	0.051	0.067	0.194
0.50	0.130	0.125	0.127	0.050	0.065	0.129
0.25	0.085	0.079	0.083	0.050	0.054	0.084
-1.0	0.079	0.066	0.055	0.052	0.063	0.078
-2.0	0.121	0.083	0.065	0.061	0.076	0.117
-3.0	0.172	0.114	0.082	0.074	0.084	0.168
-4.0	0.233	0.134	0.097	0.092	0.108	0.225
-5.0	0.294	0.156	0.122	0.113	0.141	0.286

Table 1b: Approximations to $O_p(c^3)$ from Lemma 1 of the Power Envelopes
 $H_0 : c = 0$ vs. H_1^S and H_1^E in model (S1), $T = 250$.

d_t c	β_1	$\beta_1 + \beta_2 \ln t$	$\beta_1 + \beta_2 \sqrt{t}$	$\beta_1 + \beta_2 t$	$\beta_1 + \beta_2 t^2$	$\beta_1 + \beta_2 e^t$
1.25	0.331	0.306	0.309	0.052	0.074	0.313
1.00	0.246	0.230	0.237	0.052	0.072	0.241
0.75	0.184	0.168	0.183	0.051	0.066	0.181
0.50	0.128	0.121	0.121	0.050	0.061	0.127
0.25	0.083	0.076	0.080	0.050	0.053	0.083
-1.0	0.077	0.057	0.057	0.050	0.053	0.076
-2.0	0.118	0.063	0.067	0.051	0.058	0.127
-3.0	0.183	0.077	0.101	0.051	0.061	0.178
-4.0	0.244	0.114	0.128	0.053	0.074	0.255
-5.0	0.332	0.185	0.203	0.056	0.099	0.307

Table 2: Ratio of the trend and constant Power Envelopes
for testing $H_0 : c = 0$ vs. H_1^S and H_1^E in model (S1).

α	0.01	0.01	0.01	0.05	0.05	0.05	0.10	0.10	0.10
c T	50	250	500	50	250	500	50	250	500
1.25	0.063	0.060	0.054	0.161	0.158	0.156	0.254	0.241	0.248
1.00	0.080	0.071	0.068	0.191	0.188	0.183	0.286	0.285	0.283
0.75	0.141	0.128	0.110	0.273	0.261	0.254	0.371	0.368	0.360
0.50	0.262	0.221	0.212	0.405	0.386	0.381	0.502	0.485	0.484
0.25	0.515	0.459	0.422	0.622	0.581	0.580	0.697	0.668	0.663
-1.0	0.394	0.319	0.315	0.440	0.371	0.362	0.478	0.421	0.409
-2.0	0.425	0.352	0.347	0.448	0.390	0.381	0.484	0.427	0.417
-3.0	0.455	0.409	0.396	0.479	0.434	0.436	0.501	0.466	0.457
-4.0	0.503	0.521	0.502	0.531	0.550	0.506	0.553	0.529	0.527
-5.0	0.655	0.648	0.632	0.677	0.689	0.694	0.689	0.672	0.672

Table 3: Ratio of the trend and constant Power Envelopesfor testing $H_0 : c = 0$ vs. H_1^F in model (S1).

α	0.01	0.01	0.01	0.05	0.05	0.05	0.10	0.10	0.10
c T	50	250	500	50	250	500	50	250	500
1.25	0.125	0.114	0.101	0.247	0.246	0.235	0.351	0.345	0.343
1.00	0.183	0.162	0.140	0.320	0.306	0.287	0.422	0.411	0.393
0.75	0.260	0.233	0.219	0.407	0.378	0.382	0.504	0.477	0.485
0.50	0.382	0.346	0.337	0.535	0.514	0.500	0.622	0.605	0.594
0.25	0.627	0.590	0.606	0.738	0.704	0.712	0.781	0.776	0.781
-1.0	0.787	0.797	0.796	0.765	0.768	0.774	0.740	0.778	0.783
-2.0	0.613	0.621	0.625	0.596	0.639	0.634	0.586	0.630	0.630
-3.0	0.544	0.541	0.545	0.537	0.525	0.526	0.511	0.521	0.524
-4.0	0.455	0.468	0.472	0.488	0.474	0.468	0.454	0.465	0.459
-5.0	0.411	0.419	0.426	0.428	0.408	0.415	0.440	0.405	0.407

Additional References

- Bykhovskaya, A. and P.C.B. Phillips 2017. Point optimal testing with roots that are functionally local to unity. *Cowles Foundation discussion paper, 3007*.
- Chang, Y. and J.Y. Park 2002. On the asymptotics of ADF tests for unit roots. *Econometric Reviews* **21**: 431-447.
- Hillier, G. 2001. The Density of a Quadratic Form in a Vector Uniformly Distributed on the n-Sphere. *Econometric Theory*, **17**: 1-28.
- Magnus, J.R. and Neudecker, H. 1999. *Matrix Differential Calculus, with Applications in Statistics and Econometrics*. Chichester: Wiley.
- Marsh, P. 2007b. Constructing Optimal tests on a Lagged dependent variable. *J.T.S.A.* **28**: 723-743.
- Podivinsky, J. and M.L. King 2000. The exact power envelope of tests for a unit root. *Discussion Papers in Economics and Econometrics*: University of Southampton.