*Econometric Theory*, **19**, 2003, 240–253. Printed in the United States of America.
DOI: 10.1017/S026646660319202X

# POWER FUNCTIONS AND ENVELOPES FOR UNIT ROOT TESTS

TED JUHL
*University of Kansas*

ZHIJIE XIAO
*University of Illinois at Urbana-Champaign*

This paper studies power functions and envelopes for covariate augmented unit root tests. The power functions are calculated by integrating the characteristic function, allowing accurate evaluation of the power envelope and the power functions. Using the power functions, we study the selection among point optimal invariant unit root tests. An "optimal" point optimal test is proposed based on minimizing the integrated power difference. We find that when there are covariate effects, optimal tests use a local alternative where the power envelope has an approximate value of 0.75.

## 1. INTRODUCTION

Testing for the presence of unit roots is now a common practice in empirical macroeconomic and financial time series analysis. Although the unit root hypothesis has been tested in hundreds of time series, it is widely recognized that the discriminatory power of unit root tests is generally low. In addition, oftentimes practitioners do not know whether a trend is deterministic or stochastic, so it is necessary in empirical analysis to include deterministic terms in the model when estimating the hypothesized unit root parameter. This inclusion of "extra" deterministic components in the model reduces power from its already low level.

One of the mechanisms for increasing the power of unit root tests is related to point optimal testing procedures. The point optimal procedures, which are optimal only for a specific point alternative, can also be used to test for a unit root against more complex alternative hypotheses, although it is no longer optimal for alternatives at other points. Actually, a point optimal test based on an "appropriately chosen" value of the local parameter may have higher power than the conventional unit root tests. Because there is no uniformly most powerful test for the unit root hypotheses, a point optimal test using some plausible

We thank Pentti Saikkonen and two referees for helpful comments. Address correspondence to Ted Juhl, Department of Economics, University of Kansas, 213 Summerfield Hall, Lawrence, KS 66045-2113, USA; e-mail: juhl@ukans.edu; and Zhijie Xiao, Department of Economics, University of Illinois, 1206 S. Sixth St., Champaign, IL 61820, USA; e-mail: zxiao@uiuc.edu.

 © 2003 Cambridge University Press 0266-4666/03 $12.00

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

choice of local parameter value is a reasonable choice. Notice that a point optimal test will touch the power envelope at the corresponding local parameter value, so that there is an infinitely large class of tests, all of which are asymptotically admissible. A natural question to ask is: How should we choose an "optimal" point optimal test? We attempt to address this issue for a class of unit root tests in the current paper.

When ordinary least squares (OLS) is used (with no covariates), point optimal unit root tests have been studied by Elliott, Rothenberg, and Stock (1996) and Saikkonen and Luukkonen (1993). Elliot et al. (1996) show that the power of point optimal tests is not sensitive to the choice of the alternative so long as the point optimal test has a tangency between power 0.25 and 0.75. Elliott and Jansson (2000) propose point optimal invariant tests when covariate effects are estimated. We show that there are interesting differences that arise when using covariates so that the choice of a point optimal invariant test magnifies the familiar trade-off between alternatives close to the null and more distant alternatives. In this paper, we derive characteristic functions for covariate augmented unit root tests to accurately evaluate power envelopes and power functions for point optimal tests. Using power envelopes and power functions calculated from our characteristic functions, one can evaluate differences between point optimal tests and the envelope. This allows, perhaps, formulation of an "optimal" point optimal test.

The paper is organized as follows. In Section 2, we present the asymptotic power functions derived for invariant unit root tests developed by Elliott and Jansson (2000). Section 3 provides the characteristic functions for the power functions. Section 4 studies the selection of optimal point optimal unit root tests and reports new critical values used for a feasible version of the test. Section 5 concludes.

## 2. UNIT ROOT TESTS USING COVARIATES

Elliott and Jansson (2000) extend the covariate augmented unit root tests developed by Hansen (1995) to form point optimal tests invariant to deterministic terms. Because our focus is on the evaluation of the asymptotic power curves and envelopes, we take the asymptotic distributions as given and proceed with a strategy for evaluation.

Suppose that a time series $y_t$ can be decomposed into a deterministic component and a stochastic component, $y_t = d_t + s_t$, where $d_t = \gamma' z_t$, with $z_t$ a deterministic trend of known form. The stochastic part is represented as $\Delta s_t = \delta s_{t-1} + v_t$. We also assume $\delta = -(c/T)$. Suppose we wish to test the null hypothesis $H_0: c = 0$ versus $H_{\bar{c}}: c = \bar{c}$ where $\bar{c} \geq 0$. The limiting distributions of the likelihood ratio test depend on the parameters $c$, $\bar{c}$, and $R^2$. The parameter $c$ is the true value of the local parameter, whereas $\bar{c}$ is the hypothesized local parameter. The term $R^2$ indicates the correlation between covariates and the quasi-difference of the local to unit root process (for a more detailed description of

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

the unit root model with covariates, see Hansen, 1995; Elliott and Jansson, 2000). We consider three cases with emphasis on the two empirically relevant cases where a constant is estimated and a constant and trend are estimated. In particular, case 1 represents the distribution when there are no deterministic components estimated or when they are treated as known. In case 2, an intercept is included for both the covariates and the local to unit root process, whereas case 3 includes a constant and trend in both. Our cases 1, 2, and 3 correspond to the cases 1, 3, and 5, respectively, in Elliott and Jansson (2000). Denoting the limiting distributions corresponding to these three cases as $LR_j(c,\bar{c})$, $j = 1, 2, 3$, the limiting distributions are as follows:

$$LR_1(c,\bar{c}) = (\bar{c}^2 - 2c\bar{c}) \int (W_1^c)^2 + 2\bar{c} \int W_1^c \, dW_1 + (\bar{c}^2 - 2c\bar{c}) \left( \frac{R^2}{1-R^2} \right)$$
$$\times \int (W_1^c)^2 - 2\bar{c} \frac{R}{\sqrt{1-R^2}} \int W_1^c \, dW_2,$$

$$LR_2(c,\bar{c}) = (\bar{c}^2 - 2c\bar{c}) \int (W_1^c)^2 + 2\bar{c} \int W_1^c \, dW_1 + (\bar{c}^2 - 2c\bar{c}) \left( \frac{R^2}{1-R^2} \right)$$
$$\times \int (W_1^c - \overline{W}_1^c)^2 - 2\bar{c} \frac{R}{\sqrt{1-R^2}} \int (W_1^c - \overline{W}_1^c) \, dW_2,$$

$$LR_3(c,\bar{c}) = (\bar{c}^2 - 2c\bar{c}) \int (W_1^c)^2 + 2\bar{c} \int W_1^c \, dW_1 + (\bar{c}^2 - 2c\bar{c}) \left( \frac{R^2}{1-R^2} \right)$$
$$\times \int (W_1^{c\tau})^2 + W_1^c(1)^2 - m_1 \left( (1+\bar{c}) W_1^c(1) + \bar{c}^2 \int s W_1^c \right)^2$$
$$- 2\bar{c} \frac{R}{\sqrt{1-R^2}} \int W_1^{c\tau} \, dW_2,$$

where $m_1 = 1/(1 + \bar{c}^2/3 + \bar{c})$, $W_1^c(r)$ is an Ornstein–Uhlenbeck process satisfying the stochastic differential equation

$$dW_1^c(s) = -cW_1^c(s)\, ds + dW_1(s),$$

and $W_1^{c\tau}(s) = W_1^c(s) + (6s - 4) \int W_1^c(r)\, dr - (12s - 6) \int r W_1^c(r)\, dr$, $\overline{W}_1^c = \int W_1^c(r)\, dr$.[1]

These limiting variates all have nonstandard distributions and are dependent on nuisance parameters. In practice, simulation methods are generally used to obtain approximations of these distributions. However, Nabeya and Tanaka (1990), Perron (1991), and Tanaka (1996) suggest a general approach for deriving the limiting distributions of OLS estimators for autoregressive parameters. We explore using similar methods for the covariate case in the next section.

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

## 3. CHARACTERISTIC FUNCTIONS

We find the characteristic functions of the limiting distributions and use them to find probabilities based on numerical integration. The probabilities are used to find critical values for a range of point alternatives that are then used to find the power envelope.

Evaluating the distributions using numerical integration has certain advantages over the traditional Monte Carlo based approximation. First, as documented in Nabeya and Tanaka (1990), accuracy is improved using numerical methods. For our purposes, this is important as we plan to use the derived power envelopes to find a class of optimal point optimal tests. Second, the computational time decreases substantially. This is important as we must evaluate and compare multiple power curves for point optimal tests.

The characteristic function of the random variable $LR_j(c,\bar{c})$ is defined as $E(e^{i\theta LRj})$, and we denote this by $\phi_{LRj}(\theta,c,\bar{c})$, $j=1,2,3$. We give an expression for the characteristic function in the following theorem.[2]

THEOREM 3.1.

$$\phi_{LRj}(\theta;c,\bar{c}) = e^{((c/2)-i\theta\bar{c})} \times g_j(\theta)^{-1/2} \quad j=1,2,3$$

*with*

$$g_1(\theta) = (c-2\bar{c}\theta i)\,\frac{\sin\lambda}{\lambda} + \cos\lambda,$$

$$\begin{aligned} g_2(\theta) = {} & (4c\bar{c}^2\theta i - 8c^2\bar{c}\theta i + 8\bar{c}^3\theta^2 - 16\bar{c}^3\theta^3 i - 24c\bar{c}^2\theta^2)r\,\frac{1}{\lambda^4} \\ & + (24c\bar{c}^2\theta^2 + 16\bar{c}^3\theta^3 i - 8\bar{c}^3\theta^2 + 8c^2\bar{c}\theta i - 4c\bar{c}^2\theta i)r\,\frac{\cos\lambda}{\lambda^4} \\ & + (2\bar{c}^2\theta i - 4c\bar{c}\theta i - 2c\bar{c}^2\theta i + 4c^2\bar{c}\theta i - 4\bar{c}^2\theta^2 - 4\bar{c}^3\theta^2 \\ & \quad + 8\bar{c}^3\theta^3 i + 12c\bar{c}^2\theta^2)r\,\frac{\sin\lambda}{\lambda^3} \\ & + (-2\bar{c}^2\theta i + 4c\bar{c}\theta i + 4\bar{c}^2\theta^2)r\,\frac{\cos\lambda}{\lambda^2} \\ & + (c-2\bar{c}\theta i)\,\frac{\sin\lambda}{\lambda} + \cos\lambda, \end{aligned}$$

$$g_3(\theta) = e^{i\lambda}\,|I-\Sigma\Lambda|,$$

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

*where*

$$\Sigma = \begin{pmatrix} \frac{1-e^{-2i\lambda}}{2i\lambda} & \frac{1}{-\lambda^2}\left(\frac{1}{2}-e^{-i\lambda}+\frac{e^{-2i\lambda}}{2}\right) & \frac{i\lambda-1+(i\lambda+1)e^{-2i\lambda}}{-2i\lambda^3} \\ \frac{1}{-\lambda^2}\left(\frac{1}{2}-e^{-i\lambda}+\frac{e^{-2i\lambda}}{2}\right) & \frac{1}{-i\lambda^3}\left(i\lambda-\frac{3}{2}+2e^{-i\lambda}-\frac{e^{-2i\lambda}}{2}\right) & \frac{-\lambda^2-(i\lambda+1)(1-e^{-i\lambda})^2}{2\lambda^4} \\ \frac{i\lambda-1+(i\lambda+1)e^{-2i\lambda}}{-2i\lambda^3} & \frac{-\lambda^2-(i\lambda+1)(1-e^{-i\lambda})^2}{2\lambda^4} & \frac{-2i\lambda^3+3\lambda^2+3-3(i\lambda+1)^2e^{-2i\lambda}}{6i\lambda^5} \end{pmatrix},$$

$$\Lambda = \begin{pmatrix} i\lambda - c + 2i\theta\bar{c} + 2i\theta - 2i\theta m_1(1+\bar{c})^2 & 0 & -2i\theta m_1(1+\bar{c})\bar{c}^2 \\ 0 & -8m_2 & 12m_2 \\ -2i\theta m_1(1+\bar{c})\bar{c}^2 & 12m_2 & -2i\theta m_1\bar{c}^4 - 24m_2 \end{pmatrix},$$

$$\lambda = \sqrt{2\theta\bar{c}^2 i + 2\theta(\bar{c}^2 - 2c\bar{c})ri - 4\bar{c}^2\theta^2 r - c^2},$$

$$m_1 = 1/(1 + \bar{c}^2/3 + \bar{c}),$$

$$m_2 = i\theta(\bar{c}^2 - 2c\bar{c})r - 2\theta^2\bar{c}^2 r,$$

$$r = \frac{R^2}{1-R^2}.$$

Once we have obtained the characteristic function, calculating probabilities is based on numerical integration using Gurland's (1948) formula:

$$P(LR_j(c,\bar{c}) \le x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{1}{\theta}\,\mathrm{Im}(e^{-i\theta x}\phi_{LRj}(\theta;c,\bar{c}))\,d\theta, \tag{3.1}$$

where Im is the imaginary part of the complex number.

Using the preceding formula, it is possible to calculate critical values for specific alternatives, which can then be used to obtain power envelopes and power functions for point optimal tests of the null hypothesis, $H_0: c = 0$. We compare these point optimal tests for several cases of the nuisance parameter $R^2$ in the next section.

The power envelope was calculated by first finding critical values for the test indexed by $\bar{c}$. That is, a critical value must be found for each alternative hypothesis and each value of the nuisance parameter $R^2$ considered. After critical values are generated for each of the alternatives, the power is calculated by numerical integration with $x = x_{0.05}$ so that $P(LR_j(0,\bar{c}) \le x_{0.05}) = 0.05$. The power functions are denoted

$$\Pi(c,\bar{c}) = P(LR_j(c,\bar{c}) \le x_{0.05})$$

so that the power envelope is $\Pi(\bar{c},\bar{c})$. For the numerical integration using characteristic functions, we found 2,500 values of the characteristic function and used a simple application of Simpson's rule to find the integral. The truncation of the integral was determined where the integrand consistently took a value of

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

less than $10^{-20}$ in absolute value. The power envelopes for cases 2 and 3 appear in Figures 1 and 2 and correspond to Figures 1b and 1d in Elliott and Jansson (2000).

## 4. OPTIMAL POINT OPTIMAL UNIT ROOT TESTS

A point optimal test for $H_0: c = 0$ against $H_{\bar{c}}: c = \bar{c} > 0$ also naturally provides a test for alternative $H_c: c > 0$. Although, strictly speaking, such a test is only optimal for the specified alternative, a point optimal test with an *appropriate* selection of $\bar{c}$ can be nearly optimal in the sense that the power function is very close to the power envelope. Because there is no uniformly most powerful test, point optimal tests (against the alternative of $\bar{c}$) are of interest and may provide a power gain relative to tests based simply on estimated autoregressive parameters.

Given the different behavior of tests based on different values of $\bar{c}$, an obvious question is which of these tests has the best overall power properties. As suggested in King (1988), one possibility is to choose $\bar{c}$ so that the test is more powerful over certain ranges of the parameter space. A natural criterion for measuring the closeness of a power curve to the power envelope is the integral of the difference between a power curve and the power envelope, so that we

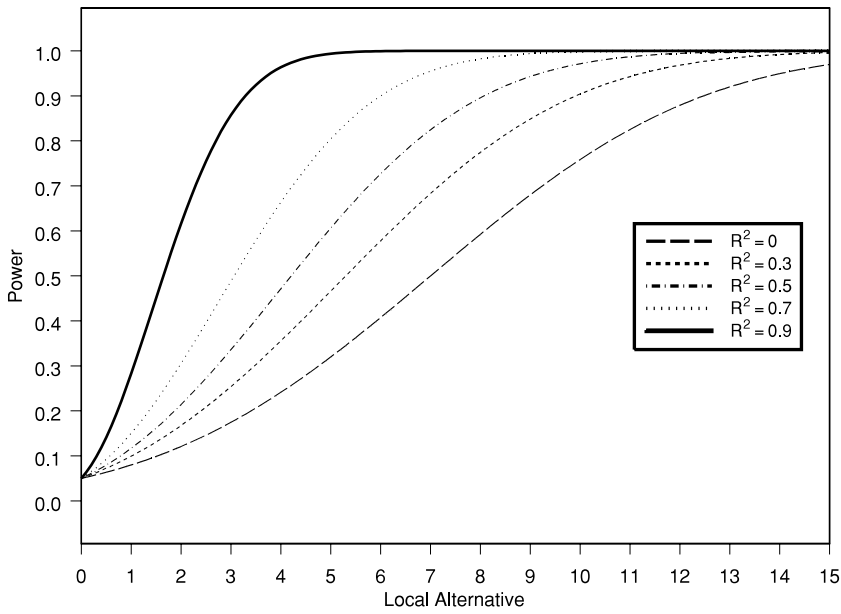



**FIGURE 1.** Case 2 envelopes.

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

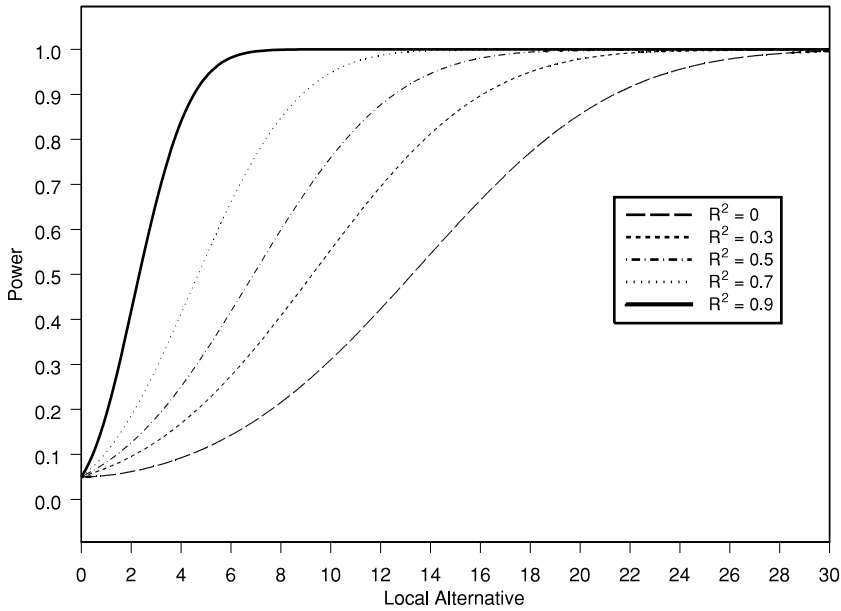



**FIGURE 2.** Case 3 envelopes.

may select the optimal point optimal test by minimizing the integrated power differences; i.e., we choose $\bar{c}$ by

$$\min_{\bar{c}} \int_0^\infty [\Pi(c,c) - \Pi(c,\bar{c})]\, dc. \tag{4.1}$$

Because the power envelope is above the power function $\Pi(c,\bar{c})$, the difference between $\Pi(c,c)$ and $\Pi(c,\bar{c})$ is always nonnegative. Notice that criterion (4.1) is equivalent to maximizing the power function $\Pi(c,\bar{c})$ integrated over $c$ and the suggested criterion can be treated as an application of Cox and Hinkley (1974) to the unit root model.[3] For general hypothesis tests against composite alternatives, Cox and Hinkley (1974) suggest selecting an alternative point that maximizes a weighted power average. Our criterion chooses the Lebesgue measure (corresponding to their $k(\theta)$ in equation (26), p. 102) and the power function of covariate augmented unit root tests.

We apply the results developed in Section 3 to select an optimal point optimal test. The optimal choices of $\bar{c}$ (and thus the corresponding test) are found for representative $R^2$ values. The integral is evaluated over the range from zero to 30 because the difference between the point optimal test and the envelope is

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

negligible at this point. Point optimal tests indexed by $\bar{c}$ from .01 to 20 are considered because the integral is decreasing and then increasing in $\bar{c}$. The integral is evaluated using Simpson's rule with 3,000 points, starting at 0.01 and 0.01 increments. The integral is evaluated for $R^2 = 0.1, \ldots, 0.9$.

The output from searching for a class of optimal $\bar{c}$ is given in Table 1. First, we point out that the minimized integral of the difference between the power envelope and the power function of the point optimal test initially increases

**TABLE 1.** Optimal $\bar{c}$ and critical values for different $R^2$

| $R^2$ | Optimal $\bar{c}$ | Critical value | $\int_0^\infty [\Pi(c,c) - \Pi(c,\bar{c})]\,dc$ |
|---|---|---|---|
| | | Case 1 | |
| 0.9 | 1.81 | 2.3686 | 0.1198 |
| 0.8 | 2.88 | 2.5714 | 0.1744 |
| 0.7 | 3.84 | 2.7362 | 0.2101 |
| 0.6 | 4.77 | 2.8717 | 0.2304 |
| 0.5 | 5.65 | 3.0023 | 0.2363 |
| 0.4 | 6.55 | 3.1110 | 0.2274 |
| 0.3 | 7.42 | 3.2190 | 0.2032 |
| 0.2 | 8.32 | 3.3116 | 0.1625 |
| 0.1 | 9.20 | 3.4049 | 0.1034 |
| | | Case 2 | |
| 0.9 | 2.45 | −0.0441 | 0.0815 |
| 0.8 | 3.66 | 1.0267 | 0.1131 |
| 0.7 | 4.59 | 1.8035 | 0.1348 |
| 0.6 | 5.54 | 2.6641 | 0.1483 |
| 0.5 | 6.40 | 3.4255 | 0.1542 |
| 0.4 | 7.15 | 4.0574 | 0.1518 |
| 0.3 | 7.88 | 4.6768 | 0.1399 |
| 0.2 | 8.60 | 5.2921 | 0.1169 |
| 0.1 | 9.45 | 6.0756 | 0.0795 |
| | | Case 3 | |
| 0.9 | 3.33 | −1.0960 | 0.0929 |
| 0.8 | 5.11 | −0.0383 | 0.1558 |
| 0.7 | 6.80 | 1.1300 | 0.2096 |
| 0.6 | 8.24 | 2.1255 | 0.2509 |
| 0.5 | 9.67 | 3.1675 | 0.2785 |
| 0.4 | 11.15 | 4.0179 | 0.2881 |
| 0.3 | 12.55 | 5.1058 | 0.2753 |
| 0.2 | 13.99 | 6.5002 | 0.2324 |
| 0.1 | 15.46 | 7.4248 | 0.1518 |

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

and eventually decreases as $R^2$ increases. In fact, when $R^2$ is nearly zero, the differences in the power envelope and the point optimal test are trivial. This is in agreement with the simulation evidence in Elliott et al. (1996) and the numerical evidence presented in Tanaka (1996). In addition, we found that the point optimal tests were tangent to the power envelope at approximately 0.75 regardless of the value of $R^2$. For a wide range of $R^2$ (except close to zero), these different choices of $\bar{c}$ are important. However, this is not of practical importance when $R^2 = 0$ (or close to 0) as the difference in point optimal tests is very small.

To illustrate the choice of tests, we plot the difference between the power envelope and several cases of point optimal invariant tests for the empirically relevant cases. Figure 3 presents the power differences for case 2 when $R^2 = 0.8$. The optimal point optimal invariant test is denoted OPO and is compared with the point optimal invariant tests associated with power at 0.5 and 0.25. In addition, the case where $\bar{c} = 7$ was suggested by Elliott and Jansson (2000) and is included for comparison.[4] None of the tests has the lowest power loss over the range of alternatives because each test is asymptotically admissible. However, the optimality criteria used in this paper dictate that we should minimize the area under the curve, and the OPO curve corresponds to that test. Figure 4 represents case 2 with $R^2 = 0.9$. The power loss is similar to the $R^2 = 0.8$ case

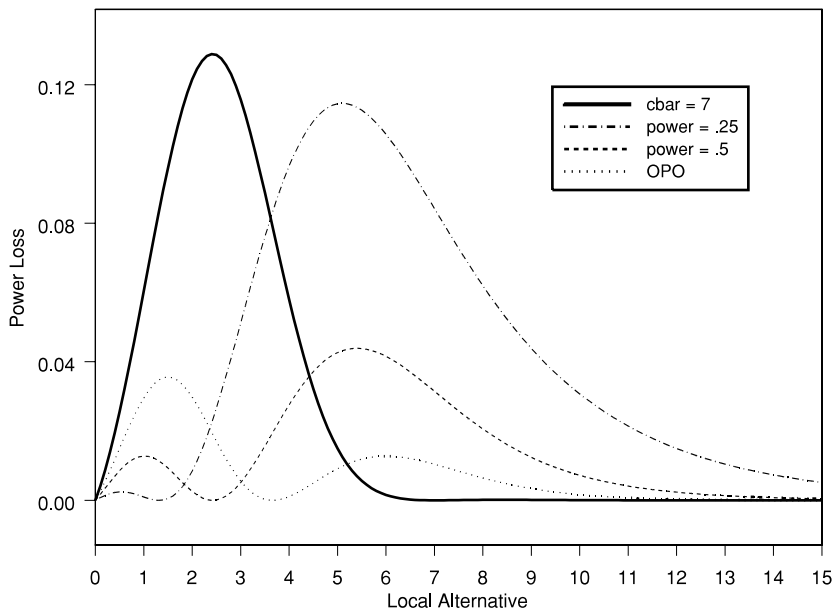



**FIGURE 3.** Case 2: $R^2 = 0.8$.

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

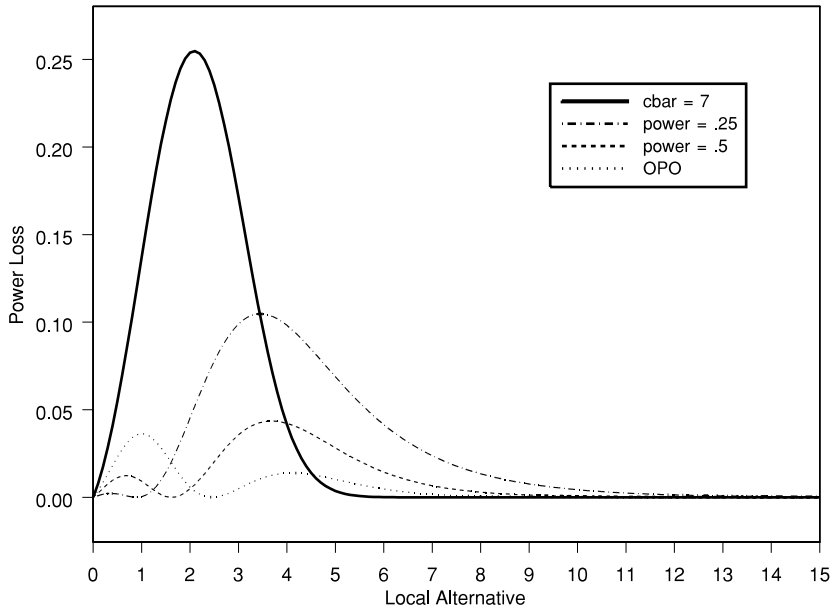



**FIGURE 4.** Case 2: $R^2 = 0.9$.

for all of the tests except when $\bar{c} = 7$. This is, of course, due to the fact that holding $\bar{c} = 7$ now forces the tangency to the envelope to lie at a higher power level because the envelope for $R^2 = 0.9$ is steeper. Hence, the power loss at more local alternatives is greater. Figures 5 and 6 repeat the analysis for case 3, but we include a test based on $\bar{c} = 13.5$ because a constant and trend are estimated.[5] Similar conclusions hold with the power loss associated with $\bar{c} = 13.5$ being even greater. When the point optimal test associated with power $= 0.25$ is examined, we see that the power loss for the middle range of alternatives is substantial, so that the area under the curve is large. This effect is magnified if we consider locally most powerful tests where $\bar{c} \to 0$.[6] In all cases, it is easy to see that the OPO test has the minimum area under the curve.

## 5. CONCLUSION

We have developed expressions for characteristic functions for the limiting distribution of point optimal invariant tests for the unit root hypothesis using covariates. Using these characteristic functions, we are able to find power envelopes and power functions by numerical integration in the manner proposed in Nabeya and Tanaka (1990) and Perron (1991) for OLS based unit root tests.

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

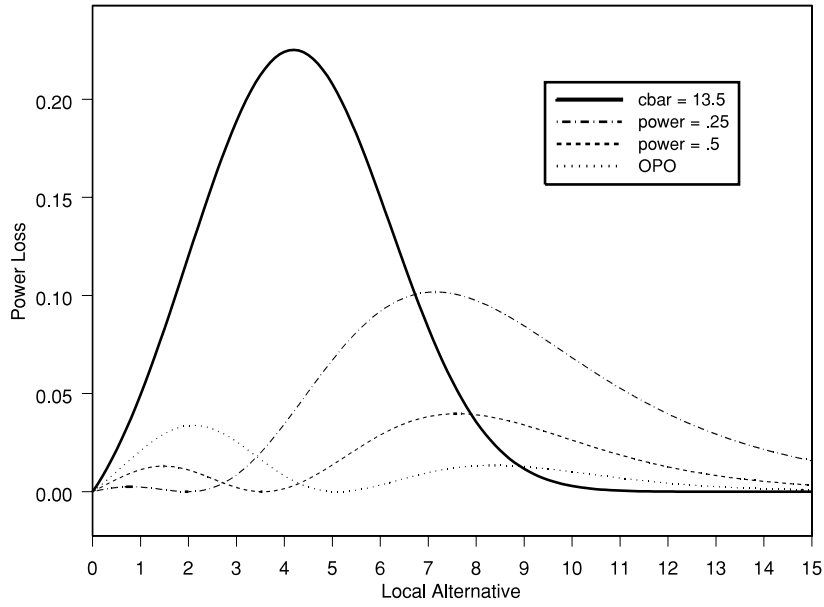



**FIGURE 5.** Case 3: $R^2 = 0.8$.

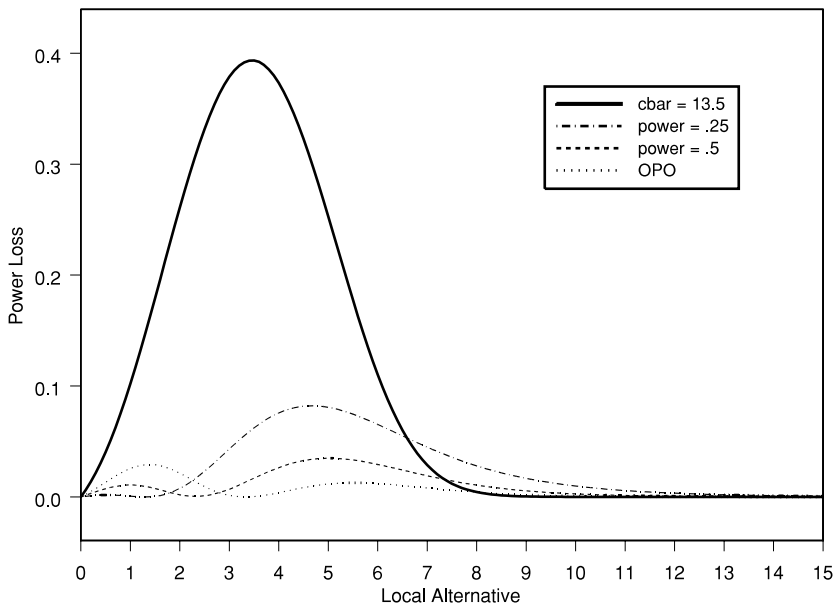



**FIGURE 6.** Case 3: $R^2 = 0.9$.

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

Employing the power envelopes and power functions for point optimal tests, we find the "optimal" point optimal tests, in the sense that the integrated difference between the envelope and the point optimal tests is minimized. Remarkably, the tangency occurs where the power envelope takes a value of approximately 0.75 regardless of the magnitude of the covariate effect. The power gains from using these optimal tests (relative to the choice of alternative when there is no covariate effect) are greater as the covariate is more effective in explaining variation. The values of $\bar{c}$ and the critical values associated with the optimal point optimal invariant tests can be used in conjunction with Elliott and Jansson (2000).

## NOTES

1. The parameters $c$ and $\bar{c}$ are taken as positive, which is the opposite of Elliott and Jansson (2000) but the same as in Hansen (1995).

2. Case 3 has a representation using $\sin \lambda$ and $\cos \lambda$. However, this representation is several pages long, and we use the determinant representation to conserve space.

3. We thank an anonymous referee for pointing out this connection.

4. When $\bar{c} = 7$, the point optimal test is tangent to the power envelope at 0.5 if there are no covariate effects. This is the same alternative used in Elliott et al. (1996).

5. For case 3, $\bar{c} = 13.5$ is tangent to the power envelope at power $= 0.5$ when there is no covariate effect and a constant and trend are estimated.

6. See Figure 1 in Elliott et al. (1996) for a comparison of tests when $R^2 = 0$.

## REFERENCES

Cox, D.R. & D.V. Hinkley (1974) *Theoretical Statistics*. London: Chapman and Hall.

Elliott, G., & M. Jansson (2000) Testing for Unit Roots with Stationary Covariates. Working paper, University of California, San Diego.

Elliott, G., T.J. Rothenberg, & J.H. Stock (1966) Efficient tests for an autoregressive unit root. *Econometrica* 64, 813–836.

Girsanov, I.V. (1960) On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability and Its Applications*, 5, 285–301.

Gurland, J. (1948) Inversion formulae for the distribution of ratios. *Annals of Mathematical Statistics* 19, 228–237.

Hansen, B.E. (1995) Rethinking the univariate approach to unit root testing. *Econometric Theory* 11, 1148–1171.

King, M. (1983) Testing for moving average regression disturbances. *Australian Journal of Statistics* 25, 23–34.

King, M. (1988) Towards a theory of point optimal testing. *Econometric Reviews* 6, 169–218.

Nabeya, S., & K. Tanaka (1990) Limiting powers of unit-root tests in time-series regression. *Journal of Econometrics* 46, 246–271.

Nelson, C.R., & C. Plosser (1982) Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics* 10, 139–162.

Perron, P. (1991) A continuous time approximation to the unstable first-order autoregressive process: The case without an intercept. *Econometrica* 59, 211–236.

Saikkonen, P., & R. Luukkonen (1993) Point optimal tests for testing the order of differencing in ARIMA models. *Econometric Theory* 9, 343–362.

Tanaka, K. (1996) *Time Series Analysis*. New York: Wiley.

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

# Appendix

**Proof of Theorem 3.1.** The theorem follows Nabeya and Tanaka (1990) and Perron (1991) closely. We prove the result for case 2 because the other cases are similar. The proof is completed by finding $E(e^{\theta LR_2})$ and then replacing $\theta$ by $i\theta$. First,

$$LR(c,\bar{c}) = (\bar{c}^2 - 2c\bar{c}) \int (W_1^c)^2 + 2\bar{c} \int W_1^c \, dW_1 + (\bar{c}^2 - 2c\bar{c}) \left( \frac{R^2}{1-R^2} \right) \int (W_1^c - \overline{W}_1^c)^2$$

$$- 2\bar{c} \, \frac{R}{\sqrt{1-R^2}} \int (W_1^c - \overline{W}_1^c) \, dW_2.$$

Condition on $W_1^c$ so that we have

$$E(e^{\theta LR}) = E\left[ \exp\left( \theta\mu + \frac{1}{2}\theta^2\sigma^2 \right) \right],$$

where

$$\mu = (\bar{c}^2 - 2c\bar{c}) \int (W_1^c)^2 + 2\bar{c} \int W_1^c \, dW_1 + (\bar{c}^2 - 2c\bar{c}) \, \frac{R^2}{1-R^2} \int (W_1^c - \overline{W}_1^c)^2$$

and

$$\sigma^2 = 4\bar{c}^2 \, \frac{R^2}{1-R^2} \int (W_1^c - \overline{W}_1^c)^2.$$

Following Perron (1991), use Girsanov's (1960) theorem to eliminate the $\int (W_1^c)^2$ terms, setting

$$\beta = \sqrt{c^2 - 2\theta\bar{c}^2 - 2\theta(\bar{c}^2 - 2c\bar{c}) \, \frac{R^2}{1-R^2} - 4\bar{c}^2\theta^2 \, \frac{R^2}{1-R^2}}.$$

The moment generating function becomes

$$e^{((c-\beta)/2 - \theta\bar{c})} E\left[ \exp\left\{ \left( \frac{\beta - c}{2} + \theta\bar{c} \right) (W_1^\beta(1))^2 \right.\right.$$

$$\left.\left. - \frac{R^2}{1-R^2} \, (2\theta^2\bar{c}^2 + \theta(\bar{c}^2 - 2c\bar{c})) \left( \int W_1^\beta \right)^2 \right\} \right].$$

Let $X = W_1^\beta(1)$ and $Y = \int W_1^\beta$. Then $V = (X,Y)^\top \sim N(0,\Sigma)$ where

$$\Sigma = \begin{pmatrix} \dfrac{1-e^{-2\beta}}{2\beta} & \dfrac{1}{\beta^2}\left( \dfrac{1}{2} - e^{-\beta} + \dfrac{e^{-2\beta}}{2} \right) \\ \dfrac{1}{\beta^2}\left( \dfrac{1}{2} - e^{-\beta} + \dfrac{e^{-2\beta}}{2} \right) & \dfrac{1}{\beta^3}\left( \beta - \dfrac{3}{2} + 2e^{-\beta} - \dfrac{e^{-2\beta}}{2} \right) \end{pmatrix}$$

(see Tanaka, 1996, p. 234).

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press

Then

$$E(e^{\theta LR}) = \exp\left(\frac{c-\beta}{2} - \theta\bar{c}\right) E\left[\exp\left(\frac{1}{2} V^\top \Lambda V\right)\right],$$

where

$$\Lambda = \begin{pmatrix} (\beta - c) + 2\theta\bar{c} & 0 \\ 0 & -4\theta^2\bar{c}^2 \frac{R^2}{1-R^2} - 2\theta(\bar{c}^2 - 2c\bar{c}) \frac{R^2}{1-R^2} \end{pmatrix}.$$

Then

$$E\left[\exp\left(\frac{1}{2} V^\top \Lambda V\right)\right] = \int \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left(-\frac{1}{2} V^\top (\Sigma^{-1} - \Lambda) V\right) dV.$$

A simple change of variables gives us

$$E(e^{\theta LR}) = \exp\left(\frac{c-\beta}{2} - \theta\bar{c}\right) \times |I - \Sigma\Lambda|^{-1/2}.$$

After algebra, we have

$$E(e^{\theta LR}) = \exp\left(\frac{c}{2} - \theta\bar{c}\right) \times g(\theta)^{-1/2},$$

where

$$\begin{aligned} g(\theta) = {} & (4c\bar{c}^2\theta - 8c^2\bar{c}\theta - 8\bar{c}^3\theta^2 - 16\bar{c}^3\theta^3 + 24c\bar{c}^2\theta^2) r \frac{1}{\beta^4} \\ & + (-24c\bar{c}^2\theta^2 + 16\bar{c}^3\theta^3 + 8\bar{c}^3\theta^2 + 8c^2\bar{c}\theta - 4c\bar{c}^2\theta) r \frac{\cosh\beta}{\beta^4} \\ & + (4c\bar{c}\theta - 2\bar{c}^2\theta + 2c\bar{c}^2\theta - 4c^2\bar{c}\theta - 4\bar{c}^2\theta^2 - 4\bar{c}^3\theta^2 - 8\bar{c}^3\theta^3 + 12c\bar{c}^2\theta^2) \\ & \times r \frac{\sinh\beta}{\beta^3} + (2\bar{c}^2\theta - 4c\bar{c}\theta + 4\bar{c}^2\theta^2) r \frac{\cosh\beta}{\beta^2} \\ & + (c - 2\bar{c}\theta) \frac{\sinh\beta}{\beta} + \cosh\beta. \end{aligned}$$

Substituting $i\theta$ for $\theta$ and noting that $\beta = i\lambda$ gives the result. ■

https://doi.org/10.1017/S026646660319202X Published online by Cambridge University Press