



MONASH University

Australia

Department of Econometrics and Business Statistics

<http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/>

**Point Optimal Testing: A Survey of the Post
1987 Literature**

Maxwell L. King and Sivagowry Srianthakumar

March 2015

Working Paper 05/15

Point Optimal Testing: A Survey of the Post 1987 Literature¹

MAXWELL L. KING

Department of Econometrics & Business Statistics

Monash University

SIVAGOWRY SRINANATHAKUMAR

School of Economics, Finance and Marketing

RMIT University

Abstract

In the absence of uniformly most powerful (UMP) tests or uniformly most powerful invariant (UMPI) tests, King (1987c) suggested the use of Point Optimal (PO) tests, which are most powerful at a chosen point under the alternative hypothesis. This paper surveys the literature and major developments on point optimal testing since 1987 and suggests some areas for future research. Topics include tests for which all nuisance parameters have been eliminated and dealing with nuisance parameters via (i) a weighted average of p values, (ii) approximate point optimal tests, (iii) plugging in estimated parameter values, (iv) using asymptotics and (v) integration. Progress on using point-optimal testing principles for two-sided testing and multi-dimensional alternatives is also reviewed. The paper concludes with thoughts on how best to deal with nuisance parameters under both the null and alternative hypotheses, as well as the development of a new class of point optimal tests for multi-dimensional testing.

JEL: C12; C50

Keywords: Local to unity asymptotics, Neyman-Pearson lemma, nuisance parameters, power envelope, unit root testing.

¹The authors are most grateful for constructive suggestions from Graham Elliott, Jiti Gao and David Harris that have helped improve the paper.

1 Introduction

Constructing hypothesis tests or choosing which test to use in econometrics can be difficult. Sometimes we are lucky and have lots of data observations at our disposal so the choice of test statistic may not be particularly crucial. On the other hand, too often the sample size is relatively small and then we want to use an accurate and powerful test. Because our data does not typically come from a controlled experiment but rather from our best efforts of observing a complicated economy, hypothesis testing has an enhanced role to play in our quest to model selected elements of an economy.

Hypotheses under test can be classified into two types; simple and composite. A simple hypothesis is one in which the observed data comes from a sole distribution with all parameters known. A composite hypothesis is made up of more than one distribution, typically involving parameters that can take a range of values. The main result that helps us construct powerful tests is the Neyman-Pearson lemma (see Lehmann and Romano, 2005, p. 60). It states that the most powerful (MP) test of a simple null hypothesis (H_0) against a simple alternative hypothesis (H_1) is based on rejecting H_0 for large values of the ratio of the H_1 density to the H_0 density. Unfortunately it is very rare that we test a simple null against a simple alternative.

When one moves to testing a simple null against a composite alternative (H_{1c}) then it can be less clear how to proceed. A useful concept for understanding the options is the power envelope. For a given level of significance (say 5%), it can be traced out by calculating the power of the MP test of H_0 against each distribution under H_{1c} . If the distributions under H_{1c} can be indexed by a parameter vector, γ , then the power envelope will be a function of γ . No test can have power above the power envelope. The best outcome is if there exists a test whose power is equal to the power envelope. This is a uniformly most powerful (UMP) test. A second best is to find a test whose power is very close to the power envelope. We might call such a test an approximately UMP test.

In the absence of a UMP test, Cox and Hinkley (1974, p.102) considered three alternative approaches. The first of these is the point optimal test which involves the MP test of H_0 against the simple hypothesis constructed by choosing γ to be “somewhat arbitrarily a ‘typical’ point”, say $\gamma = \gamma_1$. As a test of H_0 against H_{1c} , it is MP at $\gamma = \gamma_1$, or alternatively, its power curve kisses the power envelope at $\gamma = \gamma_1$. If unknowingly a UMP test does exist, then this test will be UMP. A variation on the point optimal test is Davies’ (1969) beta optimal test which maximizes power (beta) at a chosen level, say 0.5 or 0.8. Another is Schaafsma and Smid’s (1966) most stringent somewhere most powerful test which chooses the point at which power is optimized to minimize the maximum difference between the test’s power and the power envelope.

The second option is to remove any arbitrariness by constructing the test which is the limit of the point optimal tests as the chosen point moves towards the H_0 value. When γ is a scalar, this is known as a locally best (or locally MP) test. Its power curve has the steepest slope of all tests as one moves away from H_0 (see Ferguson (1967, p.235) and King and Hillier (1985)). If γ is a vector or its H_0 value is inside the range of γ values (i.e., the two-side case), then the test most likely will differ depending on the direction taken in γ space when forming the limit. If the test is invariant to the direction taken, then we have a uniformly locally best test with a power curve with steepest slope in all directions away from the null (for an example, see King (1987b) and King and Evans (1988)). An alternative solution, when there is no uniformly locally best test, is to construct the test whose power curve slope averaged over all directions in the γ space is maximized. This is known as a locally most mean powerful (LMMP) test (see Sen Gupta and Vermeire (1986) and King and Wu (1997)).

The third approach discussed by Cox and Hinkley (1974) is to choose a test which maximizes some weighted average of power. The LMMP test is a particular example of this approach and Andrews and Ploberger (1994) provide another prominent econometric example. In general, the test which maximizes a weighted average of power can be constructed using a special case of the generalized

Neyman-Pearson lemma (GNPL) (see Lehmann and Romano, 2005, p.77 and Begum and King, 2005a). Essentially H_{1c} is replaced by a weighted average of the densities to make a simple alternative and the test is the MP test of H_0 against this new alternative. It is worth noting that a point optimal test can be viewed as a test which maximizes weighted average power; this case being where all the weight is put on the point at which power is optimized.

Things become even more complicated as we move to testing a composite null (H_{0c}) against either a simple (H_1) or a composite (H_{1c}) alternative. The GNPL does provide some options in some rather special cases. If H_{0c} is made up of a finite number of completely determined densities and we are testing against a simple alternative, the GNPL provides the most powerful test if such a test exists. If the alternative is composite, then clearly it can also provide the point optimal solution or the maximized weighted power solution. In the more standard case of a composite null hypothesis with a density indexed by an unknown parameter vector, the GNPL can provide the most powerful test against a simple alternative (if such a test exists) but with the twist that average size is controlled over a countable number of subsets of the null parameter space. For a concise summary of the range of optimality properties that have been considered in the literature, see Sen Gupta (1991).

Based on a range of early applications largely involving testing the covariance matrix of the linear regression model (see Spjotvoll (1967), Davies (1969), Berenblut and Webb (1973), Fraser, Guttman and Styan (1976), Bhargava, Franzini and Narendranathan (1982), King (1981b, 1983a, 1983b, 1984, 1985a, 1985b, 1986, 1987a), Franzini and Harvey (1983), Sargan and Bhargava (1983), Evans and King (1985a, 1985b, 1988), King and Smith (1986), Shively (1986, 1988a, 1988b), Nyblom (1986) and Dufour and King (1991)), King (1987c) argued the case for the use of point optimal testing. He observed they best suit problems in which the parameter space under the alternative hypothesis can be restricted in scope by theoretical and technical (such as variances being positive) considerations. They work well when the null hypothesis can be reduced to a simple hypothesis by invariance (see, King 1980, 1987b) or similarity arguments

(see Hillier, 1987). They also allow one to trace out the maximum attainable power represented by the power envelope for a given testing problem.

It is important to note that the choice of a point optimal test does not mean that we believe the point at which power is optimized fully defines the alternative hypothesis. Rather, it is a choice of a particular test with a power curve that kisses the power envelope at the chosen point.

The aim of this paper is to update the review given in King (1987c) and outline the literature and its findings since 1987. There is a particular emphasis on how point optimal tests might be applied in cases where there are nuisance parameters that cannot be eliminated through invariance or similarity arguments. The paper also aims to make some further suggestions on solutions for problems that are less favourable to point optimal tests such as multivariate testing and the presence of nuisance parameters.

The plan of the paper is as follows. Section 2 reviews the literature since 1987 on point optimal testing where all nuisance parameters have been eliminated, typically through invariance arguments. Section 3 categorizes the various approaches to dealing with nuisance parameters including via (i) weighted averages of p values, (ii) approximate point optimal tests, (iii) plugging in estimated values, (iv) using asymptotics and (v) integrating out the nuisance parameters. Progress on using point optimal testing principles for two-sided and multi-dimensional alternatives is reviewed in Section 4. We give our thoughts in Section 5 on how best to deal with nuisance parameters under both the null and alternative hypotheses as well as presenting a new class of point optimal tests for multivariate testing. Finally, some concluding remarks are made in Section 6.

2 Tests where all nuisance parameters have been eliminated

In this section, we update King's (1987c) review of tests for problems in which all nuisance parameters have been able to be eliminated, typically through invariance arguments. A nice introduction to point optimal invariant testing in the linear regression model is given by Shively (2006). Work on improving the speed and

accuracy of numerical algorithms for calculating the p values (and critical values) of these and related tests have been reported by Shively, Ansley and Kohn (1990) and Ansley, Kohn and Shively (1992). They (Shively, Kohn and Ansley, 1994) also constructed a point optimal invariant test for nonlinearity in a semi-parametric regression model.

Since 1987, point optimal invariant tests have been proposed for a wide range of testing problems involving the covariance matrix in the linear regression model. These include (i) testing for autocorrelation in the presence of missing observations (Shively, 1993), (ii) testing for first order autoregressive (AR(1)) disturbances when the data is made up of the aggregate of a large number of small samples (Bhatti, 1992), (iii) testing for spatial autocorrelation in the disturbances (Martellosio, 2010, 2012), (iv) testing for block effects caused by random coefficients (Bhatti and Barry, 1995), (v) testing for quarter-dependent simple fourth-order autoregressive (AR(4)) disturbances (Wu and King, 1996), (vi) testing for joint AR(1)-AR(4) disturbances against joint MA(1)-MA(4) disturbances (Silvapulle and King, 1993) and (vii) testing for the presence of a particular error component (El-Bassiouni and Charif, 2004). Hwang and Schmidt (1996) extended the work of Dufour and King (1991) on testing the autocorrelation coefficient for stationary and nonstationary AR(1) disturbances while Dufour and Neifar (2008) extended it to the case of second order autoregressive (AR(2)) disturbances. Shively (2001) constructed a point optimal invariant unit root test of a random-walk-with-drift null hypothesis against a trend-stationary AR(1) alternative. This test is close to one of Dufour and King's (1991) tests, the main difference being the treatment of the initial observation. Nakatsuma et al. (2000) also derived a point optimal invariant test for a unit root in linear regression disturbances when the model is in first-differenced form. Honda (1989) showed that the class of these point optimal invariant tests is identical to the class of point optimal similar tests. This can also be concluded from Hillier's (1987) discussion of similar tests. Small (1993) observed that point optimal invariant tests for AR(1) disturbances in the linear regression can have their power tend to zero or a fraction between zero and one as the autocorrelation coefficient tends to one. This property, that is shared by the Durbin-Watson and

alternative Durbin-Watson tests (King, 1981a), confirms that the power envelope can have these properties.

Shively (1988a) devised a point optimal test for constant regression coefficients against Rosenberg's (1973) return to normalcy random coefficient model in the linear regression model. A modification to this test was suggested by Brooks (1993) who (Brooks, 1995) also investigated its robustness to Hildreth-Houck (1968) random coefficients and non-normality. Brooks (1997) studied its use, along with Brooks and King's (1994) APOI test, in a sequence of point optimal tests to select a varying coefficient model. Kurozumi (2003) derived the asymptotic distribution of a point optimal invariant test for a random walk regression coefficient in the linear regression model.

Point optimal tests (called beta-optimal tests after Davies, 1969) of the equicorrelation coefficient of a standard symmetric multivariate normal distribution was found to be approximately UMP by Bhatti and King (1990). This led to a series of papers involving point optimal testing in related settings including that of the linear regression model by Wu and Bhatti (1994) and Bhatti (1995, 2000). The problem of testing the value of the location parameter of a Cauchy density based on a single observation was investigated by Atiq-ur-Rehman and Zaman (2008) who constructed the class of point optimal tests for this problem. Davies (2001) considered testing for a unit root in an AR(1) process and also testing the stationary hypothesis against the integrated process in this setting. He observed that a time series made up of a Brownian motion sampled at equal time intervals plus white noise is exactly orthogonalized by the discrete cosine transformation-II and used this to construct beta-optimal tests.

3 Dealing with nuisance parameters when constructing point optimal tests

3.1 Weighted average of p values

An approach for dealing with unknown nuisance parameters when constructing locally best or point optimal tests that has clear potential and is worthy of further examination was suggested by King (1996). He proposed calculating p values conditional on the value of the nuisance parameters and then taking a weighted

average of these p values using either an appropriate marginal likelihood or the posterior density function of the nuisance parameters under the null hypothesis. The philosophy behind this approach is that there is information in the data about what nuisance parameter values are more likely than others, and this information should be utilized in the test procedure.

If we assume that the nuisance parameter values are known, then we can construct a point optimal test and, using Monte Carlo methods if needed, calculate the p value for this test. Then these p values, conditional on nuisance parameter values, can be averaged over the marginal likelihood or posterior density function of the nuisance parameters, typically using Monte Carlo integration. If Monte Carlo Markov chain methods are used to generate drawings from the marginal likelihood or posterior density of the nuisance parameters under the null hypothesis, then the procedure can be implemented as follows. After an appropriate burn-in period, for each drawing of the nuisance parameter vector, calculate the p value of the test conditional on that value. Following a large number of drawings (say 2,000), take as the p value of the test, the average of the calculated conditional p values.

King (1996) investigated the small sample properties of this approach for testing linear regression coefficients in the presence of AR(1) disturbances. The test used is the UMP invariant t test conditional on the value of the autoregressive parameter. This test, therefore, could be regarded as a point optimal test. King's approach was shown to be typically more accurate than the OLS based t test, the Durbin (1960) procedure outlined by King and Giles (1984), the standard maximum likelihood test and the t test based on Wooldridge's (1989) standard errors that are robust to serial correlation and hetroskedasticity.

3.2 Approximate Point Optimal Tests

The general testing problem considered by King (1987c) is one of testing

$$H_0 : x \text{ has density } f(x|\delta) \tag{1}$$

against

$$H_a : x \text{ has density } f(x|\theta), \quad (2)$$

where x is the observed sample, δ is a $w \times 1$ vector of parameters restricted to the set Δ and θ is a $q \times 1$ vector of parameters restricted to the set Θ . Any knowledge of the possible range of parameter values has been used to keep the parameter sets Δ and Θ as small as possible.

A point optimal test in this context involves choosing a value of θ , say θ_1 , at which power is to be optimized. A general, but not very explicit approach to constructing a point optimal test in this setting is discussed by Lehmann and Romano (2005, p 83-4) in the context of testing a composite null hypothesis against a simple hypothesis, in our case $H_a^* : x \text{ has density } f(x|\theta_1)$. It involves finding a probability density function over the Δ space, $h_\Delta(\delta)$, constructing

$$f_\Delta(x) = \int_\Delta f(x|\delta)h_\Delta(\delta) d\delta$$

and

$$s(\theta_1) = \frac{f(x|\theta_1)}{f_\Delta(x)}.$$

It also requires that a critical value c exists such that

$$\Pr[s(\theta_1) > c | x \sim f(x|\delta)] \leq \alpha, \text{ for all } \delta \in \Delta, \quad (3)$$

holds, where α is the desired level of significance and

$$\Pr[s(\theta_1) > c | x \sim f_\Delta(x)] = \alpha.$$

These requirements may not always be able to be met, in which case it is doubtful that a point optimal test can be found, at least by this approach. Lehmann and Romano (2005) give two examples in which $h_\Delta(\delta)$ has all its mass at a single point and one in which $h_\Delta(\delta)$ has mass at two points.

Determining what $h_{\Delta}(\delta)$ might be in a given application is not an easy task. They give the following suggestion which does provide some guidance. We should be looking for the $h_{\Delta}(\delta)$ which is of the least help in determining if H_a^* is true. In other words, we should look for the $h_{\Delta}(\delta)$ that provides the lowest power at $f(x|\theta_1)$ of the most powerful tests based on $s(\theta_1)$. If such a distribution can be found, then if (3) also holds, we have a point optimal test and $h_{\Delta}(\delta)$ is called the least favourable distribution.

King (1987c) conjectured that such a test could be constructed if one could find a point $\delta_1 \in \Delta$ and the critical value c such that

$$s(\delta_1, \theta_1) = \frac{f(x|\theta_1)}{f(x|\delta_1)} > c \quad (4)$$

is the most powerful test of the simple null

$$H_0^*: x \text{ has density } f(x|\delta_1)$$

against the simple alternative H_a^* ,

$$\Pr[s(\delta_1, \theta_1) > c | x \sim f(x|\delta_1)] = \alpha$$

and (3) holds with $s(\theta_1) = s(\delta_1, \theta_1)$. This provides an operational approach to constructing Lehmann and Romano's test in which $h_{\Delta}(\delta)$ has all its mass at a single point. For situations where appropriate values of δ_1 and c cannot be found, King (1987c) suggested an APO test which is based on (4) but requires δ_1 to be chosen such that (3) holds with $s(\theta_1) = s(\delta_1, \theta_1)$ and

$$\alpha - \Pr[s(\delta_1, \theta_1) > c | x \sim f(x|\delta_1)] \quad (5)$$

is minimized.

King (1989) constructed an APO invariant (APOI) test for simple AR(4) regression disturbances in the presence of AR(1) disturbances. For this problem, invariance arguments were used to remove most of the nuisance parameters so that $\delta = \rho_1$ where ρ_1 is the AR(1) parameter and $\theta = (\rho_1, \rho_4)'$ where ρ_4 is the simple AR(4) parameter. Note that ρ_1 is a nuisance parameter that cannot be eliminated, but its presence can be used to advantage. King's (1989) empirical power comparison of different APOI tests showed that it is important to have sensible rules for choosing $\theta_1 = (\rho_{11}, \rho_{41})'$ with a view to using the choice of ρ_{11} to help minimize (5) and therefore improve the optimality of the test. He acknowledged that the test required a lot of computation to apply.

Silvapulle and King (1991) investigated the APOI test of first order moving average (MA(1)) disturbances against AR(1) disturbances in the linear regression model. They conducted an empirical size and power comparison of their APOI test with an asymptotic test of the second-order autocorrelation coefficient of the disturbances (which is zero under the null and non-zero under the alternative) and a Lagrange multiplier (LM) test. The study led to the conclusion that their APOI test has superior small-sample size and power properties compared to the other two asymptotic tests they considered.

The problem of testing Hildreth-Houck (1968) against Rosenberg's (1973) return to normalcy random coefficients in the linear regression model was investigated by Brooks and King (1994). They were unable to construct a point optimal test, so considered the class of APOI tests. They found these tests to have good small-sample properties compared to the likelihood ratio and Wald tests in a limited empirical power comparison.

Rahman and King (1994) considered APOI tests for testing for random regression coefficients in the presence of autocorrelation in the regression disturbances. They compared the small sample properties of these tests with those of the LM

and LMMP tests based on the marginal likelihood. These latter tests were found to work well in this context and they concluded that “the extra work required to apply APOI tests hardly seems worthwhile, particularly for larger sample sizes”. An extension of this power comparison to non-normality may be found in King and Rahman (2015).

Silvapulle (1994) constructed the APOI test for AR(1) disturbances against the alternative of IMA(1,1) disturbances in the linear regression model. She compared the small sample properties of the APOI test with a test suggested by Godfrey and Tremayne (1988) and the LM test. She found for positively correlated errors, the APOI test performs best while for negatively correlated errors and larger sample sizes the LM test is best.

Overall, the literature on APO testing suggests its use does involve a lot of computation for not much extra reward. Also, the use of an APO test does not always guarantee the best test in terms of power. These sorts of conclusions have led to the search for other solutions.

3.3 Alternative approaches to approximate point optimal tests

Using the GNPL, Srianthakumar and King (2006) introduced another version of the APO test of a composite null (henceforth referred to as the g test). They found the g test has good size and power properties for the same testing problems considered by Silvapulle and King (1991) and Silvapulle (1994). Its construction involves deciding on appropriate representative points under the null hypothesis via a trial and error process and controlling multiple critical values as explained below.

In order to construct a point optimal test for testing (1) against (2), let us assume that $\theta_1 \in \Theta$ is the point under the alternative hypothesis at which we wish to optimize power. Thus, the testing problem given in (1) and (2) can now be written as testing (1) against H_a^* .

We then need to approximate $f(x|\delta)$, by a finite number of densities. Regard these as representative densities of $f(x|\delta)$, $\delta \in \Delta$.

The g test is the test with the minimum number of representative densities under the approximating null that allows the size to be sufficiently controlled over the complete null hypothesis parameter space. In the limited case of $w = 1$ (i.e., δ is a scalar) and Δ being a closed interval, experience is that at least three representative densities are needed for the approximating null. Therefore, to construct the g test, we start with three representative densities denoted

$f_1(x|\delta_1)$, $f_2(x|\delta_2)$ and $f_3(x|\delta_3)$ and find k_1 , k_2 and k_3 values such that the following size conditions (which are evaluated via the Monte Carlo method) hold simultaneously:

$$\Pr \left[f(x|\theta_1) > \sum_{i=1}^3 k_i f_i(x|\delta_i) \mid x \sim f_j(x|\delta_j) \right] = \alpha, \quad j = 1, \dots, 3. \quad (6)$$

In the case of $w = 1$ and Δ being a closed interval, δ_1 and δ_3 can be the two end points of Δ , and δ_2 can be any point in between. If three representative densities are not adequate to control the sizes of the test, the number of representative densities under the null can be increased by one and the process repeated. The critical values k_i , $i = 1, 2, 3$ can be obtained using one of two methods: either a systematic iterative procedure or via Simulated Annealing (SA)¹ (Srianthakumar and King, 2006). The g test can be computationally intensive, particularly for high dimensional testing problems. In addition, Srianthakumar (2013) showed that the g test may not be trustable in the presence of unavoidable nuisance parameters. In particular, Srianthakumar (2013) investigated the problem of testing for a linear regression model with AR(1) errors against a first-order dynamic linear regression model with white noise errors using marginal likelihood based g tests and marginal likelihood based classical (LR, LM and W) tests. She showed that for this testing problem the g tests have good power

¹ See Goffe et al. (1994) for more details about SA.

properties, particularly in the neighborhood of the chosen parameter point under the alternative hypothesis where power is optimized. However, when moving further away from this parameter point, the power of the g test becomes less desirable.

Begum and King (2005a) introduced a Most Mean Powerful (MMP) test of a composite null based on the GNPL. Their test maximizes average power subject to controlling average size over different subsets of the null hypothesis parameter space. The standard approach of controlling the maximum size over the nuisance parameter space is typically difficult and time consuming. Begum and King's approach of controlling average size over sub-regions selected to reduce variability in size seems to be a novel idea. In the context of testing for MA(1) errors against AR(1) errors in the linear regression model, their approach is shown to work well. This testing problem, after reduction through invariance arguments, becomes one dimensional. Even for this case, the MMP invariant test can be computationally intensive. Begum and King (2005b) successfully applied the MMP test to testing higher order regression disturbances, namely joint MA(1)-MA(4) against joint AR(1)-AR(4). They note that the increase in dimension does increase significantly the computational effort required to apply the test. They also (Begum and King (2006)) considered the problem of testing for heteroscedastic disturbances in the linear regression model which involves nuisance parameter space which in a one-sided infinite interval. Their test was found to have encouraging small-sample size and power properties.

3.4 An estimated parameter approach

The problem of testing for AR(1) disturbances in the linear regression model with lagged dependent variables was considered by Inder (1990). He proposed replacing the coefficients of the lagged dependent variables with estimates and then applying King's (1985a) point optimal test for AR(1) disturbances using small-disturbance asymptotic critical values. Inder (1990) reported Monte Carlo results showing the new test has superior small-sample powers compared to existing tests. A modification to his choice of critical values was suggested by King and Harris (1995) based on earlier work by King and Wu (1991).

The problem of testing for moving average unit roots in autoregressive integrated moving average (ARIMA) models was considered by Saikkonen and Luukonen (1993a, 1993b). They constructed point optimal tests and, in their more general case, used estimated values of nuisance parameters in their test statistic. They then derived the asymptotic distribution of their test statistic to allow asymptotic critical values to be obtained.

Their work (and also that of Shively, 1988b) was extended by Hwang and Schmidt (1993) to the case where the null model contains a linear trend and so is trend stationary. Hwang and Schmidt provide critical values and small-sample power for their point optimal invariant tests. Shively (2004) constructed an approximate point optimal invariant test of a unit root in the context of testing an ARIMA (p-1, 1, q) process with drift against an ARMA(p, q) trend-stationary process in which unknown nuisance parameters are replaced with estimates. Gallego and Diaz (2007) extended the tests of Saikkonen and Luukonen in univariate ARIMA models to multivariate ARIMA models.

This raises the question of why not replace unknown nuisance parameters by estimates? If we return to the general problem of testing (1) against (2), then we might regard δ as a vector of nuisance parameters and θ might be re-arranged and split into parameters of interest θ_a and those not of interest, denoted θ_b , so that $\theta' = (\theta_a', \theta_b')$. If we now write $f(x|\theta)$ as $f(x|\theta_a, \theta_b)$, the suggestion is that a point optimal like test might be based on rejecting the null for large values of the likelihood ratio

$$\frac{f(x|\theta_{a1}, \hat{\theta}_b)}{f(x|\hat{\delta})} \tag{7}$$

where θ_{a1} is the point at which one wishes to optimize power, $\hat{\theta}_b$ is the maximum likelihood estimate of θ_b under $f(x|\theta_{a1}, \theta_b)$ with θ_b being the only parameters that are estimated, and $\hat{\delta}$ is the maximum likelihood estimate of δ under H_0 .

Unfortunately such a test is no longer optimal because of the replacement of parameters with their estimates which are stochastic. However, as we shall see in Section 3.5, there can be circumstances in which this approach provides a test with optimal asymptotic properties.

There is also the issue of finding appropriate critical values. An approach that has not explicitly been raised in the literature (to the best of our knowledge) is to exploit the parallels between (7) and the Cox and related tests. Dastoor and Fisher (1987, 1988) noted the link between point optimal invariant tests of regression disturbances and Cox tests. They observed that this class of tests can be regarded as a class of Cox tests which have an exact distribution.

The problem of finding the asymptotic distribution of the log of (7), namely

$$\log f(x|\theta_a, \hat{\theta}_b) - \log f(x|\hat{\delta}) \quad (8)$$

under the null hypothesis is exactly the problem that Cox (1961, 1962) considered in his seminal papers. He proposed standardizing (8) by finding or approximating its mean and standard deviation, and treating the standardized statistic as asymptotically standard normal (see White (1982) for a discussion on the regularity conditions involved). Standardizing (8) is not always easy, but there is a very rich literature on the application of the Cox approach, see for example survey articles by MacKinnon (1983), McAleer (1987), Gouriéroux and Monfort (1994) and Pesaran and Weeks (2003). The latter paper explores two more practical approaches to this problem, involving the use of simulation methods and parametric bootstrap methods.

Two important points should be borne in mind when using this asymptotic approach. King and McAleer (1987) found this standardized version of the Cox test to have very poor small sample size and power for the problem of testing AR(1) disturbances against MA(1) disturbances in the linear regression model. For a sample size of 30 at the nominal significance level of 5%, sizes can be as high as 0.5 while, when simulation methods are used to find appropriate small

sample critical values, the power of the test does not rise much above its significance level of 0.05. As with many asymptotic tests, the standardization of the original statistic can do more harm than good if it is based on poor estimates. If a solution can be found without the standardization step (see for example, King (1998)), it is likely to produce much better small sample power and size. If standardization is needed, then it is important that the sample size is at least 100 and preferably larger.

The second point is not to think of the density $f(x|\theta_a, \theta_b)$ as that of the alternative model (as one typically would when conducting a Cox test). We are just working with that as a device for constructing a good test against the more general model given by (2).

3.5 Using an asymptotic approach

The biggest development in the last two decades has been the construction of point optimal tests with the use of asymptotics to simplify some of the problems caused by nuisance parameters. The seminal paper in this literature is Elliott, Rothenberg and Stock (1996). They considered the following data generation process:

$$y_t = d_t + u_t, \quad t = 1, \dots, T,$$

and

$$u_t = \alpha u_{t-1} + v_t,$$

where d_t is a deterministic mean component, u_t is an error term with zero mean, v_t is a stationary disturbance process with mean zero and $(v_1, \dots, v_T)'$ having covariance matrix Σ . Their interest is in testing for a unit root, namely $H_0: \alpha = 1$ against $H_a: |\alpha| < 1$. Nuisance parameters/components in this problem are d_t , u_0 , and Σ . With particular assumptions about d_t and u_0 and the restrictive assumption that $\Sigma = \sigma^2 I$, Dufour and King (1991) proposed point optimal invariant tests for this problem.

In order to analyse the qualities of competing tests of this general problem, Elliott et al. (1996) reparameterized the parameter under test to

$$\eta = T(\alpha-1)$$

and then derived the asymptotic power envelope using local to unity asymptotics and point optimal tests assuming $\eta = \bar{\eta}$, Gaussian errors and known values of u_0 and Σ . A major determinant of the behavior of the asymptotic power envelope is what is known about d_t . If d_t is known or is unknown but slowly evolving, the asymptotic power envelope remains the same. If $d_t = \beta'z_t$, where β is an unknown $q \times 1$ parameter vector and z_t is a $q \times 1$ vector of known regressors, then invariance arguments can be used to construct an asymptotic power envelope using the family of point optimal invariant tests.

The authors then consider a class of feasible point optimal invariant type tests that require a choice of $\bar{\eta}$ for when u_0 and Σ are unknown but whose asymptotic power function kisses the asymptotic power envelope constructed using known u_0 and Σ and particular forms of d_t . In that sense, their tests can be regarded as point optimal. An argument they could have used and acknowledge in subsequent papers (Elliott, Müller and Watson, 2012 and Elliott and Müller, 2014) is to use the LeCam limits of experiments approach to justify the efficiency of the resultant test. (For a textbook discussion of this approach, see van der Vaart, 1998, Chapter 9 and for an econometric testing application see Ploberger, 2004). Müller (2011) provides an excellent overview of how this alternative approach works for the types of testing problems discussed in this section. A key result is that for any limiting experiment (model), an optimal test in the limit must be the limit of optimal tests in the small sample setting. This gives an insight as to how to construct asymptotically optimal tests based on the limiting experiment (model). Elliott et al. (1996) noted that in the case of d_t known (or equivalently $d_t = 0$), Dickey and Fuller's (1979) t test has asymptotic power equal to that of the power envelope when the asymptotic power is 0.5. They gave some recommended

values of $\bar{\eta}$ for use in their test and also investigated its small sample properties. Their recommended tests, including a modified Dickey-Fuller t test, were found to largely follow their asymptotic properties, although some forms of Σ can cause poor size and power, particularly if v_t has a large moving average component. Burrige and Taylor (2000) provided further analysis of the power properties of Elliott et al.'s proposed feasible test.

There have been a number of useful extensions of this work. Rothenberg and Stock (1997) applied the methodology to a simpler AR(1) model with well-behaved but non-normal errors. They found that the asymptotic and small-sample power curves and power envelopes can be sensitive to the degree of non-normality in the errors with heavy tailed distributions being a particular problem.

A critical assumption that Elliott et al. (1996) made is that the initial error u_0 has finite variance for all values of α in the neighbourhood of $\alpha = 1$. This rules out the possibility it has variance $\sigma^2/(1-\alpha^2)$ for $\alpha < 1$ which is often assumed for a stationary AR(1) process. Elliott (1999) reworked the analysis under this latter assumption and found it changes the class of asymptotically optimal tests, confirming that these tests are sensitive to what is assumed about the distribution of u_0 . Vougas (2009) discussed the relationship between Elliott's new tests and Dufour and King's (1991) point optimal tests. He suggested a modification to the latter to improve its usefulness and tabulated critical values for two important cases.

Xiao (2001) considered estimation and testing (including point optimal type testing) in the Elliott et al. (1996) model under the assumptions of trending means and general non-Gaussian disturbance distributions.

Müller and Elliott (2003) reparameterized the dynamic model considered by Elliott (1996) et al. and showed that the power of most unit root tests depends (among other things) on a parameter ξ which they define as the deviation of the initial observation y_0 from the model's deterministic component. For the problem

of testing $H_0: \alpha = 1$, ξ is a nuisance parameter that needs to be dealt with in some way. They took the innovative approach of considering the class of tests that optimize power at a chosen value of α , $\bar{\alpha}$, averaged by a weighting function over the range of possible ξ values. This allowed them to construct a new class of tests using Elliott et al.'s approach although their main emphasis was to provide a theoretical basis for understanding the power properties of a range of existing tests. In a follow-up paper, Elliott and Müller (2006a) further investigated this dependency on ξ . They proposed an asymptotically efficient unit root test, whose power curve changes least with changes in ξ , for use when the researcher knows very little about the possible magnitude of ξ . Wang (2014) investigated the use of bootstrap methods to apply Elliott et al.'s (1996) asymptotic point optimal test used to map out the power envelope but which is infeasible because it assumes Gaussian errors and known values of u_0 and Σ . Monte Carlo simulations show that the bootstrap PO test is a feasible test that has good small-sample size and power properties.

The reverse problem of testing the null hypothesis of stationarity against the alternative of a unit root was considered by Müller (2005) who used the local to unity point optimal tests to improve the asymptotic properties of the locally best invariant tests.

Elliott et al.'s (1996) methodology was applied by Elliott and Jansson (2003) to the problem of testing for a unit root in a variable when the variable is modeled with a number of stationary related variables. They showed that good power gains can be obtained when such covariates are included in the test procedure as pointed out by Hansen (1995). This step involved the construction of the power envelope assuming knowledge of the nuisance parameters. A feasible test was constructed that can be applied by running vector auto-regressions. It was shown to have good small-sample power properties and to kiss the asymptotic power envelope at a chosen point.

Jansson (2005) applied the Elliott et al. (1996) methodology to the problem of testing the null hypothesis of cointegration against the alternative of no

cointegration in a linear dynamic model. The asymptotic power envelope assuming Gaussian errors and known nuisance parameters was derived. He then constructed a feasible point optimal test and investigated its asymptotic and small-sample performance. For an extension to this work, see Kurozumi and Arai (2005).

In another application of Elliott et al.'s (1996) approach, Elliott, Jansson and Pesavento (2005) investigated the problem of testing for a unit root in a known cointegrating vector. The feasible test they proposed was shown to be asymptotically equivalent to a point optimal invariant test. A related paper by Elliott and Pesavento (2009) considered the problem of testing the null hypothesis of no cointegration when the cointegrating variables are known to have a unit root. They traced out the power envelope using point optimal tests that maximize the weighted average power for different weightings over the unknown cointegrating vector parameter space. This provides power bounds for the evaluation of a range of existing tests. It is also another illustration of dealing with an influential nuisance parameter by constructing tests that maximize weighted average power at a point.

Elliott and Müller (2006b) investigated the problem of testing for time variation, instability or breaks in regression coefficients of the linear model. They showed that for a wide class of persistent breaking processes and assuming Gaussian errors, a range of tests designed to be efficient in small samples are asymptotically equivalent. This allowed them to recommend an asymptotically point optimal test that is very attractive because of its ease of application and its small-sample power properties. Lee (2009) reworked their analysis under weaker assumptions including the error distribution being unknown.

Using the GLS-detrending approach that is a feature of the asymptotic point optimal invariant tests of Elliott et al. (1996), Perron and Rodriguez (2003) extended the class of M-tests for unit roots proposed by Perron and Ng (1996) and Ng and Perron (2001) to allow for a change of unknown timing in the trend function. Liu and Rodriguez (2006) extended this work along the lines of Elliott

and Jansson (2003) to testing for a unit root in the presence of a structural break of unknown timing and proposed a new feasible point optimal test.

Gregoir (2006) applied Elliott et al.'s approach to testing for the presence of a pair of complex conjugate unit roots in a real time series. The feasible test that resulted allowed him to propose some new, near-efficient, seasonal unit root tests. Building on Gregoir (2006), Rodrigues and Taylor (2007) extended the results of Elliott et al. (1996) to testing for seasonal unit roots. They found that the asymptotic point optimal test of a root at a particular spectral frequency, asymptotically is independent of whether there are unit roots at other frequencies.

Moon, Perron and Phillips (2007) considered testing for unit roots in panel data models. They constructed the local asymptotic power envelope under a range of scenarios and suggested a point optimal invariant panel unit root test for each case. This was extended by Moon, Perron and Phillips (2014) to allow for the possibility of serially correlated errors.

The choice of point at which asymptotic power is optimized in Elliott et al.'s (1996) approach, is driven by asymptotic considerations. Broda, Carstensen and Paoletta (2009) asked if there are small-sample considerations that can be used to help improve power. They showed there are advantages in expressing the various tests as ratios of quadratic forms in normal variables. This allowed them to apply Juhl and Xiao's (2003) idea of using a power loss criterion to determine the chosen point under the alternative hypothesis. They also showed that there are advantages in using recursive GLS rather than conventional GLS in the feasible test.

Finally, Elliott et al.'s (1996) emphasis on the power envelope and the small sample point optimal test allowing the power envelope to be traced out has led to a new standard in the evaluation of new tests, that is to include a comparison of the new test's power with a particular power envelope.

3.6 Integrating out nuisance parameters

Elliott, Müller and Watson (2012) used a weighting function to integrate out nuisance parameters under the alternative hypothesis. This results in a test that maximizes average power and any weighting function can be used. For nuisance parameters under the null hypothesis, this approach requires considerable care because as noted in Section 3.3, the weighting function needs to be the least favourable distribution. Elliott et al. (2012) propose an approximate least favourable distribution be used and that it be chosen to minimize power at the chosen point (a requirement of the least favourable distribution). Müller and Watson (2013) apply this approach to cointegration testing while Elliott and Müller (2014) apply it to the problem of testing hypotheses about the pre and post break values of a parameter when there is a single break in a time series with unknown timing. A third application is provided by Müller (2014) and involves heteroscedasticity and autocorrelation standard errors for time series inference.

4. Point optimal testing against two-sided and multi-dimensional alternatives

Andrews, Moreira and Stock (2006) considered two-sided testing of the coefficient of a single included endogenous regressor in an instrumental variables regression. They constructed a two-sided power envelope for invariant similar tests via point optimal invariant similar two-sided test. This allowed them to assess the properties of a range of existing tests and make recommendations on which are best to use. Their two-sided power envelope was obtained via a class of two-point optimal invariant tests which involve maximizing the average power at two chosen points, one on each side of the null hypothesis. Care needs to be taken in how these points are chosen – they used an asymptotic efficiency requirement. They also briefly mentioned two other approaches to constructing two-sided power envelopes, both of which give similar (or the same) power envelopes. These findings were extended to the class of non-similar tests by Andrews, Moreira and Stock (2008).

Dufour and Iglesias (2008) suggested a novel approach to point optimal (and locally best) testing involving a potentially multidimensional composite alternative. Their approach requires splitting the sample into two parts, a smaller sample (approximately 10%) that is used to decide on the alternative hypothesis

point for the point optimal test and the remainder of the sample that is used to conduct the test. The alternative hypothesis point is determined either via a consistent estimator (if one is known to exist) or by maximizing the asymptotic power. They called this the split-sample Monte Carlo adaptive optimal test and demonstrated its application to a range of volatility models with Gaussian or heavy-tailed errors. Their test has attractive features in that it does not require the existence of moments and can be applied in a range of settings such as non-normality and non-stationarity. The negative is the power loss that comes from not using all the observations in the actual test. The hope is that this loss will be small and more than compensated by optimizing power with the remaining observations at the most likely alternative hypothesis point.

Dufour and Taamouti (2010) constructed point optimal sign-based tests in linear and nonlinear regression models that are valid under non-normality and heteroscedasticity of unknown form. A split-sample approach is used in order to choose the alternative point in a way that brings the power curve close to the power envelope.

5 Fertile areas for future research

There is no doubt that the theory of point optimal testing has come a long way since 1987. The use of power envelopes as a benchmark for the power function of new tests has become more prevalent, particularly in the unit root testing literature. Clearly learning how best to deal with nuisance parameters has been a significant thrust of the literature. A second issue that to date has received very little attention (see the previous section), is how the principle of point optimal testing might best be employed against two-sided and multi-dimensional alternatives.

5.1 Dealing with nuisance parameters under the null hypothesis

It is our view that different approaches to handling nuisance parameters might be needed depending on whether they occur under the null or alternative hypothesis. We turn first to the null hypothesis case.

Using the notation of Section 3.2, assume after the problem has been reduced down to its smallest dimensions through invariance and other arguments, the null hypothesis is given by (1). Effectively, δ is a vector of nuisance parameters. The problem with δ is the difficulty it can cause one when controlling the probability of a Type I error (PTIE). If the point optimal test of interest is a similar test (has the same PTIE for all parameter points, δ , under the null hypothesis), there is no issue. If there are a range of point optimal tests to choose from, say of the form of (4) and indexed by δ_1 , then one might choose the δ_1 value that maximizes power of the resultant test at θ_1 . (An example of this approach of choosing nuisance parameter values to maximize power is choosing band-width parameters used in a test statistic; see Gao and Gijbels, 2008, Sun, Phillips and Jin, 2008, Gao et al., 2009a, 2009b and Gao and King, 2014).

A more likely scenario is that the preferred test statistic is non-similar. The next obvious approach is to see if asymptotic arguments (using Müller, 2011 for guidance) allow one to replace the remaining nuisance parameters with estimates for an asymptotically optimal test. In the remainder of this section, we will assume this is not the case. The conventional approach to non-similar testing is to find the critical value that makes the PTIE less than or equal to the desired test size (say 5%) over the entire null hypothesis parameter space, Δ . If the null hypothesis holds, then δ will have a true value which we will denote as δ_0 . If the desired size of the test is α and the PTIE at δ_0 is α_0 , then rather than applying an α level test, we are applying an α_0 level test with a consequential loss in power. A test of the form of (4) is no longer point optimal if $\alpha_0 < \alpha$ because of this power loss. It could be that α_0 is close to zero which might result in a rather extreme drop in power. One solution already discussed in Section 3.2 (and by Elliott et al., 2012) is to look for the least favourable distribution over the δ parameter space, $h_\Delta(\delta)$. Lehmann and Romano (2005) note that we are essentially looking for the weighting function, $h_\Delta(\delta)$, that gives the lowest power at θ_1 . This seems to be the opposite of what we should be doing, particularly given there is information in the observed sample, x , about what likely δ_0 values

might be when the null hypothesis is true. If we knew what δ_0 was, then the appropriate test would be (4) with $\delta_1 = \delta_0$ and we could easily find the appropriate critical value, c , by simulating the null distribution from $f(x|\delta_0)$.

Typically δ_0 is unknown but almost always we can find its posterior density function, at least empirically. This might be used to average p values conditional on different δ_0 values as suggested by King (1996). Whether there are other ways of using the information in the data about δ_0 to help build a powerful test is clearly an important area for future research. The main idea is that we only need to worry about controlling the size of a test for reasonably likely δ_0 values. As the sample size grows, this neighbourhood of concern should shrink to the true value δ_0 .

5.2 Dealing with nuisance parameters under the alternative hypothesis

Turning to the problem of nuisance parameters under the alternative hypothesis, we think the most fruitful approach is to optimize average power over the nuisance parameter space at the chosen value of the parameter vector of interest.

Using the notation of Section 3.4, suppose $\theta' = (\theta'_a, \theta'_b)$ where θ_a is the parameter vector of interest and θ_b is the nuisance parameter vector. We now write $f(x|\theta)$ as $f(x|\theta_a, \theta_b)$ and let $f_b(\theta_b)$ be an appropriate (or chosen) weighting function over the θ_b parameter space which we will denote as Θ_b .

If we denote the rejection region (the complete set of x values for which the null hypothesis is rejected) of a test of (1) against (2) by ω , then its PTIE is given by

$$\int_{\omega} f(x|\delta) dx$$

and its power is given by

$$\int_{\omega} f(x|\theta_a, \theta_b) dx$$

which is clearly a function of θ_a and θ_b . The power averaged by $f_b(\theta_b)$ over Θ_b is therefore

$$\begin{aligned} & \int_{\Theta_b} \int_{\omega} f(x|\theta_a, \theta_b) dx f_b(\theta_b) d\theta_b \\ &= \int_{\omega} \int_{\Theta_b} f(x|\theta_a, \theta_b) dx f_b(\theta_b) d\theta_b dx \\ &= \int_{\omega} f_a(x|\theta_a) dx \end{aligned} \quad (9)$$

$$\text{where } f_a(x|\theta_a) = \int_{\Theta_b} f(x|\theta_a, \theta_b) f_b(\theta_b) d\theta_b. \quad (10)$$

Observe that (9) can be interpreted as the power of the test with rejection region ω when the data has been generated from the distribution with density $f_a(x|\theta_a)$ given by (10).

As noted by Begum and King (2005a, p 1083), the GNPL implies that the test of $H_0: \delta = \delta_1$ against (2) that maximizes average power over Θ_b at $\theta_a = \theta_{a1}$ involves rejecting H_0 for

$$\frac{f_a(x|\theta_{a1})}{f(x|\delta_1)} > c \quad (11)$$

where c is the appropriate critical value.

It may be that there is a closed form solution to the integral in (10). If $f_b(\theta_b)$ is considered to be a prior then the literature on conjugate prior distributions might help find a class of weighting functions $f_b(\theta_b)$ for which a closed form of (10) is known.

If a closed form solution to (11) is not available, we can proceed as follows. Let $\theta_{b,1}, \dots, \theta_{b,m}$ be a simple random sample from $f_b(\theta_b)$ of size m , then via Monte Carlo integration, the left-hand side of (11) can be approximated by

$$\frac{1}{m} \sum_{i=1}^m \frac{f(x|\theta_{a1}, \theta_{bi})}{f(x|\delta_1)} \quad (12)$$

and the test that maximizes average power over Θ_b at $\theta_a = \theta_{a1}$ involves rejecting the null hypothesis for large values of (12). If need be, the critical value c can be found by simulating (12) for repeated samples of x from the $f(x|\delta_1)$ distribution.

The literature on optimizing average power does suggest that some care is needed in choosing $f_b(\theta_b)$. Future research is needed to see how well these tests might work in practice.

5.3 Handling multi-dimensional parameter spaces under the alternative hypothesis

With respect to multidimensional testing, we will illustrate a potential approach to point optimal testing by considering the problem of testing $\theta = 0$ against $\theta \neq 0$ when x has density $f(x|\theta)$, where θ is $q \times 1$ and Θ is R^q . Observe that θ can be reparameterized into polar coordinates $(r, \phi_1, \dots, \phi_{q-1})'$, $r \geq 0$, $\phi_j \in [0, \pi]$, $j = 1, \dots, q-2$ and $\phi_{q-1} \in [0, 2\pi]$, via

$$r = (\theta' \theta)^{1/2},$$

$$\theta_1 = r \cos \phi_1,$$

$$\theta_j = \left(\prod_{i=1}^{j-1} \sin \phi_i \right) \cos \phi_j, \quad 2 \leq j \leq q-1$$

$$\theta_q = \prod_{i=1}^{q-1} \sin \phi_i.$$

The problem of testing $\theta = 0$ against $\theta \neq 0$ now becomes one of testing whether the scalar $r = 0$ against $r > 0$, in the presence of nuisance parameters $\phi_1, \dots, \phi_{q-1}$ which collectively determine a direction from the origin in R^q space. For

examples of this kind of transformation being usefully used in hypothesis testing, see King and Shively (1993) and King and Edwards (1989).

Along the lines of (11), a test can be constructed that has maximum average power across the nuisance parameters space, at $r = r_1$. If the weighting function is chosen to be uniform overall direction from the origin in R^q space, then our test statistic (12) becomes the sum of likelihood ratios sampled over random directions from $\theta = 0$. The choice r_1 could be that which makes the average power at $r = r_1$ equal to 0.75. Again further research is needed to see how well this class of optimal tests might work in practice.

6 Concluding Remarks

As the list of references that follow attest, there has been considerable innovation and research on point optimal testing since 1987. A high proportion of this new literature has been in the very highly researched area of unit root testing. This has proved to be an extremely difficult testing problem that point optimal testing and particularly its application in a local-to-unity asymptotic setting by Elliott et al. (1996) and more recently Müller (2011) have helped solve, although we continue to see innovations that result in power improvements. In particular, Broda et al. (2009) have reminded us of the importance of small-sample consideration on power by using Juhl and Xiao's (2003) optimal approach to selecting the point at which power is optimized and recursive GLS detrending rather than conventional GLS detrending in the feasible test.

Juhl and Xiao's investigation of "optimal" point optimal testing does provide some guidance on the application of Davies' (1969) beta optimal test. We now have a better idea of what level of power one should choose to optimize power at. Davies (1969) originally recommended 0.8, King (1985b) suggested 0.65 and many other researchers have used 0.5. We now recommend 0.75 as a consequence of Juhl and Xiao's (2003) finding.

The split sample testing approach to point optimal testing suggested by Dufour and Iglesias (2008) may have merit, but more research is needed to see if this is indeed the case. The biggest issue is where to split the sample between that used to choose the point at which power is “optimized” and that used to conduct the test. Juhl and Xiao’s (2003) optimal approach could be used to determine the optimal split as well as the point at which power is “optimized”. The resulting power function could then be compared with other point optimal tests which use the full sample for testing to see whether the power loss from splitting the sample is too great. We guess it might be.

The modern literature on point optimal testing has provided a greater emphasis on the power envelope. A number of recent papers proposing new tests have compared the power of their test with a particular power envelope. This approach to test evaluation should be encouraged where possible.

In this paper we have made a series of suggestions for future research. These include the need to handle nuisance parameters differently under the null and alternative hypotheses. The literature on the Cox and related non-nested tests may help with finding appropriate critical values if nuisance parameters are replaced with their maximum likelihood estimates. Optimizing average power across the nuisance parameter space under the alternative at the chosen point for the parameter(s) of interest, in our view, is an approach worthy of further scrutiny. We give a general formula for the construction of such tests in a general setting. We also discuss how a simple transformation from Cartesian to polar coordinates can transform a multivariate testing problem into a one-sided scalar testing problem with nuisance parameters that is amenable to such a solution.

It is fair to conclude from our review of the recent literature and our suggestions for new generic point optimal tests, that point optimal testing has much greater applicability than was apparent in 1987.

References

- Andrews, D.W.K., M.J. Moreira and J.H. Stock (2006). Optimal two-sided invariant similar tests for instrumental variables regression, *Econometrica*, 74, 715-752.
- Andrews, D.W.K., M.J. Moreira and J.H. Stock (2008). Efficient two-sided nonsimilar invariant tests in IV regression with weak instruments, *Journal of Econometrics*, 146, 241-254.
- Andrews, D.W.K. and W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, 62, 1383-1414.
- Ansely, C.F., R. Kohn and T.S. Shively (1992). Computing p values for the generalized Durbin-Watson and other test statistics, *Journal of Econometrics*, 54, 277-300.
- Atiq-ur-Rehman, A. and A. Zaman (2008). Most stringent test for location parameter of a random number from Cauchy density, *IIIE*, International Islamic University, Islamabad.
- Begum, N. and M.L. King (2005a). Most mean powerful test of a composite null against a composite alternative, *Computational Statistics and Data Analysis*, 49, 1079-1104.
- Begum, N. and M.L. King (2005b). Most mean powerful invariant test for testing two-dimensional parameter spaces, *Journal of Statistical Planning and Inference*, 134, 536-548.
- Begum, N. and M.L. King (2006). Most mean powerful test for testing heteroscedastic disturbances in the linear regression model, *Model Assisted Statistics and Applications*, 1, 9-16.
- Berenblut, I.I. and G.I. Webb (1973). A new test for autocorrelated errors in the linear regression model, *Journal of the Royal Statistical Society, Series B*, 35, 33-50.
- Bhargava, A., Franzini, L. and W. Narendranathan (1982). Serial correlation and the fixed effects model, *Review of Economic Studies*, 49, 533-549.
- Bhatti, M.I. (1992). Optimal testing for serial correlation in a large number of small samples, *Biometrical Journal*, 34, 57-66.

- Bhatti, M.I. (1995). Optimal testing for equicorrelated linear regression models, *Statistical Papers*, 36, 299-312.
- Bhatti, M.I. (2000). Some tests for the equicorrelated models: An example of loss in power, *Statistical Papers*, 41, 345-352.
- Bhatti, M.I. and A.M. Barry (1995). On testing for equicorrelation in random coefficient models, *Biometrical Journal*, 37, 383-391.
- Bhatti, M.I. and M.L. King (1990). A beta optimal test for equicorrelation coefficient, *Australian Journal of Statistics*, 32, 87-97.
- Broda, S., K. Carstenen and M.S. Paoella (2009). Assessing and improving the performance of nearly efficient unit root tests in small samples, *Econometric Reviews*, 28, 468-494.
- Brooks, R.D. (1993). Alternative point optimal tests for regression coefficient stability, *Journal of Econometrics*, 57, 365-376.
- Brooks, R.D. (1995). The robustness of point optimal testing for Rosenberg random regression coefficients, *Econometric Reviews*, 14, 35-53.
- Brooks, R.D. (1997). Using a sequence of point optimal tests to select a varying coefficient model, *Communications in Statistics – Simulation and Computation*, 26, 671-685.
- Brooks, R.D. and M.L. King (1994). Testing Hildreth-Houck against return to normalcy random regression coefficients, *Journal of Quantitative Economics*, 10, 33-52.
- Burridge, P. and A.M.R. Taylor (2000). On the power of GLS-type unit root tests, *Oxford Bulletin of Economics and Statistics*, 62, 633-645.
- Cox, D.R. (1961). Tests of separate families of hypotheses, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*.
- Cox, D.R. (1962). Further results on tests of separate families of hypotheses, *Journal of Royal Statistical Society, Series B*, 24, 406-424.
- Cox, D.R. and D.V. Hinkley (1974). *Theoretical Statistics*. London: Chapman and Hall.
- Dastoor, N.K. and G. Fisher (1987). Comment on towards a theory of point optimal testing, *Econometric Reviews*, 6, 219-229.
- Dastoor, N.K. and G. Fisher (1988). On point-optimal Cox tests, *Econometric Theory*, 4, 97-107.

- Davies, R.B. (1969). Beta-optimal tests and an application to the summary evaluation of independent experiments, *Journal of the Royal Statistical Society, Series B*, 31, 524-538.
- Davies, R.B. (2001). Integrated process and the discrete cosine transform, *Journal of Applied Probability*, 38, 105-121.
- Dickey, D.A. and W.A. Fuller (1979). Distribution of estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-431.
- Dufour, J.M. and E.M. Iglesias (2008). Finite-sample and optimal inference in possibly nonstationary general volatility models with Gaussian and heavy-tailed errors, Technical Report, Montreal University, Quebec.
- Dufour, J.M. and M.L. King (1991). Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors, *Journal of Econometrics*, 47, 115-143.
- Dufour, J.M. and M. Neifar (2008). Optimal invariant tests for autoregressive coefficients in linear regressions with Gaussian possibly nonstationary AR(2) errors, Department of Economics, McGill University.
- Dufour, J.M. and A. Taamouti (2010). Exact optimal inference in regression models under heteroscedasticity and non-normality of unknown form, *Computational Statistics and Data Analysis*, 54, 2532-2553.
- Durbin, J. (1960). Estimation of parameters in time series regression models, *Journal of the Royal Statistical Society, Series B*, 22, 139-153.
- El-Bassiouni, M.Y. and H.A. Charif (2004). Testing a null variance ratio in mixed models with zero degrees of freedom for error, *Computational Statistics and Data Analysis*, 46, 707-719.
- Elliott, G. (1999). Efficient tests for a unit root when initial observation is drawn from its unconditional distribution, *International Economic Review*, 40, 767-783.
- Elliott, G. and M. Jansson (2003). Testing for unit roots with stationary covariates, *Journal of Econometrics*, 115, 75-89.
- Elliott, G., M. Jansson and E. Pesavento (2005). Optimal power for testing potential cointegrating vectors with known parameters for nonstationarity, *Journal of Business and Economic Statistics*, 23, 34-48.

- Elliott, G. and U.K. Müller (2006a). Minimizing the impact of the initial condition on testing for unit roots, *Journal of Econometrics*, 135, 285-310.
- Elliott, G. and U.K. Müller (2006b). Efficient tests for general persistent time variation in regression coefficients, *Review of Economic Studies*, 73, 907-940.
- Elliott, G. and U.K. Müller (2014). Pre and post break parameter inference, *Journal of Econometrics*, 180, 141-157.
- Elliott, G., U.K. Müller and M.W. Watson (2012). Nearly optimal tests when a nuisance parameter is present under the null hypothesis, University of California San Deigo.
- Elliott, G. and E. Pesavento (2009). Testing the null of no cointegration when covariates are known to have a unit root, *Econometric Theory*, 25, 1829-1850.
- Elliott, G., T.J. Rothenberg and J.H. Stock (1996). Efficient tests for an autoregressive unit root, *Econometrica*, 64, 813-836.
- Evans, M.A. and M.L. King (1985a). A point optimal test for heteroscedastic disturbances, *Journal of Econometrics*, 27, 163-178.
- Evans, M.A. and M.L. King (1985b). Critical value approximations for tests of linear regression disturbances, *Australian Journal of Statistics*, 27, 68-83.
- Evans, M.A. and M.L. King (1988). A further class of tests for heteroscedasticity, *Journal of Econometrics*, 37, 265-276.
- Ferguson, T.S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York.
- Fraser, A.S., Guttman, I. and G.P.H. Styan (1976). Serial correlation and distributions on the sphere, *Communications in Statistics – Theory and Methods A*, 5, 97-118.
- Franzini, L. and A.C. Harvey (1983). Testing for deterministic trend and seasonal components in time series models, *Biometrika*, 70, 673-82.
- Gallego, J.L. and C. Diaz (2007). Testing for noninvertibility and cointegration in vector autoregressive moving average models, Department of Economics, University of Cantabria.
- Gao, J. and I. Gijbels (2008). Bandwidth selection in nonparametric kernel testing. *Journal of the American Statistical Association*, 103, 1584-1594.
- Gao, J. and M.L. King (2014). Specification testing in parametric trending models with unknown errors. *Advances in Econometrics*, 33, 151-202.

- Gao, J., M.L. King, Z. Lu and D. Tjostheim (2009a). Nonparametric specification testing for nonlinear time series with nonstationarity, *Econometric Theory*, 25, 1869-1892.
- Gao, J., M.L. King, Z. Lu and D. Tjostheim (2009b). Specification testing in nonlinear and nonstationary time series autoregression, *The Annals of Statistics*, 37, 3893-3928.
- Godfrey, L.G. and A.R. Tremayne (1988). Checks of model adequacy for univariate time series models and their application to econometric relationships, *Econometric Reviews*, 7, 1-42.
- Goffe, W.L., G.D. Ferrier and J. Rogers (1994). Global optimization of statistical functions with simulated annealing, *Journal of Econometrics*, 60, 65-99.
- Gourieroux, C. and A. Monfort (1994). Testing nonnested hypotheses, In R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics*, 4, Elsevier: Oxford.
- Gregoir, S. (2006). Efficient tests for the presence of a pair of complex conjugate unit roots in real time series, *Journal of Econometrics*, 130, 45-100.
- Hansen, B.E. (1995). Rethinking the univariate approach to unit root testing: Using covariates to increase power, *Econometric Theory*, 11, 1148-1171.
- Hildreth, C. and J.P. Houck (1968). Some estimators for a linear model with random coefficients, *Journal of the American Statistical Association*, 63, 584-595.
- Hillier, G.H. (1987). Classes of similar regions and their power properties for some econometric testing problems, *Econometric Theory*, 3, 1-44.
- Honda, Y. (1989). On the optimality of some tests of the error covariance matrix in the linear regression model, *Journal of the Royal Statistical Society, Series B*, 51, 71-79.
- Hwang, J. and P. Schmidt (1993). On the power of point optimal tests of the trend stationarity hypothesis, *Economics Letters*, 43, 143-147.
- Hwang, J. and P. Schmidt (1996). Alternative methods of detrending and the power of unit root tests, *Journal of Econometrics*, 71, 227-248.
- Inder, B.A. (1990). A new test for autocorrelation in the disturbances of the dynamic linear regression model, *International Economic Review*, 31, 341-354.
- Jansson, M. (2005). Point optimal tests of the null hypothesis of cointegration, *Journal of Econometrics*, 124, 187-201.

- Juhl, T. and Z. Xiao (2003). Power functions and envelopes for unit root tests, *Econometric Theory*, 19, 240-253.
- 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. (1981a). The alternative Durbin-Watson test: An assessment of Durbin and Watson's choice of test statistic, *Journal of Econometrics*, 17, 51-66.
- King, M.L. (1981b). Testing for random walk disturbances in the linear regression model, Mimeo, Monash University.
- King, M.L. (1983a). Testing for autoregressive against moving average errors in the linear regression model, *Journal of Econometrics*, 21, 35-51.
- King, M.L. (1983b). Testing for moving average regression disturbances, *Australian Journal of Statistics*, 25, 23-34.
- King, M.L. (1984). A new test for fourth-order autoregressive disturbances, *Journal of Econometrics*, 24, 269-277.
- King, M.L. (1985a). A point optimal test for autoregressive disturbances, *Journal of Econometrics*, 27, 21-37.
- King, M.L. (1985b). A point optimal test for moving average regression disturbances, *Econometric Theory*, 1, 211-222.
- King, M.L. (1986). Efficient estimation and testing of regressions with a serially correlated error component, *Journal of Quantitative Economics*, 2, 231-247.
- King, M.L. (1987a). Testing for autocorrelation in linear regression models: A survey. In M.L. King and D.E.A. Giles, eds., *Specification Analysis in the Linear Model*. London: Rutledge and Kegan Paul, 19-73.
- King, M.L. (1987b). An alternative test for regression coefficient stability. *Review of Economics and Statistics*, 69, 379-381.
- King, M.L. (1987c). Towards a theory of point optimal testing, *Econometric Reviews*, 6, 169-218.
- King, M.L. (1989). Testing for fourth-order autocorrelation in regression disturbances when first-order autocorrelation is present, *Journal of Econometrics*, 41, 285-301.
- King, M.L. (1996). Hypothesis testing in the presence of nuisance parameters, *Journal of Statistical Planning and Inference*, 50, 103-120.

- King, M.L. (1998). The likelihood ratio test of non-nested linear regression models, *Applied Statistical Science*, 3, 365-375.
- King, M.L. and P.M. Edwards (1989). Transformations for an exact goodness-of-fit test of structural change in the linear regression model, *Empirical Economics*, 14, 113-121.
- King, M.L. and M.A. Evans (1988). Locally optimal properties of the Durbin-Watson test, *Econometric Theory*, 4, 509-516.
- King, M.L. and D.E.A. Giles (1984). Autocorrelation pre-testing in the linear model: Estimation, testing and prediction, *Journal of Econometrics*, 25, 35-48.
- King, M.L. and D.C. Harris (1995). The application of the Durbin-Watson test to the dynamic regression model under normal and non-normal errors, *Econometric Reviews*, 14, 487-510.
- King, M.L. and G.H. Hillier (1985). Locally best invariant tests of the error covariance matrix of the linear regression model, *Journal of the Royal Statistical Society, Series B*, 47, 98-102.
- King, M.L. and M. McAleer (1987). Further results on testing AR(1) against MA(1) disturbances in the linear regression model, *The Review of Economic Studies*, 54, 649-663.
- King, M.L. and S. Rahman (2015). Robustness of marginal likelihood based tests for random regression coefficients, Department of Econometrics and Business Statistics, Monash University.
- King, M.L. and T.S. Shively (1993). Locally optimal testing when a nuisance parameters is present only under the alternative, *The Review of Economics and Statistics*, 75, 1-7.
- King, M.L. and M.D. Smith (1986). Joint one-sided test of linear regression coefficients, *Journal of Econometrics*, 32, 367-383.
- King, M.L. and P.X. Wu (1991). Small-disturbance asymptotics and the Durbin-Watson and related tests in the dynamic regression model, *Journal of Econometrics*, 47, 145-152.
- King, M.L. and P.X. Wu (1997). Locally optimal one-sided tests for multiparameter hypotheses, *Econometric Reviews*, 16, 131-156.
- Kurozumi, E. (2003). Some properties of the point optimal invariant test for the constancy of parameters, *Journal of the Japan Statistical Society*, 33, 169-180.

- Kurozumi, E. and Y. Arai (2005). Point optimal test for cointegration with unknown variance-covariance matrix, Discussion Paper 2005-08, Graduate School of Economics, Hitotsubashi University.
- Lee, D.J. (2009). Parametric and semiparametric efficient tests for parameter instability, Working Paper 2008-40 R, Department of Economics, University of Connecticut.
- Lehmann, E.L. and J.P. Romano (2005). Testing Statistical Hypotheses, 3rd edition, Springer, New York.
- Liu, H. and G. Rodriguez (2006). Unit root tests using GLS detrended data and covariates in structural change models, Department of Economics, University of Ottawa.
- MacKinnon, J. (1983). Model specification against non-nested alternatives, *Econometric Reviews*, 2, 85-111.
- Martellosio, F. (2010). Power properties of invariant tests for spatial autocorrelation in linear regression, *Econometric Theory*, 26, 152-186.
- Martellosio, F. (2012). Testing for spatial autocorrelation: The regressors that make the power disappear, *Econometric Reviews*, 31, 215-240.
- McAleer, M. (1987). Specification tests for separate models: A survey, In M.L. King and D.E.A. Giles (eds.), *Specification Analysis in the Linear Model: Essays in honour of Donald Cochrane*. Routledge and Kegan Paul; London, 146-196.
- Moon, H.R., B. Perron, P.C.B. Phillips (2007). Incidental trends and the power of panel unit root tests, *Journal of Econometrics*, 141, 416-459.
- Moon, H.R., B. Perron and P.C.B. Phillips (2014). Point optimal panel unit root tests with serially correlated errors, *Econometrics Journal*, 17, 338-372.
- Müller, U.K. (2005). Size and power of tests of stationarity in highly autocorrelated time series, *Journal of Econometrics*, 128, 195-213.
- Müller, U.K. (2011). Efficient tests under a weak convergence assumption, *Econometrica*, 79, 395-435.
- Müller, U.K. (2014). HAC corrections for strongly autocorrelated time series, *Journal of Business and Economic Statistics*, 32, 311-322.
- Müller, U.K. and G. Elliott (2003). Tests for unit roots and the initial condition, *Econometrica*, 71, 1269-1286.

- Müller, U.K. and M.K. Watson (2013). Low-frequency robust cointegration testing, *Journal of Econometrics*, 174, 66-81.
- Nakatsuma, T., E. Gousskova, J. Uemura and H. Tsurumi (2000). Ratio tests of a unit root, *Communications in Statistics – Theory and Methods*, 29, 2547-2571.
- Ng, S. and P. Perron (2001). Lag length selection and the construction of unit root tests with good size and power, *Econometrica*, 69, 1519-1554.
- Nyblom, J. (1986). Testing for deterministic linear trend in time series, *Journal of the American Statistical Association*, 81, 545-549.
- Pesaran, M.H. and M. Weeks (2003). Nonnested hypothesis testing: An overview. In B.H. Baltagi (eds.), *A Companion to Theoretical Econometrics*, Blackwell Publishing: Malden, 279-309.
- Perron, P. and G. Rodriguez (2003). GLS detrending efficient unit root tests and structural change, *Journal of Econometrics*, 115, 1-27.
- Perron, P. and S. Ng (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties, *Review of Economics Studies*, 63, 435-463.
- Ploberger, W. (2004). A complete class of tests when the likelihood is locally asymptotically quadratic, *Journal of Econometrics*, 118, 67-94.
- Rahman, S. and M.L. King (1994). A comparison of marginal likelihood based and approximate point optimal tests for random regression coefficients in the presence of autocorrelation, *Pakistan Journal of Statistics*, 10, 375-394.
- Rodrigues, P.M.M. and A.M.R. Taylor (2007). Efficient tests of the seasonal unit root hypothesis, *Journal of Econometrics*, 141, 548-573.
- Rosenberg, B. (1973). The analysis of a cross-section of time series by stochastically convergent parameter regression, *Annals of Economic and Social Measurement*, 2, 399-428.
- Rothenberg, T.J. and J.H. Stock (1997). Inference in a nearly integrated autoregressive model with nonnormal innovations, *Journal of Econometrics*, 80, 269-286.
- Saikkonen, P. and R. Luukkonen (1993a). Testing for a moving average unit root in autoregressive integrated moving average models, *Journal of the American Statistical Association*, 88, 596-601.
- Saikkonen, P. and R. Luukkonen (1993b). Point optimal tests for testing the order of differencing in ARIMA models, *Econometric Theory*, 9, 343-362.

- Sargan, J.D. and A. Bhargava (1983). Maximum likelihood estimation of regression models with first order moving average errors when the root lies on the unit circle, *Econometrica*, 51, 799-820.
- Schaafsma, W. and L.J. Smid (1966). Most stringent somewhere most powerful tests against alternatives restricted by a number of linear inequalities, *The Annals of Mathematical Statistics*, 37, 1161-1172.
- Sen Gupta, A. and L. Vermeire (1986). Locally optimal tests for multiparameter hypotheses, *Journal of American Statistical Association*, 81, 819–825.
- Sen Gupta, A. (1991). A review of optimality of multivariate tests, *Statistics and Probability Letters*, 12, 527-535.
- Shively, P.A. (2001). Trend-Stationary GNP: Evidence from a new exact point wise most powerful invariant unit root test, *Journal of Applied Econometrics*, 16, 537-551.
- Shively, P.A. (2004). Testing for a unit root in ARIMA processes, *Journal of Applied Statistics*, 31, 785-798.
- Shively, T.S. (1986). Issues in the analysis of stochastic coefficient regression models. Unpublished Ph.D. thesis, University of Chicago.
- Shively, T.S. (1988a). An analysis of tests for regression coefficient stability, *Journal of Econometrics*, 39, 367-386.
- Shively, T.S. (1988b). An exact test for a stochastic coefficient in a time series regression model, *Journal of Time Series Analysis*, 9, 81-88.
- Shively, T.S. (1993). Testing for autoregressive disturbances in a time series regression with missing observations, *Journal of Econometrics*, 57, 233-255.
- Shively, T.S. (2006). Point optimal invariant tests, In *Encyclopedia of Statistical Sciences*, S. Kotz, C.B. Read and D.L. Banks (eds.), Wiley and Sons, 586-589.
- Shively, T.S., C.F. Ansely and R. Kohn (1990). Fast evaluation of the distribution of the Durbin-Watson and other invariant test statistic in time series regression, *Journal of the American Statistical Association*, 85, 676-685.
- Shively, T.S., R. Kohn and C.F. Ansely (1994). Testing for nonlinearity in a semiparametric regression model, *Journal of Econometrics*, 64, 77-96.
- Silvapulle, P. and M.L. King (1991). Testing moving average against autoregressive disturbances in the linear-regression model, *Journal of Business and Economic Statistics*, 9, 329–335.

- Silvapulle, P. and M.L. King (1993). Nonnested testing for autocorrelation in the linear regression model, *Journal of Econometrics*, 58, 295-314.
- Silvapulle, P. (1994). Testing for AR(1) against IMA(1,1) disturbances in the linear regression model, *Communication in Statistics - Theory and Methods*, 23, 701–720.
- Small, J.P. (1993). The limiting power of point optimal autocorrelation tests, *Communications in Statistics – Theory and Methods*, 22, 2463-2470.
- Spjotvoll, E. (1967). Optimum invariant tests in unbalanced variance components models, *The Annals of Mathematical Statistics*, 38, 422-428.
- Sriananthakumar, S. (2013). Testing linear regression model with AR(1) errors against a first-order dynamic linear regression model with white noise errors: A point optimal testing approach, *Economic Modelling*, 33, 126–136.
- Sriananthakumar, S. and M.L. King (2006). A new approximate point optimal test of a composite null hypothesis, *Journal of Econometrics*, 130, 101–122.
- Sun, Y., P.C.B. Phillips and S. Jin (2008). Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing, *Econometrica*, 76, 175-194.
- Van der Vaart, A.W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.
- Vougas, D.V. (2009). Modification of the point optimal unit root test, *Applied Economics Letters*, 16, 349-352.
- Wang, L. (2014). Bootstrap point optimal unit root tests, *Journal of Time Series Econometrics*, 6, 1–31.
- White, H. (1982). Regularity conditions for Cox’s test of nonnested hypotheses, *Journal of Econometrics*, 19, 301-318.
- Wooldridge, J.M. (1989). A computationally simple heteroskedasticity and serial correlation robust standard error for the linear regression model, *Economics Letters*, 31, 239-243.
- Wu, P.X. and M.I. Bhatti (1994). Testing for block effects and misspecification in regression models based on survey data, *Journal of Statistical Computation and Simulation*, 50, 75-90.
- Wu, P.X. and M.L. King (1996). Small-sample power of tests for inequality restrictions: The case of quarter-dependent regression errors, *Economics Letters*, 52, 121-127.

Xiao, Z. (2001). Likelihood-based inference in trending time series with a root near unity, *Econometric Theory*, 17, 1082-1112.