IOWA STATE UNIVERSITY

Powerful Trend Function Tests That are Robust to Strong Serial Correlation with an Application to the Prebisch-Singer Hypothesis

Helle Bunzel, Timothy J. Vogelsang

April 2003

Working Paper # 03018

Department of Economics Working Papers Series

Ames, Iowa 50011

Powerful Trend Function Tests That are Robust to Strong Serial Correlation with an Application to the Prebisch-Singer Hypothesis

Helle Bunzel
Iowa State University and CentER, Tilburg

Timothy J. Vogelsang*
Cornell University

April 28, 2003, Revised September 21, 2004

Abstract

In this paper we propose tests for hypotheses regarding the parameters of the deterministic trend function of a univariate time series. The tests do not require knowledge of the form of serial correlation in the data and they are robust to strong serial correlation. The data can contain a unit root and the tests still have the correct size asymptotically. The tests we analyze are standard heteroskedasticity autocorrelation (HAC) robust tests based on nonparametric kernel variance estimators. We analyze these tests using the fixed-b asymptotic framework recently proposed by Kiefer and Vogelsang (2002). This analysis allows us to analyze the power properties of the tests with regards to bandwidth and kernel choices. Our analysis shows that among popular kernels, there are specific kernel and bandwidth choices that deliver tests with maximal power within a specific class of tests. Based on the theoretical results, we propose a data dependent bandwidth rule that maximizes integrated power. Our recommended test is shown to have power that dominates a related test proposed by Vogelsang (1998). We apply the recommended test to the logarithm of a net barter terms of trade series and we find that this series has a statistically significant negative slope. This finding is consistent with the well known Prebisch-Singer hypothesis.

Keywords: HAC Estimator, Fixed-b Asymptotics, Power Envelope, Unit Root, Nearly Integrated, Partial Sum, Deterministic Trend, Linear Trend, Data Dependent Bandwidth.

*Corresponding Author: Tim Vogelsang, Department of Economics, Uris Hall, Cornell University, Ithaca, NY 14853-7601, Phone: 607-255-5108, Fax: 607-255-2818, email: tjv2@cornell.edu

1 Introduction

In this paper we propose tests of linear hypotheses on the parameters in a univariate deterministic trend model. The tests are designed to be size-robust to strong serial correlation in the errors including the case of a unit root in the errors. Robustness to serial correlation is obtained using well known nonparametric heteroskedasticity autocorrelation (HAC) robust standard errors. Tests using the HAC robust standard errors may still have significant size distortion, however. One source of this distortion is the fact that the finite sample distributions of HAC robust tests are highly dependent on the choice of bandwidth and kernel, while the asymptotic distributions of the tests do not depend on these choices. Using the newly developed fixed bandwidth (fixed-b) asymptotics of Kiefer and Vogelsang (2002) we develop an asymptotic theory that captures the choice of kernel and bandwidth. Fixed-b asymptotics can be used to reduce some of the size distortions. The second source of size distortion is the possibility of strong serial correlation (possibly a unit root) in the errors. We show that the fixed-b asymptotic distributions are free of serial correlation nuisance parameters regardless of the bandwidth or kernel used to compute the HAC robust standard errors. This asymptotic pivotal result holds for stationary errors as well as nearly integrated errors although the limiting distributions are different in the two cases. Using this asymptotically pivotal property, we are able to control the over-rejection problem caused by strong serial correlation by implementing the scaling correction approach proposed by Vogelsang (1998). Therefore, the tests we propose have well behaved size even when the errors have strong serial correlation.

For the special case of the simple linear trend model, we use a local asymptotic power analysis to guide the choice of kernel and bandwidth. Confining attention to tests with asymptotically correct size, we consider a class of well known and popular kernels and we compute asymptotic power envelopes that represent maximal power across the kernels and bandwidths within the class. We then show that tests based on the Daniell kernel have power that effectively attains the power envelope. We address the traditionally difficult issue of HAC bandwidth choice using fixed-b asymptotics in conjunction with local to unity asymptotics. For a given value of the local to unity parameter, we numerically determine the bandwidth that maximizes integrated power. To our knowledge, this is the first detailed HAC bandwidth analysis that focuses on power of tests rather than the mean square error of the HAC estimate although Hall (2004) does point out a potential link between bandwidth choice and power of over-identifying restrictions tests in generalized method of moments models. Our analysis provides a data dependent bandwidth that maximizes integrated power. Finite sample simulations suggest that a feasible version of this data dependent bandwidth rule works well in practice. Our asymptotic and finite sample analysis points to one HAC based test that we recommend in practice.

We compare our recommended test with the related tests proposed by Vogelsang (1998). One of those tests, the $t-PSW$ test, is very similar to the tests analyzed in this paper. We show that the $t-PSW$ test is dominated in terms of power by the recommended test. Therefore, the tests analyzed in this paper are an improvement over the tests proposed by Vogelsang (1998).

We use the recommended test to investigate the well known Prebisch (1950) and Singer (1950) hypothesis that postulates that over time the net barter terms of trade should be declining between countries that primarily export commodities and countries that primarily export manufactures. This empirical conjecture has received considerable attention in the international economics literature. See Ardeni and Wright (1992), Cuddington and Urzua (1989), Grilli and Yang (1988), Lutz (1999), Powell (1991), Sapsford (1985), Spraos (1980) and Trivedi (1995) among others. The empirical results in this literature have been mixed. Many authors have interpreted evidence in support of the Prebisch-Singer hypothesis with caution because of the potential over-rejection problem caused by strong serial correlation/unit root in the errors. In fact, many authors have focused on, and in our opinion been distracted by, the question as to whether or not the innovations have a unit root or are stationary. Because a time series can have a decreasing deterministic trend whether the innovations are stationary or have a unit root, the unit root issue is simply a nuisance parameter in the context of the Prebisch-Singer hypothesis. One advantage of our approach is that it allows a direct test on the slope coefficient of the linear trend that is robust to the unit root question. When applied to the net barter terms of trade series of Grilli and Yang (1988) as extended by Lutz (1999) we find strong and consistent evidence to support the Prebisch-Singer hypothesis. Our results are not subject to the usual "over-rejection problem" critique because of the robust properties of the tests. Further tests indicate that the trend function of this series is stable over time. Our results confirm what many authors have been saying for over 20 years: Prebisch and Singer were right!

The rest of the paper is organized as follows. In Section 2 the trend function model is described in detail, the required assumptions are stated, and some of the basic asymptotic results are presented. Section 3 describes the scaling procedure that is used to control the over-rejection problem caused by strong serial correlation. Section 4 briefly describes the $t-PSW$ test proposed by Vogelsang (1998). In Section 5 we derive and discuss asymptotic results obtained under the new fixed-b asymptotics. In Section 6 we examine the asymptotic properties of the test statistics in the simple linear trend model. We compute asymptotic power envelopes and determine kernels and bandwidths that deliver tests with power close to the envelopes. Section 7 proposes a feasible data dependent bandwidth rule that builds on the theoretical results. In Section 8 the results of some finite sample simulation experiments are reported. The empirical results on the Prebisch-Singer hypothesis are given in Section 9. Section 10 concludes and proofs of important results are collected in the appendix.

2 The Model Setup

We are interested in the following model of a time series with deterministic trends:

$$y_t = f(t)' \beta + u_t, \qquad t = 1, ..., T, \tag{1}$$

where $f(t)$ denotes a $(k \times 1)$ vector of trend functions, β is a $(k \times 1)$ vector of parameters, and " $'$ " denotes the transpose, when used in the context of a vector. This type of model is used frequently in macroeconomics and finance to determine the composition of univariate time series. When performing tests on β, for example to determine whether a given trend should be included, the presence of serial correlation and heteroskedasticity in the errors must be taken into account. In this paper, we will concern ourselves with the situation where the exact error structure is not of interest. In that case, there is no need to model the error structure explicitly, as hypotheses regarding the coefficients on the trends can be tested without doing so. Such testing is virtually always done by using HAC estimators to estimate the asymptotic variance of the parameter estimates, and we follow that approach in this paper.

Throughout the paper we assume that u_t is a scalar, mean zero time series. The time series process $\{u_t\}$ is allowed to have serial correlation and may be stationary or have a unit root or a root close to one. For the purpose of studying the impact of these various error specifications on the testing procedures, we make the following flexible assumptions about u_t.

Assumption 1

$$u_t = \alpha u_{t-1} + \varepsilon_t, \qquad t = 2, 3, ..., T, \qquad u_1 = \varepsilon_1,$$

$$\varepsilon_t = d(L) e_t, \quad d(L) = \sum_{i=0}^{\infty} d_i L^i, \quad \sum_{i=0}^{\infty} i |d_i| < \infty, \quad d(1)^2 > 0,$$

where $\{e_t\}$ is a martingale difference sequence with $E\left(e_t^2 | e_{t-1}, e_{t-2}, ...\right) = 1$ and $\sup_t E\left(e_t^4\right) < \infty$. Under this specification, the errors are stationary when $|\alpha| < 1$. In this case, α is not modeled as a function of the sample size. Alternatively, the errors can be modeled as nearly integrated by letting $\alpha = \left(1 - \frac{\overline{\alpha}}{T}\right)$ where $\overline{\alpha} = 0$ corresponds to a pure unit root process.

Under Assumption 1 the following functional central limit theorems follow from well known results (see Chan and Wei (1988), Phillips (1987) and Phillips and Solo (1992)):

$$T^{-1/2} \sum_{t=1}^{[rT]} u_t \Rightarrow \sigma w(r) \quad \text{if} \quad |\alpha| < 1$$

$$T^{-1/2} u_{[rT]} \Rightarrow d(1) V_{\overline{\alpha}}(r) \quad \text{if} \quad \alpha = 1 - \frac{\overline{\alpha}}{T},$$

where $\sigma^2 = d(1)^2/(1-\alpha)^2$, $w(r)$ is a standard Wiener process, $V_{\overline{\alpha}}(r) = \int_0^r \exp(-\overline{\alpha}(r-s))\, dw(s)$ and $\Rightarrow$ denotes weak convergence.

At times it will be useful to stack the equations in (1) and rewrite them as

$$y = \mathbf{f}\left(T\right)\beta + u. \tag{2}$$

Here $\mathbf{f}\left(T\right)$ is the $(T \times k)$ stacked vector of trend functions. The following assumptions on the trend are sufficient to obtain the main results of the paper:

Assumption 2 $f(t)$ includes a constant, there exists a $(k \times k)$ diagonal matrix τ_T and a vector of functions F, such that $\tau_T f\left(t\right) = F\left(\frac{t}{T}\right) + o\left(1\right)$, $\int_0^1 F_i\left(r\right)dr < \infty$, $i = 1, ..., k$, and $\det\left[\int_0^1 F\left(r\right)F\left(r\right)' dr\right] > 0$.

Assumption 2 is fairly standard and it is essentially the same as the one used by Vogelsang (1998). We include the additional assumption that an intercept is included in the model.

Model (1) is estimated using Ordinary Least Squares (OLS) and $\widehat{\beta} = \left(\mathbf{f}(T)'\mathbf{f}(T)\right)^{-1}\mathbf{f}(T)'y$ denotes the OLS estimate of β, while $\widehat{u} = y - \mathbf{f}(T)\widehat{\beta}$ denotes the OLS residuals. The limiting distribution of $\widehat{\beta}$ is well known for both stationary and unit root errors:

$$T^{1/2}\tau_T^{-1}\left(\widehat{\beta} - \beta\right) \Rightarrow \sigma\left(\int_0^1 F(r)F(r)'dr\right)^{-1}\int_0^1 F\left(r\right)dw\left(r\right) \quad \text{if} \quad |\alpha| < 1,$$

$$T^{-1/2}\tau_T^{-1}\left(\widehat{\beta} - \beta\right) \Rightarrow d(1)\left(\int_0^1 F(r)F(r)'dr\right)^{-1}\int_0^1 F\left(r\right)V_{\overline{\alpha}}\left(r\right)dr \quad \text{if} \quad \alpha = 1 - \frac{\overline{\alpha}}{T}.$$

Notice that when the errors are stationary, the only unknown nuisance parameter in the limiting distribution is σ. The fact that a single nuisance parameter appears in the limiting distribution of the OLS estimates occurs because the regressors are deterministic. In a regression model with random regressors, the asymptotic variance of the OLS estimates depends on a zero-frequency spectral density matrix with rank equal to the number of regression parameters. In that case, the HAC robust standard errors are computed using a vector of time series comprised of products of the regressors and OLS residuals.

When the errors are nearly integrated, the only unknown nuisance parameters are $d(1)$ and $\overline{\alpha}$. The dependence of the tests on $\overline{\alpha}$ when the errors are nearly integrated is the reason that HAC robust tests tend to be over-sized in practice when errors have strong serial correlation. We control the over-rejection problem using the scaling factor approach proposed by Vogelsang (1998).

To construct the usual HAC robust t or Wald tests, an estimator of σ^2 is often used. We consider the case where σ^2 is estimated nonparametrically using the OLS residuals, $\widehat{u}_t$:

$$\widehat{\sigma}^2 = \widehat{\gamma}_0 + 2\sum_{j=1}^{T-1} k\left(j/M\right)\widehat{\gamma}_j, \tag{3}$$

where $\widehat{\gamma}_j = T^{-1}\sum_{t=j+1}^{T}\widehat{u}_t\widehat{u}_{t-j}$ and $k\left(x\right)$ is a kernel function satisfying $k\left(x\right) = k\left(-x\right)$, $k\left(0\right) = 1$, $|k\left(x\right)| \leq 1$, $k\left(x\right)$ continuous at $x = 0$ and $\int_0^1 k^2\left(x\right)dx < \infty$. M is called the bandwidth or the truncation lag. For $\widehat{\sigma}^2$ to be consistent, it is necessary to downweight or eliminate the sample autocovariances for high values of j. Specifically, it is necessary that $M \to \infty$ and $M/T \to 0$ as $T \to \infty$. Most commonly used kernel functions have the property that $k(x) = 0$ for $|x| > 1$, effectively eliminating the sample autocovariances for all values of j greater than M, inspiring the name truncation lag.

We are interested in testing hypotheses of the form $H_0 : R\beta = d$ where R is a $q \times k$ matrix of known constants, d is a $q \times 1$ vector of known constants and $q \leq k$. Typically, R is a matrix selecting single entries of β, and d is a vector of zeros, but we maintain the hypothesis in its general form. As a rule, the test statistics used to test this type of hypothesis on the trend function are either t or Wald statistics of the form:

$$W_T = \left(R\widehat{\beta} - r\right)' \left[\widehat{\sigma}^2 R \left(\mathbf{f}(T)' \mathbf{f}(T)\right)^{-1} R'\right]^{-1} \left(R\widehat{\beta} - r\right)$$

$$t = \frac{R_1\widehat{\beta} - r}{\sqrt{\widehat{\sigma}^2 R_1 \left(\mathbf{f}(T)' \mathbf{f}(T)\right)^{-1} R_1'}},$$

where the subscript in R_1 signifies that the restriction matrix is a vector in the case of the $t-$test. If the errors are stationary and $\widehat{\sigma}^2$ is a consistent estimator, then the t test has a standard normal limiting distribution and W has a limiting chi-square distribution. Unfortunately, when there is strong serial correlation in the errors, these standard asymptotic approximations are often inaccurate and the tests suffer from severe over-rejection problems (see Vogelsang (1998, Table I)). In addition, the finite sample behavior of the tests are sensitive to the choice of bandwidth and kernel, yet the standard asymptotics is the same regardless of the kernel or bandwidth. We address both of these issues. We control the over-rejection problem using a scaling factor proposed by Vogelsang (1998). We address the bandwidth and kernel problem by deriving the limiting distributions of the scaled tests under the fixed-b asymptotic framework proposed by Kiefer and Vogelsang (2002).

3 Scaled Statistics

We now describe the scaling procedure proposed by Vogelsang (1998) and introduce a new variant of the approach. The basic idea is to multiplicatively scale the t and W tests by a factor that converges to one when the errors are stationary but converges to a nuisance parameter free random variable when the errors have a unit root. We consider two scaling factors based on two unit root tests. Let J denote the unit root test proposed by Park (1990) and Park and Choi (1988). Consider the regression

$$y_t = f(t)' \beta + \sum_{i=p}^{9} \alpha_i t^i + u_t, \tag{4}$$

where t^{p-1} is the highest order polynomial of t included in $f(t)$. Then the J statistic is defined as

$$J = \frac{SSR_{(1)} - SSR_{(4)}}{SSR_{(4)}},$$

where $SSR_{(4)}$ is the sum of squared residuals obtained from the estimation of (4) by OLS, and $SSR_{(1)}$ be the sum of squared residuals from the OLS estimation of (1). The second unit root test is the test proposed by Breitung (2002) defined as

$$BG = \frac{T^{-2} \sum_{t=1}^{T} \widehat{S}_t^2}{SSR_{(1)}},$$

where $\widehat{S}_t = \sum_{j=1}^{t} \widehat{u}_j$ are the partial sums of the OLS residuals from Model (1). Both the J and BG statistics share the property that they converge to zero when the errors are stationary. When the errors are nearly integrated, the asymptotic distributions of J and BG are non-degenerate and depend on $\overline{\alpha}$ but otherwise do not depend on nuisance parameters such as $d(1)$.

Let UR generically denote either J or BG and let c denote a constant. The scaling factor

$$\exp(-c\, UR),$$

converges to a well defined random variable when the errors have a unit root but converges to one when the errors are stationary. Using the scaling factor we now redefine the t and W statistics as

$$t = \left(\frac{R_1 \widehat{\beta} - q}{\sqrt{\widehat{\sigma}^2 R_1 \left(\mathbf{f}(T)' \mathbf{f}(T) \right)^{-1} R_1'}} \right) \exp\left(-c\, UR\right),$$

$$W_T = \left(\left(R\widehat{\beta} - q \right)' \left[\widehat{\sigma}^2 R \left(\mathbf{f}(T)' \mathbf{f}(T) \right)^{-1} R' \right]^{-1} \left(R\widehat{\beta} - q \right) \right) \exp\left(-c\, UR\right). \tag{5}$$

The limiting distributions of t and W are unaffected by the scaling when the errors are stationary. When the errors are nearly integrated, the scaling factor affects the limiting distribution. Given $\overline{\alpha}$, for a specific percentage point, it is possible to compute the constant c such that the asymptotic critical values of each statistic are the same for stationary errors and nearly integrated errors. We follow Vogelsang (1998) and compute c for the case of $\overline{\alpha} = 0$. By making the asymptotic critical values the same for stationary errors and unit root errors, the scaling factors solve the over-rejection problem caused by strong serial correlation in the errors.

The versions of the statistics given by (5) will be used for the remainder of the paper. Note that the value of c used in practice depends on the significance level of the test and depends on the unit root statistic used for the scaling factor. A detailed discussion of the choice of c is given below for the simple linear trend model.

4 Comparisons to the PSW Test

The HAC robust tests defined by (5) are very closely related to the PSW and $t - PSW$ tests proposed by Vogelsang (1998). The PSW and $t-PSW$ tests are defined in the same manner except that $\widehat{\sigma}^2$ is replaced by $T^{-1}s_z^2$ where s_z^2 is the OLS error variance estimator from the regression

$$z_t = g(t)'\beta + S_t, \tag{6}$$

where $z_t = \sum_{j=1}^{t} y_j$, $g(t) = \sum_{j=1}^{t} f(j)$ and $S_t = \sum_{j=1}^{t} u_j$, i.e. from the regression obtained by computing the partial sums of regression (1). Tedious algebra can be used to show that $T^{-1}s_z^2$ is closely related to the Bartlett kernel estimator of σ^2 using residuals $\widetilde{u}_t = y_t - f(t)'\widetilde{\beta}$ where $\widetilde{\beta}$ is the OLS estimate from (6). Therefore, $T^{-1}s_z^2$ is of the same stochastic order as $\widehat{\sigma}^2$. In the case of the simple linear regression model (see Section 6), we compare the asymptotic size and power of the $t - PSW$ test and show that it is dominated by the Daniel kernel HAC robust test. Details are given in Section 6.

Vogelsang (1998) also proposed a second test labeled $t - PS$ based on the OLS estimates of regression (6). We do not provide comparisons to the $t - PS$ test given that the $t - PSW$ test has higher power than $t - PS$ when the errors are stationary.

5 Limiting Distributions Under Fixed-b Asymptotics

In this section we provide the limiting null distributions of t and W as defined in (5) under the assumption that $M = bT$ where $b \in (0, 1]$. This asymptotic nesting for the bandwidth was proposed by Kiefer and Vogelsang (2002) and results were obtained for stationary models estimated by generalized method of moments. The results in Kiefer and Vogelsang (2002) do not apply to parameters associated with deterministic trends, nor to errors that contain unit roots. Therefore, the results given here are new.

Before we proceed, some additional notation and definitions are required. As is well known, estimators of coefficients on different trends will often converge at different rates. Specifically, the coefficients entering the constraint which converge the slowest will dominate the asymptotic distribution. In order to formalize this, let μ_i be the largest non-positive power of time, t, in the nonzero elements in the i'th row of $R\tau_T$. Then define the $q \times q$ diagonal matrix A in such a way that $A_{ii} = T^{\mu_i}$, and let $R^* = \lim_{T\to\infty} A^{-1}R\tau_T$. In the case when $q = 1$ we use R_1^* to denote R^*. Under fixed-b asymptotics, the limiting distributions depend on the type of kernel used in computing $\widehat{\sigma}^2$. The following definition describes the types of kernel we analyze.

Definition 1 *A kernel is labelled Type 1 if $k(x)$ is twice continuously differentiable everywhere and as a Type 2 kernel if $k(x)$ is continuous, $k(x) = 0$ for $|x| \geq 1$ and $k(x)$ is twice continuously differentiable everywhere except at $|x| = 1$.*

In addition to kernels which fall in these two categories, we consider the Bartlett kernel (which is neither Type 1 or 2) separately.

The limiting distributions are expressed in terms of the following functions and random variables.

Definition 2

$$N^F = \begin{cases} \int_0^1 F(s)\,dw(s), & \text{if } |\alpha| < 1 \\ \int_0^1 F(s)\,V_{\overline{\alpha}}(s)\,ds, & \text{if } \alpha = 1 - \frac{\overline{\alpha}}{T} \end{cases}$$

$$H(r) = \begin{cases} w(r) & \text{if } |\alpha| < 1 \\ \int_0^r V_{\overline{\alpha}}(s)ds & \text{if } \alpha = 1 - \frac{\overline{\alpha}}{T} \end{cases}$$

$$Q^F(r) = H(r) - \int_0^r F(s)'\,ds \left(\int_0^1 F(s)\,F(s)'\,ds \right)^{-1} N^F$$

$$k^*(x) = k\left(\frac{x}{b}\right),$$

$k_-^{*'}$ is the first derivative of k^* from below

$$\Phi^F(b,k) = \begin{cases} \int_0^1 \int_0^1 -k^{*\prime\prime}(r-s)\,Q^F(r)\,Q^F(s)'\,drds & \text{if } k(x) \text{ is Type 1} \\ \int\int_{|r-s|<b} -k^{*\prime\prime}(r-s)\,Q^F(r)\,Q^F(s)\,drds & \\ \quad +2k_-^{*\prime}(b)\int_0^{1-b} Q^F(r+b)\,Q^F(r)\,dr & \text{if } k(x) \text{ is Type 2} \\ \frac{2}{b}\int_0^1 Q^F(r)^2\,dr - \frac{2}{b}\int_0^{1-b} Q^F(r+b)\,Q^F(r)\,dr & \text{if } k(x) \text{ is Bartlett} \end{cases}$$

In the case of nearly integrated errors, the limiting distributions of the tests depend on the limiting distributions of the unit root tests used in the scaling factors. Let $\widehat{V}_{\overline{\alpha}}(r)$ denote the residuals from the projection of $V_{\overline{\alpha}}(r)$ onto the space spanned by $F(r)$, and let $V_{\overline{\alpha}}^*(r)$ denote the residuals from the projection of $V_{\overline{\alpha}}(r)$ onto the space spanned by $\left(F(r)', r^p, r^{p+1}, ..., r^9\right)'$. The following lemma follows directly from Park (1990), Park and Choi (1988) and Breitung (2002).

Lemma 1 *Suppose Assumptions 1 and 2 hold. If $|\alpha| < 1$, then as $T \to \infty$, $J \Rightarrow 0, BG \Rightarrow 0$. If $\alpha = 1 - \frac{\overline{\alpha}}{T}$, then as $T \to \infty$,*

$$J \Rightarrow \frac{\int_0^1 \widehat{V}_{\overline{\alpha}}(r)^2\,dr - \int_0^1 V_{\overline{\alpha}}^*(r)^2\,dr}{\int_0^1 V_{\overline{\alpha}}^*(r)^2\,dr}$$

$$BG \Rightarrow \frac{\int_0^1 Q^F(r)^2\,dr}{\int_0^1 \widehat{V}_{\overline{\alpha}}(r)^2\,dr}.$$

We generically denote these limiting distributions by UR_∞ in what follows.

We can now state the main theorem.

Theorem 1 *Let $M = bT$, $b \in (0,1]$. Then under Assumptions 1 and 2 as $T \to \infty$*
a)

$$\widehat{\sigma}^2 \Rightarrow \sigma^2 \Phi^F(b,k) \quad \text{if } |\alpha| < 1,$$
$$T^{-2}\widehat{\sigma}^2 \Rightarrow d(1)^2 \Phi^F(b,k) \quad \text{if } \alpha = 1 - \frac{\overline{\alpha}}{T},$$

b)

$$W_T \Rightarrow \left(R^* \left(\int_0^1 F(s) F(s)' ds \right)^{-1} N^F \right)' \left[\Phi^F(b,k) R^* \left(\int_0^1 F(s) F(s)' ds \right)^{-1} R^{*\prime} \right]^{-1}$$
$$\times \left(R^* \left(\int_0^1 F(s) F(s)' ds \right)^{-1} N^F \right) \exp(-cUR_\infty),$$

c)

$$t \Rightarrow \left(\frac{R_1^* \left(\int_0^1 F(s) F(s)' ds \right)^{-1} N^F}{\sqrt{\Phi^F(b,k) R_1^* \left(\int_0^1 F(s) F(s)' ds \right)^{-1} R_1^{*\prime}}} \right) \exp(-cUR_\infty).$$

Theorem 1 demonstrates that pivotal test statistics are obtained under fixed-b asymptotics regardless of kernel or bandwidth, although the limiting distributions of the test statistics depend upon the choice of kernel and bandwidth. The limiting distributions are clearly different when the errors are stationary compared to when the errors are nearly integrated. For each combination of kernel, bandwidth, scaling factor and percentage point, c can be chosen so that the critical values are the same for both stationary errors and unit root errors ($\overline{\alpha} = 0$). The critical values corresponding to the asymptotic distributions in Theorem 1 along with the values of c are simple to compute numerically. A power analysis in the next section indicates specific kernels and bandwidth values that lead to tests with optimal power properties in a model with a simple linear trend. Critical values and details of their computation are given for the recommended tests in the simple linear trend model following a discussion of power.

6 Optimal Kernels and Bandwidths in the Simple Linear Trend Model

In this section extensive analysis of local asymptotic size and power of the simple linear trend model is provided. We focus on tests of the slope parameter and we derive limiting distributions under a local alternative. This allows us to compute local asymptotic power for a wide range of kernels and bandwidths. Because size is well controlled, we base the choice of kernel and bandwidth on how they affect power.

The simple linear trend model is given by

$$y_t = \beta_1 + \beta_2 t + u_t, \ t = 1, ..., T. \tag{7}$$

The null hypothesis under consideration is $H_0 : \beta_2 \leq \beta_0$. The alternative is given by

$$H_A : \beta_2 = \beta_0 + \delta g(T),$$

where $g(T) = T^{-3/2}$ if $|\alpha| < 1$ and $g(T) = T^{-1/2}$ if $\alpha = 1 - \frac{\overline{\alpha}}{T}$. The t statistic for this test is given by

$$t = \left(\frac{T^{3/2}\left(\widehat{\beta}_2 - \beta_0\right)}{\sqrt{\widehat{\sigma}^2 \left(T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2\right)^{-1}}} \right) \exp\left(-cUR\right). \tag{8}$$

The limiting null distribution of t follows from Theorem 1. Note that $\widehat{\sigma}^2$, J and BG are exactly invariant (invariant for all T) to the true value of β_2 and are hence exactly invariant to the value of δ. Therefore, only $\widehat{\beta}_2 - \beta_0$ depends on the local alternative. The following theorem gives the limiting distribution of t under the local alternative.

Theorem 2 *Let $M = bT$, $b \in (0,1]$. Suppose Assumptions 1 and 2 hold. Let t be given by (8) and let $F(r) = (1,r)'$ and $R_1^* = (0,1)'$. Then under the local alternative, H_A, as $T \to \infty$*

$$t \Rightarrow \left(\frac{\nu + R_1^*\left(\int_0^1 F(s)\,F(s)'\,ds\right)^{-1} N^F}{\sqrt{\Phi^F(b,k)\,R_1^*\left(\int_0^1 F(s)\,F(s)'\,ds\right)^{-1} R_1^{*\prime}}} \right) \exp(-cUR_\infty),$$

where $\nu = \delta/\sigma$ if $|\alpha| < 1$ and $\nu = \delta/d(1)$ if $\alpha = 1 - \overline{\alpha}/T$.

Using the results of this theorem, it is easy to simulate asymptotic power of the t statistic for different choices of kernels and bandwidths. The first step is to simulate asymptotic critical values under the null hypothesis. This was done using 50,000 replications. For each replication, we approximated the Wiener processes implicit in the limiting distributions using normalized partial sums of 1,000 iid N(0,1) random deviates. We focused on five well known kernels: Bartlett, Parzen, Bohman, Daniell and Quadratic Spectral (QS). Formulas for the kernels are given in an appendix. We considered the grid of bandwidths given by $b = 0.02,\ 0.04, ..., 1$. Given a percentage point, for a given bandwidth and kernel we computed values of c such that the asymptotic critical values are the same for $|\alpha| < 1$ and $\alpha = 1$. These values of c are different for the J and BG scaling factors. Given the values of c and the critical values, the second step is to compute rejection probabilities for a grid of values of ν using simulation methods thus producing asymptotic power curves.

To guide the choice of kernel and bandwidth, we computed power for the five kernels and the bandwidth grid for a grid of values of ν. These calculations were done for $\overline{\alpha} = 0, 1, 2, ..., 49, 50$. For a given value of $\overline{\alpha}$ and for each value of ν, maximal power across the kernels, bandwidths and choice of scaling factor was found thus providing power envelopes of the HAC robust tests. We label these power envelopes 'robust envelopes'. Similar calculations were performed for the case of stationary errors (note that the choice of scaling factor is irrelevant in this case). Given the robust envelopes, we then searched for specific kernel and bandwidth choices that give tests with power that is close to the corresponding robust envelope. In a preliminary analysis we found that the Daniell kernel generally delivers tests with power closest to the robust envelopes and we focus on the Daniell kernel for the remainder of the paper.

Focusing on the Daniell kernel, for a given value of $\overline{\alpha}$ we computed bandwidths that maximize power across the values of ν. We found that most of the time no single bandwidth choice maximizes power for all values of ν. A common pattern we found is that a relatively small bandwidth will maximize power when ν is small whereas a relatively big bandwidth will maximize power when ν is large. In other words, power curves using small bandwidths often cross power curves using large bandwidths. Because no single bandwidth gives a test with uniformly maximal power in this situation, we used the slightly weaker criterion of integrated power to choose the bandwidth. If we denote the power function of our one-sided test by $p(\nu;\overline{\alpha},b)$, integrated power is given by

$$\int_0^\infty p(\nu;\overline{\alpha},b)d\nu$$

which is easily approximated by numerical integration methods. Integrated power is simply a measure of average power over the alternative space and is similar to other average power criteria used in the econometrics literature, e.g. Andrews and Ploberger (1994). Unique bandwidths can be found that are optimal in the sense of maximizing integrated power. Note that if there exists a bandwidth that gives a uniformly most powerful test, the integrated power criterion delivers the same bandwidth.

Given a scaling factor, J or BG, for each value of $\overline{\alpha}$ in our grid we computed integrated power using simulation methods described above and we determined the bandwidth, b_{opt}, that maximizes integrated power. We used relatively fine grids for ν and we truncated the integral at large enough values of ν so that, for all bandwidths, power was at least 0.999 for values of ν above the truncation. In Figure 1 we plot b_{opt} for both the J and BG scaling factors. In both cases, b_{opt} is relatively large when $\overline{\alpha}$ is close to zero and b_{opt} declines as $\overline{\alpha}$ increases and the errors become more stationary. The decline in b_{opt} occurs in somewhat discrete drops and this is likely due to the relative coarseness of the grid for $\overline{\alpha}$. Smoother functions for b_{opt} could be obtained with a finer grid at a substantially higher computational cost. Once $\overline{\alpha}$ is big enough, b_{opt} drops to 0.02 and stays there. Calculations done for the case of stationary errors confirm that $b_{opt} = 0.02$ for stationary errors. In the remainder of the paper, Dan-J and Dan-BG are always implemented using optimal bandwidths and are labeled Dan-J and Dan-BG.

To show that Dan-J and Dan-BG have asymptotically correct size (due to the use of the scaling factors), Table 1 provides asymptotic null rejection probabilities of the Dan-J and Dan-

BG statistics. The $t - PSW$ statistic is also included for comparison. Results are given for $\overline{\alpha} = 0, 2, 4, ..., 18, 20$ at the 5% significance level. In all cases rejection probabilities are no larger than 0.05 indicating that the scaling factors deliver tests that have asymptotically correct size.

Our power calculations reduce the number of potential tests in our class of tests from an infinite number to two. We are now in a position to address two interesting and practical questions. First, how does the power of Dan-J and Dan-BG compare with each other and compare with power of $t - PSW$? Second, how close is the power of Dan-J, Dan-BG, and $t - PSW$ to the robust envelopes?

In Figure 2 we plot integrated power of the three tests relative to the integrated power of Dan-J across the grid for $\overline{\alpha}$. Obviously, the integrated power of Dan-J relative to itself is always one. Two patterns stand out. First, notice that Dan-J dominates $t - PSW$ for all values of $\overline{\alpha}$ although power is similar for $\overline{\alpha} \approx 20$. Second, the integrated power of Dan-J and Dan-BG cross once at $\overline{\alpha} = 10$ and it is obvious that the BG scaling factor gives a more powerful test when $\overline{\alpha}$ is close to zero and the J scaling factor gives a more powerful test for more stationary errors.

Figures 3-5 plot the asymptotic power (not the integrated power) of the tests along with the robust envelopes. Stationary errors are depicted in Figure 3 and nearly integrated errors are depicted in Figures 4 and 5 with $\overline{\alpha} = 0, 10$. As an additional benchmark, Figures 4-5 also include power of a test based on infeasible generalized least squares (GLS). This power curve was easily computed analytically using theoretical results from Canjels and Watson (1997). In general, tests based on infeasible GLS dominate OLS based tests for $\overline{\alpha}$ close to zero whereas GLS and OLS are asymptotically equivalent for stationary errors as is known from the classic results of Grenander and Rosenblatt (1957). Figure 3 shows in the case of stationary errors that the Daniell kernel test using b_{opt} delivers a test with power equal to the robust envelope whereas power of $t - PSW$ lies below the robust envelope. Figure 4 shows that for $\overline{\alpha} = 0$, the Dan-BG test has power equal to the envelope. Dan-J has less power although both Dan-BG and Dan-J dominate $t - PSW$. All of the OLS HAC tests are dominated by infeasible GLS as expected. In Figure 5 we see that when $\overline{\alpha} = 10$ none of the tests attain the envelope. Dan-J slightly dominates Dan-BG and Dan-J and $t - PSW$ cross each other. These results are not surprising given the results in Figure 2.

The results given in Figures 2-5 should be interpreted with caution because implementation of infeasible GLS, Dan-J and Dan-BG requires a value of $\overline{\alpha}$ which is an unknown parameter that cannot be consistently estimated. Canjels and Watson (1997) propose a feasible version of GLS based on simple proxy (inconsistent estimate) for $\overline{\alpha}$. They deal with the uncertainty generated by using a proxy for $\overline{\alpha}$ using Bonferoni bounds. In subsequent sections we propose a straightforward data dependent bandwidth rule that makes the Dan-J and Dan-BG tests feasible in practice, and we compare the finite sample performance of the feasible tests with the $t - PSW$ statistic.

7 A Feasible Data Dependent Bandwidth Rule

There are two challenges to implementing feasible versions of the Dan-J and Dan-BG tests. The first challenge is finding a simple function that can approximate the relationship between b_{opt} and $\overline{\alpha}$ as depicted in Figure 1. The second challenge is dealing with the fact that $\overline{\alpha}$ is an unknown parameter that cannot be consistently estimated.

Let $b_{opt}(\overline{\alpha})$ denote the function that gives optimal bandwidths in terms of $\overline{\alpha}$. Given a value of $\overline{\alpha}$, it is easy, but computationally intensive, to compute $b_{opt}(\overline{\alpha})$. A practical alternative is to approximate $b_{opt}(\overline{\alpha})$ using the step-like functions depicted in Figure 1. Let $\mathbf{1}_{(x\leq a)}$ denote the indicator function that takes on the value 1 if $x \leq a$ and takes on the value 0 otherwise. The step function approximations for $b_{opt}(\overline{\alpha})$, using the same simulations as for Figure 1, are as follows. For Dan-J

$$\begin{aligned} b_{opt}(\overline{\alpha}) &= 0.02 + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 21)} + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 20)} + 0.04 \cdot \mathbf{1}_{(\overline{\alpha}\leq 19)} + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 18)} + 0.12 \cdot \mathbf{1}_{(\overline{\alpha}\leq 17)} \\ &\quad + 0.1 \cdot \mathbf{1}_{(\overline{\alpha}\leq 14)} + 0.1 \cdot \mathbf{1}_{(\overline{\alpha}\leq 12)} + 0.06 \cdot \mathbf{1}_{(\overline{\alpha}\leq 11)} + 0.12 \cdot \mathbf{1}_{(\overline{\alpha}\leq 10)} + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 7)} + 0.2 \cdot \mathbf{1}_{(\overline{\alpha}\leq 4)}, \end{aligned}$$

and for Dan-BG

$$b_{opt}(\overline{\alpha}) = 0.02 + 0.1 \cdot \mathbf{1}_{(\overline{\alpha}\leq 29)} + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 23)} + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 19)} + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 17)} + 0.02 \cdot \mathbf{1}_{(\overline{\alpha}\leq 2)}.$$

Using the fact that $\overline{\alpha} = T(1-\alpha)$ we can use these step functions to write the optimal bandwidth in terms of the sample size and an estimate of α. The simplest estimator of α is given by

$$\widehat{\alpha} = \frac{\sum_{t=2}^{T} \widehat{u}_t \widehat{u}_{t-1}}{\sum_{t=2}^{T} \widehat{u}_{t-1}^2}$$

where $\widehat{u}_t$ are the OLS residuals from (7). Plugging into the formula for $\overline{\alpha}$ gives $\widehat{\overline{\alpha}} = T(1-\widehat{\alpha})$ which can be used to define a data dependent bandwidth rule given by

$$\widehat{b}_{opt} = b_{opt}(\widehat{\overline{\alpha}}). \tag{9}$$

The bandwidth, M, used in the formula for (3) is given by

$$\widehat{M} = \max(\widehat{b}_{opt}T, 2), \tag{10}$$

where the lower bound of 2 is placed on $\widehat{M}$ to ensure that $\widehat{M}$ is not too small when T and $\widehat{b}_{opt}$ are both small.

We could follow Canjels and Watson (1997) and attempt to derive Bonferoni bounds when using $\widehat{b}_{opt}$. However, this calculation would be much more difficult than the calculations carried out by Canjels and Watson (1997). This is true because of the very complicated manner in which $\overline{\alpha}$ enters the asymptotic distribution theory. By contrast, in Canjels and Watson (1997) $\overline{\alpha}$ only entered the asymptotic distributions through the variance of a normal random variable thus greatly simplifying the computation of Bonferoni bounds.

While a theoretical analysis of the impact of using $\widehat{b}_{opt}$ on the asymptotic distributions of Dan-J and Dan-BG is well beyond the scope of this paper, we can assess through finite sample simulations whether the following naive but simple asymptotic approximation works. Recall that the asymptotic critical value and the scaling factor, c, depend on the value of b. A simple alternative to Bonferoni bounds is to use an asymptotic critical value and a value for c treating $\widehat{b}_{opt}$ as constant. This approach is similar in spirit to the common practice of treating asymptotic variance estimators as known when using standard first order asymptotics. Although $\widehat{b}_{opt}$ is clearly not constant nor even a consistent estimate of b_{opt}, finite sample simulation results in the next section indicate that treating it as consistent provides a remarkably accurate asymptotic approximation.

A final practical matter that requires discussion is a convenient way of obtaining asymptotic critical values and the corresponding values of c for a given value of $\widehat{b}_{opt}$. To that end, we took simulated asymptotic critical values and values of c for the grid of $b = 0.02, 0.04, ..., 0.98, 1.0$ and estimated the following polynomial functions using OLS:

$$cv(b) = \theta_0 + \theta_1 b + \theta_2 b^2 + \theta_3 b^3 + \theta_4 b^4 + \theta_5 b^5,$$
$$c(b) = \lambda_0 + \lambda_1 b + \lambda_2 b^2 + \lambda_3 b^3 + \lambda_4 b^4 + \lambda_5 b^5 + \lambda_6 b^6 + \lambda_7 b^7.$$

The estimated coefficients are given in Tables 2 and 3 along with the R^2 from the regressions. In all cases, the fits are excellent. In practice, given a value of $\widehat{b}_{opt}$ and given a significance level, the value of c used for the scaling factor is given by $c(\widehat{b}_{opt})$ and the rejection rule is carried out using the asymptotic critical value given by $cv(\widehat{b}_{opt})$. For convenience sake, we also report $cv(b)$ and $c(b)$ functions for the t-statistic on the intercept parameter in (7).

8 Finite Sample Evidence

In this section, we discuss some finite sample simulations designed to assess the accuracy of the asymptotic approximations and to compare the finite sample performance of the tests. The Dan-J and Dan-BG tests were implemented using the feasible data dependent bandwidth described in the previous section.

For the finite sample simulations, we continue to use model (7). We test the hypothesis that $\beta_2 \leq 0$ against $\beta_2 > 0$ at the 5% significance level. The errors are generated according to $u_t = \alpha u_{t-1} + e_t + \phi e_{t-1}$, where e_t is $i.i.d.$ $N(0,1)$. The first set of simulations assesses the accuracy of asymptotic approximations under the null. Simulations are reported for $\alpha = 0.0$, 0.7, 0.8, 0.9, 0.95, 1.0, and $\phi = -0.8$, -0.4, -0.0, 0.4, 0.8 and for sample sizes 50, 100, and 200. In all cases, $5,000$ replications were used. Table 4 provides empirical null rejection probabilities of the t-tests. It is clear that unless a large negative MA-term and a unit root are simultaneously present, all of these tests have empirical rejection probabilities either close to 0.05 or below. Therefore, the J and BG scaling factors work well in practice. This contrasts with standard HAC robust tests were it is well known that strong serial correlation causes over-rejections that can be severe. See Vogelsang (1998, Table I). The reason the tests over-reject when there is a unit root and a large negative MA component is because the J and BG statistics are oversized as unit root tests. In other words, J and BG tend to be too small in finite samples and they do not scale down the t-statistics enough to control the over-rejection problem.

The overall performance of Dan-J and Dan-BG in terms of size is similar to $t - PWS$ and is quite impressive. These results suggest that treating $\widehat{b}_{opt}$ as constant is a reasonable approach.

We also report some finite sample power results to show that power in practice is qualitatively similar to that implied by the local asymptotic analysis. For comparison purposes, we also include power of a conservative feasible GLS test suggested by Canjels and Watson (1997). We implement this test in exactly the same manner as in the finite sample power simulations reported by Vogelsang (1998). Figures 6-8 plot power for $\alpha = 0.7$, 0.9, 1.0, $\phi = 0$, for $T = 100$. The results show that the asymptotic patterns are also reflected in the finite sample results. Dan-J dominates $t - PSW$ in all cases whereas Dan-J dominates Dan-BG except when $\alpha = 1$. Perhaps as expected, when $\alpha = 1$, the feasible GLS test has much higher power than the OLS based tests. However, for $\alpha < 1$ feasible GLS can have much lower power than the OLS based tests.

Because of the well known downward bias of $\widehat{\alpha}$ in models with deterministic trends, we experimented using the median unbiased estimator of α proposed by Andrews (1993) in place of $\widehat{\alpha}$ when computing $\widehat{b}_{opt}$. While size results were similar, power was often lower using the median unbiased estimator of α. It is an interesting topic for future research to more carefully compare the relative merits of various methods of estimating α when constructing $\widehat{b}_{opt}$.

Based on the results of this section we recommend that the Dan-J test be used in practice. The power of Dan-J dominates $t - PSW$ and it has much higher power than feasible GLS when α is not close to one. We do not recommend that Dan-BG be used in practice because in the one situation where Dan-BG has higher power than Dan-J, feasible GLS has much higher power than both.

9 Evidence on the Prebisch-Singer Hypothesis

In this section we provide empirical evidence on the Prebisch-Singer hypothesis. The time series we analyze is the logarithm of the net barter terms of trade series constructed by Grilli and Yang (1988) and extended by Lutz (1999). See Grilli and Yang (1988) and Lutz (1999) for details on the construction of this time series. The data is annual from 1900-1995. The net barter terms of trade is the ratio of a non-fuel primary commodities price index to a manufacturing price index. The Prebisch-Singer hypothesis asserts that the net barter terms of trade should be falling over time. We plot the data in Figure 9 and it is clear from the plot that the logarithm of net barter terms of trade has been decreasing over time. Is this decrease systematic? If we take regression (7) as a reasonable model of the statistical time series behavior of the logarithm of the net barter terms of trade, then the Prebisch-Singer hypothesis asserts that the trend slope coefficient is negative. If we take as the null hypothesis that the Prebisch-Singer hypothesis does not hold against the alternative that the Prebisch-Singer hypothesis holds, then we can parameterize the hypothesis as $H_0 : \beta_2 \leq 0$, $H_1 : \beta_2 > 0$.

Note that the Prebisch-Singer hypothesis is an empirical notion about the long run behavior of a time series; namely that the time series is steadily decreasing over time. It is important to keep in mind that this notion has nothing to do with the correlation in the data. More specifically, the Prebisch-Singer hypothesis has nothing to do with whether the error term is stationary or has a unit root. In our opinion, the empirical literature on the Prebisch-Singer hypothesis has become distracted by the unit root issue. This is not surprising given the technical difficulties the presence of a unit root brings with it. The advantage of the test proposed in this paper is that it allows a direct and very simple test of the Prebisch-Singer hypothesis that does not depend on whether or not a unit root is in the errors.

Using the logarithm of the net barter terms of trade series, we estimated regression (7) by OLS and obtained $\widehat{\beta}_2 = -0.0645$. We computed the Dan-$J$, the recommended test, using the data dependent bandwidth. Recall that the value of c used for the scaling factors depends on the significance level of the tests and we provide results for significance levels 5% and 2.5%. The results are given in Table 5. Also reported in Table 5 are $\widehat{\alpha}$ and $\widehat{b}_{opt}$. The null hypothesis that the Prebisch-Singer hypothesis does not hold can be rejected at the 5% level but not at the 2.5% level. This rejection is robust because the tests do not suffer from over-rejection problems even if the errors have a unit root. Our empirical result suggests that there is relatively strong evidence that the Prebisch-Singer hypothesis holds implying that Prebisch and Singer were right.

As an additional robustness check, we applied the partial sum trend function structural change tests proposed by Vogelsang (1999). We computed variants of the Vogelsang (1999) tests designed to jointly detect a shift in intercept and/or slope in the deterministic trend function. The break date was treated as unknown. The tests also use the J scaling factor to control the over-rejection problem caused by strong serial correlation. We computed the mean, mean-exponential and supremum statistics using 1% trimming (see Vogelsang (1999) for details). The results were: mean=0.084, mean-exponential=0.0103 and supremum=0.0948. The 5% asymptotic critical values for these tests when using the J scaling factor are 2.0917, 1.3325 and 5.1651 respectively. Therefore, the null hypothesis that the trend function is stable over time cannot be rejected.

10 Conclusion

In this paper we have proposed tests for hypotheses regarding the parameters of the deterministic trend function of a univariate time series. The tests do not require knowledge of the form of serial correlation in the data and they are robust to strong serial correlation. The data can even contain a unit root and the tests still have the correct size asymptotically. The tests we analyze are standard OLS HAC robust tests based on nonparametric variance estimators. We extend the fixed-b asymptotic framework for HAC robust tests recently proposed by Kiefer and Vogelsang (2002). This allows us to analyze the power properties of the tests with regards to bandwidth and kernel choices. Our analysis shows that among popular kernels, the Daniell kernel delivers tests with optimal power within a specific class of tests that have the correct asymptotic size whether the errors are stationary or have a unit root. We achieve this size robustness using the J scaling factor proposed by Vogelsang (1998) and a new scaling factor, BG, based on the unit root test of Breitung (2002). Based on our asymptotic and finite sample analysis we recommend that the J correction be used over the BG correction in practice. Our results also suggest that the Dan-J test dominates the $t - PSW$ test of Vogelsang (1998) in terms of power and Dan-J is recommended in practice.

We address the traditionally difficult issue of HAC bandwidth choice using fixed-b asymptotics in conjunction with local to unity asymptotics. For a given value of the local to unity parameter, we numerically determining the bandwidth that maximizes integrated power. To our knowledge, this is the first bandwidth analysis that focuses on power of tests rather than the mean square error of the HAC estimate. Our analysis provides a data dependent bandwidth that maximizes integrated power. Finite sample simulations suggest that a feasible version of the optimal bandwidth rule works well in practice.

We applied the Dan-J test to the logarithm of a net barter terms of trade series and the test suggest that this series has a statistically significant negative slope. This finding is consistent with the well known Prebisch-Singer hypothesis. Because our tests are robust to strong serial correlation or a unit root in the data, our results in support of the Prebisch-Singer hypothesis are robust.

11 Acknowledgements

We are grateful to Alastair Hall, an associate editor and two referees for thoughtful and constructive comments that lead to improvements of the paper.We thank Jan Jacobs for useful discussions, as well as seminar participants at the University of Groningen, University of Tilburg, Iowa State University (Statistics), Cornell University and the American Statistical Association Meetings (August 2000) for beneficial feedback. Vogelsang thanks the Center for Analytic Economics at Cornell and gratefully acknowledges support from the National Science Foundation under grant SES-0095211. We are grateful to Matthias Lutz who kindly provided the data used in the paper.

Appendix

In this appendix we give the proof of Theorem 1. Theorem 2 follows easily from Theorem 1 using simple algebra and details are omitted.

Proof of Theorem 1.

Proof of part a):

Following Kiefer and Vogelsang (2002), we define

$$\Delta^2 \kappa_{ij} = \left\{ k\left(\frac{i-j}{[bT]}\right) - k\left(\frac{i-j-1}{[bT]}\right) \right\} - \left\{ k\left(\frac{i-j+1}{[bT]}\right) - k\left(\frac{i-j}{[bT]}\right) \right\},$$

and use this expression to rewrite $\widehat{\sigma}^2$ as

$$\widehat{\sigma}^2 = -T^{-1} \sum_{i=1}^{T-1} T^{-1} \sum_{j=1}^{T-1} T^2 \Delta^2 \kappa_{ij} \left(T^{-1/2}\widehat{S}_i\right)\left(T^{-1/2}\widehat{S}_j\right). \tag{11}$$

For (11) to be valid it must be the case that the residuals sum to zero. So, for the asymptotic results to hold, a constant must be included in the model. The following lemma provides the distribution of $T^{-1/2}\widehat{S}_t$.

Lemma 2 $T^{-1/2}\widehat{S}_{[rT]} \Rightarrow \sigma Q^F(r)$.

Proof of Lemma 2: Simple matrix manipulations yield:

$$T^{-1/2}\widehat{S}_{[rT]} = T^{-1/2} \sum_{t=1}^{[rT]} u_t - \left(T^{-1} \sum_{t=1}^{[rT]} f(t)' \tau_T \right) T^{1/2} \tau_T^{-1} \left(\widehat{\beta} - \beta\right). \tag{12}$$

where

$$T^{1/2}\tau_T^{-1}\left(\widehat{\beta} - \beta\right) = \left(T^{-1}\tau_T \mathbf{f}(T)' \mathbf{f}(T) \tau_T\right)^{-1} \left(T^{-1/2}\tau_T \mathbf{f}(T)' u\right). \tag{13}$$

Clearly the terms consisting only of trend functions will have limiting distributions which do not depend on whether or not u_t is stationary. It is well know that these terms have the following limits:

$$T^{-1}\tau_T \mathbf{f}(T)' \mathbf{f}(T) \tau_T \Rightarrow \int_0^1 F(s) F(s)' ds, \text{ and} \tag{14}$$

$$T^{-1} \sum_{t=1}^{[rT]} f(t)' \tau_T \Rightarrow \int_0^r F(s)' ds. \tag{15}$$

The last term in (13) and the first term in (12) depend on u_t and therefore their limiting distributions will depend on whether or not u_t is stationary. Again using standard results, those asymptotic distributions are:

$$T^{-1/2}\tau_T \mathbf{f}(T)' u \Rightarrow \sigma \int_0^1 F(s)\, dw(s) \quad \text{if} \quad |\alpha| < 1,$$
$$T^{-3/2}\tau_T \mathbf{f}(T)' u \Rightarrow d(1) \int_0^1 F(s)\, V_{\overline{\alpha}}(s)\, ds \quad \text{if} \quad \alpha = 1 - \frac{\overline{\alpha}}{T},$$
$$T^{-1/2}\sum_{t=1}^{[rT]} u_t \Rightarrow \sigma w(r) \quad \text{if} \quad |\alpha| < 1,$$
$$T^{-3/2}\sum_{t=1}^{[rT]} u_t \Rightarrow d(1) \int_0^r V_{\overline{\alpha}}(s)\, ds \quad \text{if} \quad \alpha = 1 - \frac{\overline{\alpha}}{T}.$$

Using these limits the asymptotic distribution of $\widehat{S}_{[rT]}$ is as follows.

$$\begin{aligned} T^{-1/2}\widehat{S}_{[rT]} &\Rightarrow \sigma \left(w(r) - \int_0^r F(s)'\, ds \left(\int_0^1 F(s) F(s)'\, ds \right)^{-1} \int_0^1 F(s)\, dw(s) \right) \\ &= \sigma Q^F(r) \quad \text{if} \quad |\alpha| < 1, \\ T^{-3/2}\widehat{S}_{[rT]} &\Rightarrow d(1) \left(\int_0^r V_{\overline{\alpha}}(s)\, ds - \int_0^r F(s)'\, ds \left(\int_0^1 F(s) F(s)'\, ds \right)^{-1} \int_0^1 F(s)\, V_{\overline{\alpha}}(s)\, ds \right) \\ &= d(1) Q^F(r) \quad \text{if} \quad \alpha = 1 - \frac{\overline{\alpha}}{T}. \end{aligned}$$

The rest of the proof is split into three cases, corresponding to Type 1, Type 2 and the Bartlett kernels.

Case 1: $k(x)$ is a Type 1 kernel. By definition of the second derivative, $T^2\Delta^2\kappa_{ij} \to k''$, and using Lemma (2) it follows easily for the case when $|\alpha| < 1$ that

$$\begin{aligned} \widehat{\sigma}^2 &= T^{-1}\sum_{i=1}^{T-1} T^{-1}\sum_{j=1}^{T-1} -T^2\Delta^2\kappa_{ij} T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_j \\ &\Rightarrow \sigma^2 \int_0^1 \int_0^1 -k''(r-s)\, Q^F(r)\, Q^F(s)\, dr ds. \end{aligned}$$

When $\alpha = 1 - \frac{\overline{\alpha}}{T}$ we have

$$\begin{aligned} T^{-2}\widehat{\sigma}^2 &= T^{-1}\sum_{l=1}^{T-1} T^{-1}\sum_{i=1}^{T-1} -T^2\Delta^2\kappa_{il} T^{-3/2}\widehat{S}_i T^{-3/2}\widehat{S}_l \\ &\Rightarrow d(1)^2 \int_0^1 \int_0^1 -k''(r-s)\, Q^F(r)\, Q^F(s)\, dr ds. \end{aligned}$$

Case 2: $k(x)$ is a Type 2 kernel. Following Kiefer and Vogelsang (2002), we use simple algebra and the definition of $\Delta^2\kappa_{ij}$ to establish that when $|i-j| > [bT]$, $\Delta^2\kappa_{ij} = 0$, and when $|i-j| = [bT]$, $\Delta^2\kappa_{ij} = -k\left(\frac{[bT]-1}{[bT]}\right)$. When $|i-j| < [bT]$, $T^2\Delta^2\kappa_{ij} \to k''$. First consider the case when $|\alpha| < 1$. We split up the expression of $\widehat{\sigma}^2$ as follows:

$$
\begin{aligned}
\widehat{\sigma}^2 &= T^{-1}\sum_{i=1}^{T-1} T^{-1}\sum_{j=1}^{T-1} -T^2\Delta^2\kappa_{ij}T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_j \\
&= T^{-1}\sum_{i=1}^{T-1} T^{-1}\sum_{j=1}^{T-1} -1_{\{|i-j|<[bT]\}}T^2\Delta^2\kappa_{ij}T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_j \\
&\quad + 2T^{-2}\sum_{i=1}^{T-[bT]-1} T^2 k\left(\frac{[bT]-1}{[bT]}\right)T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_{i+[bT]} \\
&= T^{-1}\sum_{i=1}^{T-1} T^{-1}\sum_{j=1}^{T-1} -1_{\{|i-j|<[bT]\}}T^2\Delta^2\kappa_{ij}T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_j \\
&\quad + 2k\left(1-\frac{1}{[bT]}\right)\sum_{i=1}^{T-[bT]-1} T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_{i+[bT]} \\
&\Rightarrow \sigma^2\left(\int\int_{|r-s|<b} -k^{*\prime\prime}(r-s)\,Q^F(r)\,Q^F(s)\,drds + 2k_-^{*'}(b)\int_0^{1-b} Q^F(r+b)\,Q^F(r)\,dr\right),
\end{aligned}
$$

where the asymptotic distribution follows directly from Lemma (2) and Kiefer and Vogelsang (2002). The result when $\alpha = 1 - \frac{\overline{\alpha}}{T}$ follows analogously for $T^{-2}\widehat{\sigma}^2$ where $\widehat{S}_i$ is normalized by $T^{-3/2}$ instead of $T^{-1/2}$.

Case 3: $k(x)$ is the Bartlett Kernel. Here again using simple algebra following Kiefer and Vogelsang (2002), it can be verified that when $|i-j| = 0$, $\Delta^2\kappa_{ij} = \frac{2}{[bT]}$, and when $|i-j| = [bT]$, $\Delta^2\kappa_{ij} = -\frac{1}{[bT]}$. Using these expressions and Lemma (2) in (11), we obtain the following limiting distribution when $|\alpha| < 1$:

$$
\begin{aligned}
\widehat{\sigma}^2 &= T^{-1}\sum_{j=1}^{T-1} T^{-1}\sum_{i=1}^{T-1} T^2\Delta^2\kappa_{ij}T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_j \\
&= \frac{2}{[bT]}\sum_{i=1}^{T-1}\left(T^{-1/2}\widehat{S}_i\right)^2 - \frac{2}{bT}\sum_{i=1}^{T-[bT]-1} T^{-1/2}\widehat{S}_i T^{-1/2}\widehat{S}_{i+[bT]} \\
&\Rightarrow \sigma^2\left(\frac{2}{b}\int_0^1 Q^F(r)^2\,dr - \frac{2}{b}\int_0^{1-b} Q^F(r+b)\,Q^F(r)\,dr\right),
\end{aligned}
$$

The result when $\alpha = 1 - \frac{\overline{\alpha}}{T}$ follows analogously for $T^{-2}\widehat{\sigma}^2$ where $\widehat{S}_i$ is normalized by $T^{-3/2}$ instead of $T^{-1/2}$. Comparing the distributions from Cases 1-3 with the definition of $\Phi^F(b,k)$ completes the proof of **a)**.

Proof of part b): First note that W_T can be written as

$$W_T = \left(R\widehat{\beta} - q\right)' \left[\widehat{\sigma}^2 R \left(\mathbf{f}(T)' \mathbf{f}(T)\right)^{-1} R'\right]^{-1} \left(R\widehat{\beta} - q\right) \exp(-cUR).$$

$$= T\left(R\widehat{\beta} - q\right)' \left[\widehat{\sigma}^2 R\tau_T \left(\frac{1}{T}\tau_T \mathbf{f}(T)' \mathbf{f}(T)\tau_T\right)^{-1} \tau_T R'\right]^{-1} \left(R\widehat{\beta} - q\right) \exp(-cUR)$$

$$= \left[\left(A^{-1}R\tau_T\right)\tau_T^{-1}T^{1/2}\left(\widehat{\beta} - \beta\right)\right]' \left[\widehat{\sigma}^2 \left(A^{-1}R\tau_T\right)\left(\frac{1}{T}\tau_T \mathbf{f}(T)' \mathbf{f}(T)\tau_T\right)^{-1}\left(\tau_T R' A^{-1}\right)\right]^{-1}$$
$$\times \left[\left(A^{-1}R\tau_T\right)\tau_T^{-1}T^{1/2}\left(\widehat{\beta} - \beta\right)\right] \exp(-cUR),$$

By definition $A^{-1}R\tau_T \to R^*$. Furthermore we established the asymptotic distribution of $\widehat{\sigma}^2$ in **a)**. It therefore directly follows that when $|\alpha| < 1$

$$W_T \Rightarrow \left[\sigma R^* \left(\int_0^1 F(s)F(s)'\,ds\right)^{-1} N^F\right]' \left[\sigma^2 \Phi^F(b,k) R^* \left(\int_0^1 F(s)F(s)'\,ds\right)^{-1} (R^*)'\right]^{-1}$$

$$\times \left(\sigma R^* \left(\int_0^1 F(s)F(s)'\,ds\right)^{-1} N^F\right) \exp(-cUR_\infty)$$

$$= \left[R^* \left(\int_0^1 F(s)F(s)'\,ds\right)^{-1} N^F\right]' \left[\Phi^F(b,k) R^* \left(\int_0^1 F(s)F(s)'\,ds\right)^{-1} (R^*)'\right]^{-1}$$
$$\times \left(R^* \left(\int_0^1 F(s)F(s)'\,ds\right)^{-1} N^F\right) \exp(-cUR_\infty).$$

When $\alpha = 1 - \frac{\overline{\alpha}}{T}$ the desired result follows by normalizing $\left(\widehat{\beta} - \beta\right)$ by $T^{-1/2}$ and normalizing $\widehat{\sigma}^2$ by T^{-2}. Part c) of the theorem follows directly from part b). ■

A List of Kernels

The kernels we use:

$$\text{Bartlett} \quad k(x) = \begin{cases} 1-|x| \ \text{ for } \ |x| \le 1 \\ 0 \text{ otherwise} \end{cases}$$

$$\text{Parzen (a)} \quad k(x) = \begin{cases} 1-6x^2+6|x|^3 \ \text{ for } \ |x| \le \frac{1}{2} \\ 2(1-|x|)^3 \ \text{ for } \frac{1}{2} \le |x| \le 1 \\ 0 \text{ otherwise} \end{cases}$$

$$\text{Quadratic Spectral (QS)} \quad k(x) = \frac{25}{12\pi^2 x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right)$$

$$\text{Daniell} \quad k(x) = \frac{\sin(\pi x)}{\pi x}$$

$$\text{Bohman} \quad k(x) = \begin{cases} (1-|x|)\cos(\pi x) + \sin(\pi|x|)/\pi \ \text{ for } \ |x| \le 1 \\ 0 \text{ otherwise} \end{cases}$$

The second derivatives of the kernels we use are:

$$\text{Parzen (a)} \quad k''(x) = \begin{cases} -12+36|x| \ \text{ for } \ |x| \le \frac{1}{2} \\ 12(1-|x|) \ \text{ for } \frac{1}{2} \le |x| \le 1 \end{cases}$$

$$\text{QS} \quad k''(x) = \begin{cases} -\frac{36\pi^2}{125} \ \text{ for } x = 0 \\ \frac{125}{72\pi^3 x^5}\left[\left(12 - \frac{36\pi^2 x^2}{5}\right)\sin(6\pi x/5) + \right. \\ \left. \left(\frac{216\pi^3 x^3}{125} - \frac{72\pi x}{5}\right)\cos(6\pi x/5)\right] \ \text{ otherwise} \end{cases}$$

$$\text{Daniell} \quad k''(x) = \begin{cases} -\frac{1}{3}\pi^2 \ \text{ for } x = 0 \\ \frac{2(\sin(\pi x) - \pi x\cos(\pi x))}{\pi x^3} - \frac{\pi \sin(\pi x)}{x} \ \text{ otherwise} \end{cases}$$

$$\text{Bohman} \quad k''(x) = \pi\sin(\pi|x|) - \pi^2(1-|x|)\cos(\pi x)$$

Note that in the case of the Bartlett kernel, the asymptotic distribution is not expressed in terms of $k''(x)$ because $k''(0)$ does not exist for the Bartlett kernel. The fact that $k''(0)$ does not exist does not pose any technical problems because results for the Bartlett kernel are obtained through direct calculations.

References

Andrews, D. W. K. (1993), Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models, *Econometrica* **61**, 139–165.

Andrews, D. W. K. and Ploberger, W. (1994), Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative, *Econometrica* **62**, 1383–1414.

Ardeni, P. G. and Wright, B. (1992), The Prebisch-Singer Hypothesis: A Reappraisal Independent of Stationarity Hypotheses, *Economic Journal* **102**, 803–812.

Breitung, J. (2002), Nonparametric Tests for Unit Roots and Cointegration, *Journal of Econometrics* **108**, 343–363.

Canjels, E. and Watson, M. W. (1997), Estimating Deterministic Trends in the Presence of Serially Correlated Errors, *Review of Economics and Statistics* **May**, 184–200.

Chan, N. H. and Wei, C. (1988), Limiting Distribution of Least Squares Estimates of Unstable Autoregressive Processes, *Annals of Statistics* **16**, 367–401.

Cuddington, J. T. and Urzua, C. M. (1989), Trends and Cycles in the Net Barter Terms of Trade: A New Approach, *Economic Journal* **99**, 426–442.

Grenander, U. and Rosenblatt, M. (1957), *Statistical Analysis of Stationary Time Series*, Wiley, New York.

Grilli, E. R. and Yang, C. (1988), Primary Commodity Prices, Manufactured Goods Prices, and the Terms of Trade of Developing Countries, *World Bank Economic Review* **2**, 1–47.

Hall, A. (2004), *Generalized Method of Moments*, Oxford University Press, Oxford.

Kiefer, N. M. and Vogelsang, T. J. (2002), A New Asymptotic Theory for Heteroskedasiticy-Autocorrelation Robust Tests, Working Paper, Center for Analytic Economics, Cornell University.

Lutz, M. G. (1999), A General Test of the Prebisch-Singer Hypothesis, *Review of Development Economics* **3**, 44–57.

Park, J. Y. (1990), Testing for Unit Roots and Cointegration by Variable Addition, *in* T. Fomby and F. Rhodes (eds), *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, London: Jai Press, pp. 107–134.

Park, J. Y. and Choi, I. (1988), A New Approach to Testing for a Unit Root, Center for Analytic Economics, Cornell University, Working Paper 88-23.

Phillips, P. C. B. (1987), Time Series Regression with Unit Roots, *Econometrica* **55**, 277–302.

Phillips, P. C. B. and Solo, V. (1992), Asymptotics for Linear Processes, *The Annals of Statistics* **20**, 971–1001.

Powell, A. (1991), Commodity and Developing Country Terms of Trade: What Does the Long Run Show?, *Economic Journal* **101**, 1485–1496.

Prebisch, R. (1950), *The Economic Development of Latin America and Its Principle Problems*, United Nations Publications, New York.

Sapsford, D. (1985), The Statistical Debate on the Net Barter Terms of Trade Between Primary Commodities and Manufactures: A Comment and Some Additional Evidence, *Economic Journal* **95**, 781–788.

Singer, H. (1950), The Distributions of Gains Between Investing and Borrowing Countries, *American Economic Review, Papers and Proceedings* **40**, 473–485.

Spraos, J. (1980), The Statistical Debate on the Net Barter Terms of Trade Between Primary Commodities and Manufactures, *Economic Journal* **90**, 107–128.

Trivedi, P. K. (1995), Tests of Some Hypotheses About Time Series Behavior of Commodity Prices, *Advances in Econometrics and Quantitative Economics: A Volume in Honor of C. R. Rao*, Blackwell, Oxford, pp. 383–412.

Vogelsang, T. J. (1998), Trend Function Hypothesis Testing in the Presence of Serial Correlation Correlation Parameters, *Econometrica* **65**, 123–148.

Vogelsang, T. J. (1999), Testing for a Shift in Trend When Serial Correlation is of Unknown Form, Center for Analytic Economics Working Paper 97-11, Cornell University.

Table 1: Asymptotic Null Rejection Probabilities in the Simple Trend Model Nearly Integrated Errors, 5% Nominal Level, 50,000 Replications

$\overline{\alpha}$	Dan-J, $b = b_{opt}$	Dan-BG, $b = b_{opt}$	t-PSW-J
0	0.050	0.050	0.050
2	0.028	0.023	0.027
4	0.025	0.018	0.022
6	0.024	0.017	0.022
8	0.026	0.017	0.023
10	0.027	0.017	0.024
12	0.027	0.017	0.025
14	0.026	0.018	0.027
16	0.025	0.018	0.028
18	0.023	0.017	0.029
20	0.025	0.016	0.030

Table 2: Asymptotic Critical Value Function Coefficients of HAC Robust t-tests in the Simple Linear Trend Model Using the Daniell Kernel

$$y_t = \beta_1 + \beta_2 t + u_t.$$

$$cv(b) = \theta_0 + \theta_1 b + \theta_2 b^2 + \theta_3 b^3 + \theta_4 b^4 + \theta_5 b^5$$

Intercept	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5	R^2
90%	1.2768	3.0917	-4.6941	35.2194	-26.1069	7.9022	0.9999
95%	1.6600	3.1416	7.7632	12.4836	11.4220	-8.8793	0.9999
97.5%	1.9402	6.2830	-6.3219	82.6107	-64.6229	22.1838	0.9999
99%	2.4016	0.4538	64.5686	-66.8889	124.2655	-54.7427	0.9999
Slope	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5	R^2
90%	1.2802	2.4100	1.1323	17.1458	-4.8840	-0.6734	0.9999
95%	1.6383	3.5083	3.1079	31.3777	-16.0674	3.6881	0.9999
97.5%	1.9659	4.0603	11.6626	34.8269	-13.9506	3.2669	0.9999
99%	2.3259	6.5916	8.8314	99.0511	-73.3258	26.2719	0.9999

Table 3: Asymptotic $c(b)$ Function Coefficients of HAC Robust t-tests in the Simple Linear Trend Model Using the Daniell Kernel

$$y_t = \beta_1 + \beta_2 t + u_t.$$

$$c(b) = \lambda_0 + \lambda_1 b + \lambda_2 b^2 + \lambda_3 b^3 + \lambda_4 b^4 + \lambda_5 b^5 + \lambda_6 b^6 + \lambda_7 b^7$$

J Scaling Factor

Intercept	λ_0	λ_1	λ_2	λ_3	λ_4
90%	0.7499	-9.7621	65.796	-247.4533	528.4285
95%	0.9610	-13.1521	90.3259	-343.243	736.7416
97.5%	1.1890	-16.7142	115.2884	-438.3496	940.2971
99%	1.5870	-24.2006	172.7681	-664.5252	1429.5649

Intercept	λ_5	λ_6	λ_7	R^2
90%	-634.3561	398.1199	-101.5139	0.9970
95%	-886.1946	556.2731	-141.7148	0.9965
97.5%	-1130.4387	709.5573	-180.8483	0.9967
99%	-1717.8948	1076.238	-273.5812	0.9957

Slope	λ_0	λ_1	λ_2	λ_3	λ_4
90%	1.1531	-10.7044	69.5348	-255.9725	540.5918
95%	1.5765	-14.479	95.252	-356.2578	762.0497
97.5%	2.1582	-20.7712	142.0705	-541.8446	1164.2989
99%	2.9487	-27.6477	189.1506	-735.8488	1615.5392

Slope	λ_5	λ_6	λ_7	R^2
90%	-644.6063	402.3978	-102.0847	0.9974
95%	-918.8257	579.6667	-148.584	0.9970
97.5%	-1400.0856	878.4994	-223.8275	0.9969
99%	-1979.9895	1262.2460	-325.801	0.9969

Table 3: (continued)

BG Scaling Factor

Intercept	λ_0	λ_1	λ_2	λ_3	λ_4
90%	234.2704	-2868.4772	18811.8039	-69731.2231	147463.4333
95%	286.3570	-3743.1267	25162.8573	-94152.1563	199276.5117
97.5%	352.1194	-4943.2405	34469.1564	-131258.8606	279704.2615
99%	444.8561	-6743.2306	48990.1553	-190438.6854	409721.2544

Intercept	λ_5	λ_6	λ_7	R^2
90%	-175665.7364	109539.4716	-27779.3683	0.9981
95%	-236651.833	146849.1402	-37028.7816	0.9980
97.5%	-332970.252	206834.2458	-52197.0985	0.9980
99%	-489543.7975	304130.5416	-76587.1552	0.9974

Slope	λ_0	λ_1	λ_2	λ_3	λ_4
90%	354.9211	-3192.5493	20104.5508	-71791.6215	149125.4086
95%	458.5817	-4320.4082	28274.2022	-101854.4156	210053.0637
97.5%	577.7701	-5475.1849	35431.5583	-122293.2595	238753.4663
99%	739.8640	-7189.0036	48398.4732	-172687.4149	347792.5436

Slope	λ_5	λ_6	λ_7	R^2
90%	-176601.0298	110082.5451	-27966.1682	0.9970
95%	-245890.3349	151589.8660	-38163.0692	0.9972
97.5%	-264497.7570	154988.5891	-37301.4541	0.9978
99%	-397333.0643	239910.1386	-59396.9866	0.9970

TABLE 4: Empirical Null Rejection Probabilities in the Simple Trend Model
5% Nominal Level, 5,000 Replications

$y_t = \beta_1 + \beta_2 t + u_t$, $u_t = \alpha u_{t-1} + e_t + \phi e_{t-1}$, e_t, $i.i.d.N(0,1)$, $u_0 = 0$, $H_0 : \beta_2 \leq 0$, $H_A : \beta_2 > 0$.

		$T = 50$			$T = 100$			$T = 200$		
ϕ	α	Dan-J	Dan-BG	t-PSW	Dan-J	Dan-BG	t-PSW	Dan-J	Dan-BG	t-PSW
-0.8	0.00	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.001	0.005
	0.70	0.009	0.005	0.017	0.010	0.006	0.027	0.018	0.015	0.037
	0.80	0.026	0.014	0.033	0.041	0.026	0.051	0.049	0.038	0.050
	0.90	0.066	0.040	0.066	0.128	0.080	0.089	0.129	0.092	0.072
	0.95	0.102	0.061	0.092	0.196	0.116	0.125	0.193	0.122	0.102
	1.00	0.184	0.112	0.165	0.297	0.185	0.219	0.314	0.185	0.212
-0.4	0.00	0.014	0.008	0.019	0.014	0.011	0.030	0.036	0.033	0.040
	0.70	0.034	0.014	0.031	0.071	0.037	0.043	0.064	0.039	0.045
	0.80	0.034	0.017	0.034	0.077	0.038	0.043	0.075	0.041	0.043
	0.90	0.039	0.023	0.039	0.071	0.036	0.042	0.076	0.038	0.037
	0.95	0.050	0.034	0.049	0.070	0.037	0.043	0.058	0.031	0.031
	1.00	0.099	0.074	0.099	0.107	0.070	0.093	0.081	0.059	0.071
0.0	0.00	0.027	0.012	0.031	0.034	0.021	0.042	0.043	0.034	0.045
	0.70	0.016	0.011	0.018	0.052	0.026	0.032	0.052	0.026	0.039
	0.80	0.015	0.010	0.017	0.044	0.022	0.028	0.055	0.024	0.034
	0.90	0.015	0.013	0.016	0.031	0.017	0.022	0.044	0.017	0.027
	0.95	0.022	0.022	0.021	0.026	0.015	0.021	0.031	0.015	0.019
	1.00	0.048	0.051	0.045	0.054	0.051	0.052	0.053	0.047	0.048
0.4	0.00	0.020	0.008	0.027	0.036	0.018	0.040	0.040	0.027	0.043
	0.70	0.016	0.012	0.015	0.039	0.019	0.029	0.047	0.022	0.037
	0.80	0.016	0.012	0.011	0.031	0.017	0.024	0.046	0.018	0.032
	0.90	0.013	0.012	0.011	0.022	0.015	0.019	0.031	0.015	0.024
	0.95	0.023	0.019	0.014	0.021	0.015	0.018	0.024	0.014	0.016
	1.00	0.040	0.043	0.031	0.047	0.046	0.042	0.046	0.046	0.043
0.8	0.00	0.015	0.009	0.024	0.033	0.016	0.039	0.038	0.024	0.042
	0.70	0.020	0.012	0.014	0.033	0.018	0.028	0.046	0.020	0.036
	0.80	0.016	0.011	0.011	0.027	0.019	0.024	0.043	0.017	0.031
	0.90	0.014	0.012	0.009	0.023	0.014	0.018	0.030	0.014	0.023
	0.95	0.021	0.017	0.012	0.020	0.014	0.017	0.022	0.014	0.015
	1.00	0.037	0.042	0.028	0.044	0.045	0.041	0.045	0.046	0.042

Note: The feasible optimal bandwidth given by (10) was used for Dan-J and Dan-BG.

Table 5: Empirical Results for the Logarithm of Net Barter Terms of Trade Annual Data, 1900-1995, Dan-J Statistic

$\widehat{\beta}_2$	$\widehat{\alpha}$	$\widehat{b}_{opt}$	Dan-J (5% c)	Dan-J (2.5% c)
-0.0645	0.702	0.02	-2.445	-1.549

Note: The 5% and 2.5% asymptotic critical values are -1.710 and -2.052.

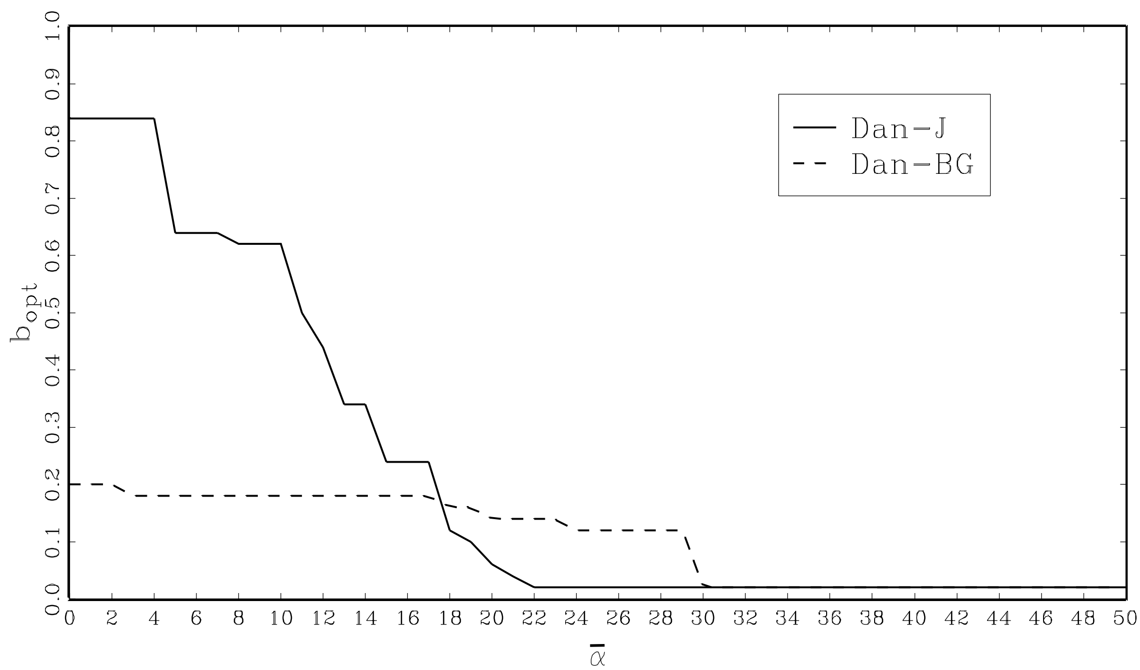

Figure 1: Asymptotic Optimal Bandwidths for Daniell Kernel

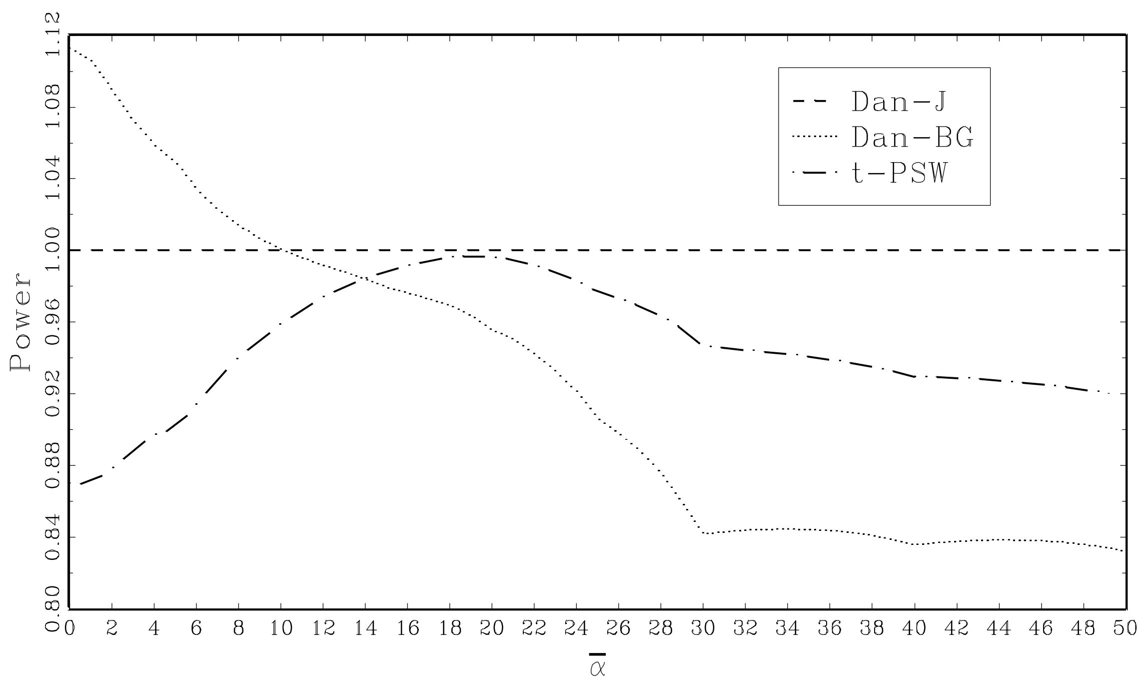

Figure 2: Integrated Asymptotic Power Relative to Dan-J

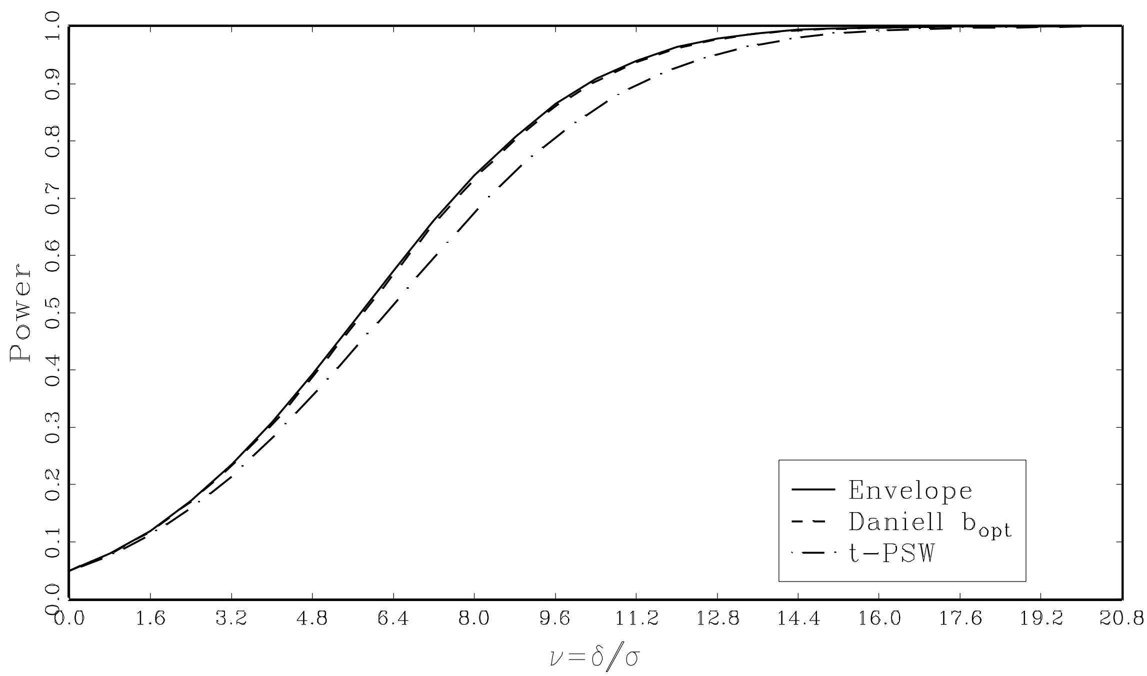

Figure 3: Asymptotic Power Stationary Errors

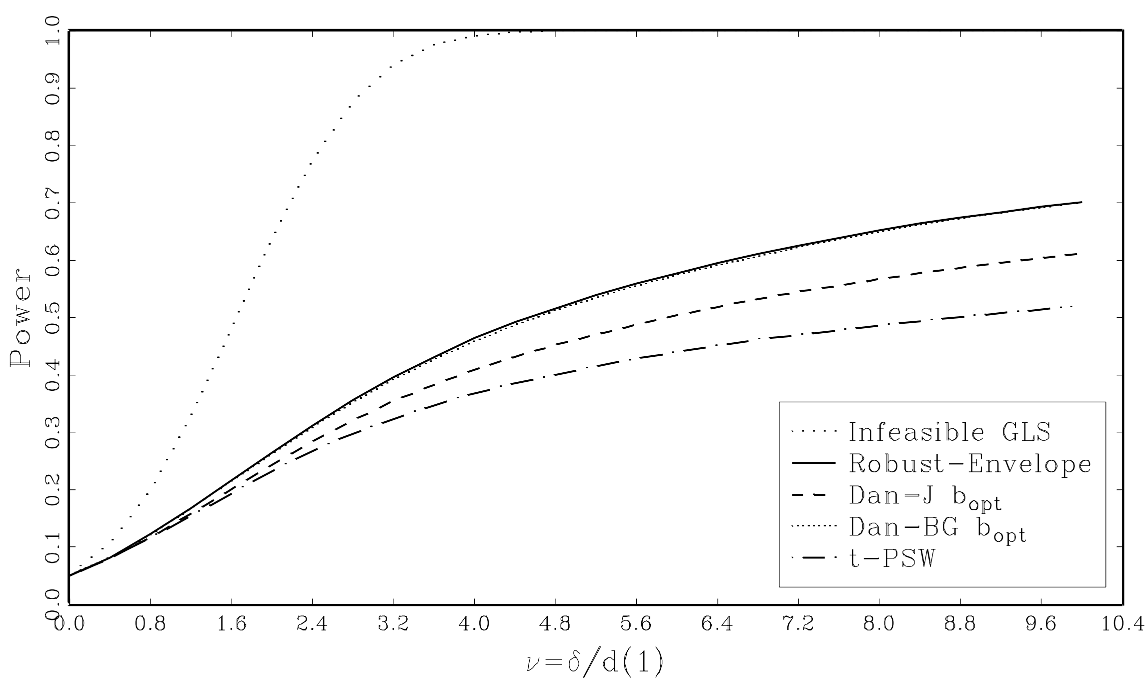

Figure 4: Asymptotic Power I(1) Errors, $\overline{\alpha}=0$

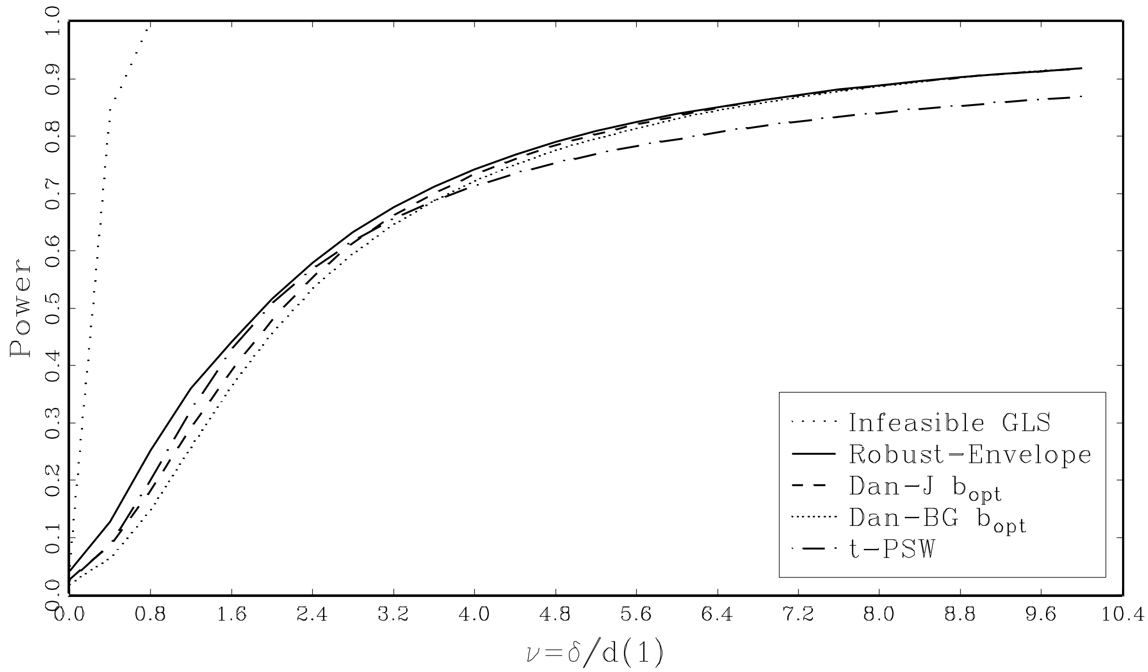

Figure 5: Asymptotic Power I(1) Errors, $\bar{\alpha}=10$

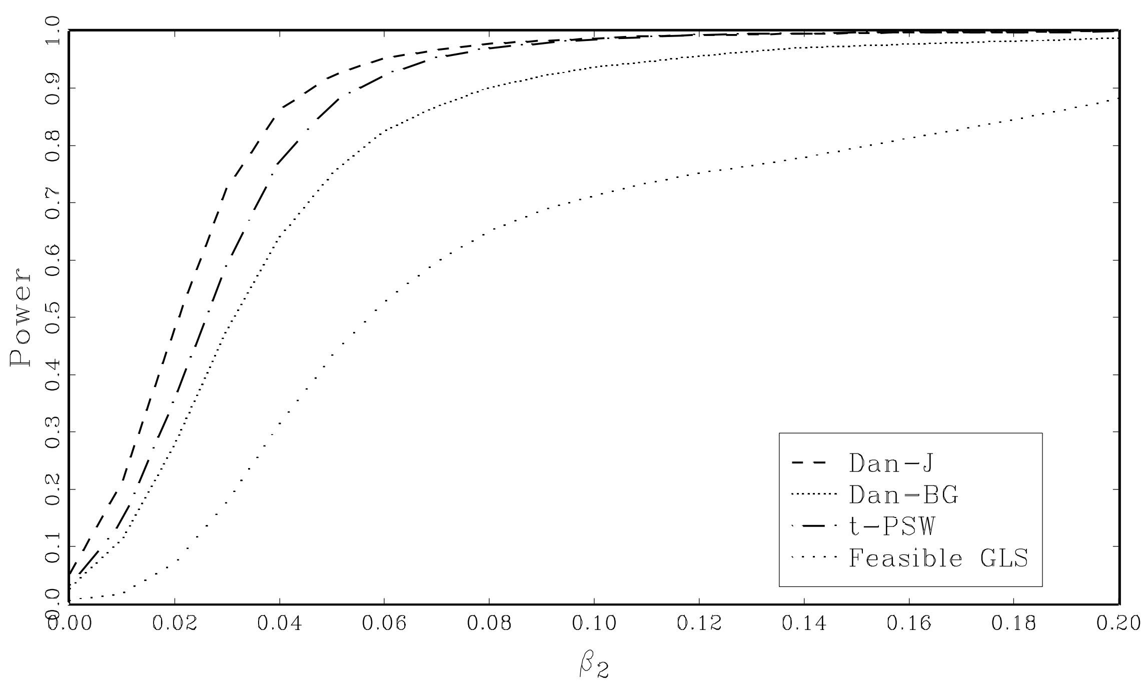

Figure 6: Finite Sample Power, AR(1) Errors, $\alpha=0.7$, T=100.

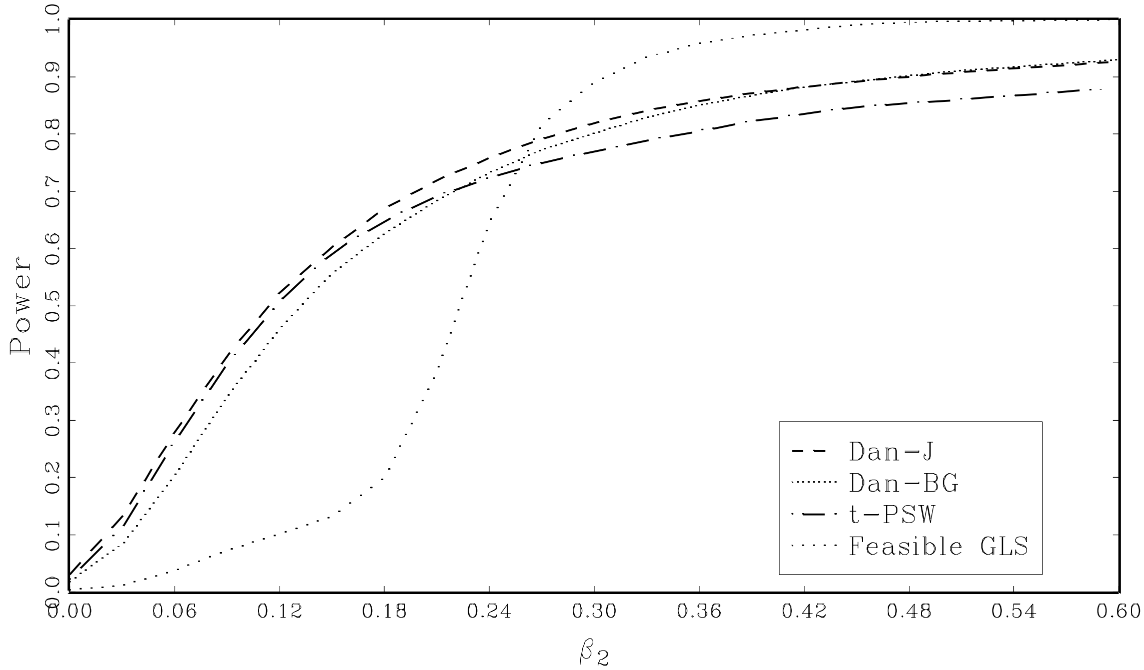

Figure 7: Finite Sample Power, AR(1) Errors, α=0.9, T=100.

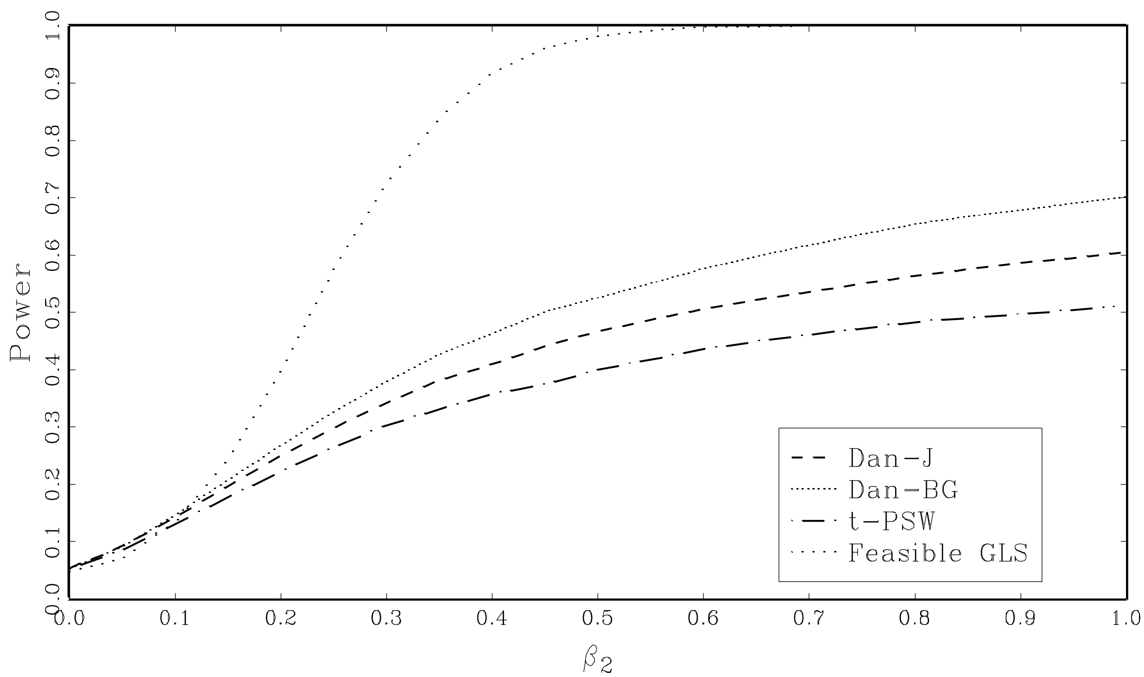

Figure 8: Finite Sample Power, AR(1) Errors, α=1.0, T=100.

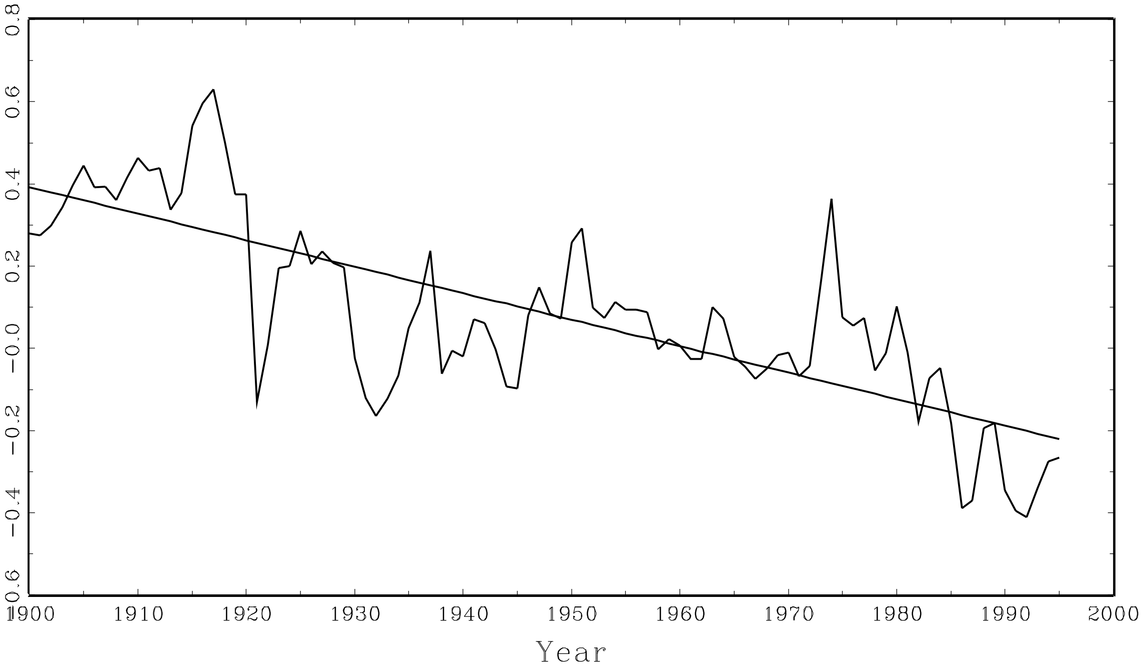

Figure 9: Logarithm of Net Barter Terms of Trade and Fitted Trend