Université de Montréal
Faculté des arts et des sciences
Département de sciences économiques

CAHIER 9427

USEFUL MODIFICATIONS TO SOME UNIT ROOT TESTS WITH DEPENDENT ERRORS AND THEIR LOCAL ASYMPTOTIC PROPERTIES

Pierre PERRON[1] and Serena NG[1]

[1] Département de sciences économiques and Centre de recherche et développement en économique (C.R.D.E.), Université de Montréal.

December 1994

The authors acknowledge grants from the Social Sciences and Humanities Research Council of Canada (SSHRC). The second author would like to thank also the Fonds pour la formation de chercheurs et l'aide à la recherche du Québec (FCAR) and the National Science Foundation for financial support.

C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7

Télécopieur (FAX): (514) 343-5831
Courrier électronique (E-Mail): econo@tornade.ERE.Umontreal.CA

Ce cahier a également été publié au Centre de recherche et développement en économique (C.R.D.E.) (publication no 3194).

Dépôt légal - 1994
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISSN 0709-9231

RÉSUMÉ

Plusieurs tests de racine unitaire ont des niveaux exacts déformés quand le processus d'erreur a une racine près du cercle unitaire. Cet article analyse les propriétés des tests de Phillips-Perron et quelques-unes de leurs variantes dans l'espace problématique. On utilise des analyses asymptotiques locales pour montrer pourquoi les tests de Phillips-Perron souffrent de distorsions de niveaux sévères, peu importe le choix de l'estimateur de la densité spectrale, et comment les statistiques modifiées occasionnent des améliorations dramatiques du niveau, lorsque utilisées en conjonction avec une formulation particulière d'un estimateur autorégressif de densité spectrale. On explique pourquoi les estimateurs de densité spectrale basés sur un noyau aggravent le problème de niveau dans les tests de Phillips-Perron et n'amènent aucune amélioration pour les statistiques modifiées. Les puissances asymptotiques locales des statistiques modifiées sont aussi évaluées. Ces statistiques modifiées sont recommandées dans les travaux empiriques étant donné qu'ils sont exempts du problème de niveau et qu'ils permettent aussi une puissance respectable.

Mots-clés : racine unitaire, asymptotique local, quasi-intégré, doublement intégré, intégration à fréquence saisonnière, Phillips-Perron.

ABSTRACT

Many unit root tests have distorted sizes when the root of the error process is close to the unit circle. This paper analyzes the properties of the Phillips-Perron tests and some of their variants in the problematic parameter space. We use local asymptotic analyses to explain why the Phillips-Perron tests suffer from severe size distortions regardless of the choice of the spectral density estimator but that the modified statistics show dramatic improvements in size when used in conjunction with a particular formulation of an autoregressive spectral density estimator. We explain why kernel-based spectral density estimators aggravate the size problem in the Phillips-Perron tests and yield no size improvement to the modified statistics. The local asymptotic power of the modified statistics is also evaluated. These modified statistics are recommended as being useful in empirical work since they are rid of size problems which have plagued many unit root tests and retain respectable power.

Key words: unit root, local asymptotic, near integrated, twice integrated, seasonally integrated, Phillips-Perron.

1. Introduction.

Testing for the presence of unit roots and cointegration is now a common practice in applied macroeconomics. Often, one is required to use statistics which appropriately account for serial correlation in the error process. Among statistics in this class, the augmented Dickey Fuller and the Phillips-Perron tests are perhaps the most popular, as they are implemented in many statistical software packages. However, it is also by now a well documented fact that the Phillips-Perron tests, as originally defined, suffer from severe size distortions when there are negative moving average errors. Although the size of the Dickey Fuller test is more accurate, the problem is not negligible.

The objectives of our paper are twofold. The first is to provide an understanding for the sources of size distortions in the Phillips-Perron tests and to explain how these distortions relate to the choice of the spectral density estimator. We use local asymptotic frameworks to analyze the properties of the statistics when the autoregressive or the moving average error process has a root close to the unit circle. We find that although no spectral density estimator can eliminate size distortions in these cases, kernel based spectral density estimators tend to aggravate the size problem.

Our second objective is to compare the properties of the Phillips-Perron tests with selected statistics which can be viewed as modified Phillips-Perron tests. These modified statistics, based originally on the work of Stock (1990), are found to have exact sizes much closer to the nominal size *when used in conjunction with a particular formulation of the autoregressive spectral density estimator*. However, these seemingly attractive properties do not generalize to kernel based spectral density estimators. Using local asymptotic analyses as developed in Nabeya and Perron (1994), we provide an explanation for these results, qualify the conditions when the modifications will alleviate size distortions, and when they will be vacuous.

This paper is organized as follows. Definitions of the statistics and estimators for the spectral density at frequency zero are given in Sections 1.1 and 1.2 respectively. The empirical properties of the statistics are presented in Section 1.3. Section 1.4 sets up the framework for analyzing data generating processes whose error has a root close to the unit circle. The next three sections analyze the theoretical properties of the statistics for different choices of the spectral density estimator. The local asymptotic size and power of the statistics are analyzed in Section 5. A conclusion completes the analysis. All proofs are contained in a mathematical appendix. Properties of the autoregressive spectral density estimator are discussed in more detail in Perron and Ng (1994).

1.1 The Test Statistics.

Consider the data generating process

$$y_t = \alpha\ y_{t-1} + u_t, \tag{1.1}$$

where $\{u_t\}$ is $i.i.d.(0,\sigma^2)$. White (1958) showed that the normalized least squares statistic, $T(\hat{\alpha}-1)$, and the t statistic for $\hat{\alpha}$, defined as $t_{\hat{\alpha}} = (\hat{\alpha}-1)/s_u(\sum_{t=1}^{T} y_{t-1}^2)^{-1/2}$ with $s_u^2 = T^{-1}\sum_{t=1}^{T}\hat{u}_t^2$, have the following asymptotic distributions:

$$T(\hat{\alpha}-1) \Rightarrow \left(\int_0^1 W(r)dW(r)\right)\left(\int_0^1 W(r)^2dr\right)^{-1}, \tag{1.2}$$

$$t_\alpha \Rightarrow \left(\int_0^1 W(r)dW(r)\right)\left(\int_0^1 W(r)^2dr\right)^{-1/2}, \tag{1.3}$$

where $W(r)$ is a standard Brownian motion on $C[0,1]$, the space of continuous functions on the interval [0,1], and $\Rightarrow$ denotes weak convergence in distribution. When $\{u_t\}$ is serially correlated, Phillips (1987) showed that, under some regularity conditions, the limiting distributions of the statistics become

$$T(\hat{\alpha}-1) \Rightarrow \left(\int_0^1 W(r)dW(r)+\lambda\right)\left(\int_0^1 W(r)^2dr\right)^{-1},$$

$$t_\alpha \Rightarrow (\sigma/\sigma_u)\left(\int_0^1 W(r)dW(r)+\lambda\right)\left(\int_0^1 W(r)^2dr\right)^{-1/2},$$

where $\lambda = (\sigma^2-\sigma_u^2)/2\sigma^2$, $\sigma_u^2 = \lim_{T\to\infty}\sum_{t=1}^{T} T^{-1}E[u_t^2]$, $\sigma^2 = \lim_{T\to\infty} E[T^{-1}S_T^2]$, and $S_T = \sum_{j=1}^{T} u_j$. When $\{u_t\}$ is stationary, $\sigma^2 = 2\pi f_u(0)$, where $f_u(0)$ is the non-normalized spectral density function of $\{u_t\}$ evaluated at frequency zero. In the case of martingale difference innovations, we have $\sigma^2 = \sigma_u^2$. To remove the dependence of the asymptotic distributions on the nuisance parameters σ^2 and σ_u^2, Phillips (1987) and Phillips and Perron (1988) proposed the statistics Z_α and Z_t, defined in the case of regression (1.1), as

$$Z_\alpha = T(\hat{\alpha}-1) - (s^2-s_u^2)\left(2T^{-2}\sum_{t=1}^{T} y_{t-1}^2\right)^{-1}, \tag{1.4}$$

$$Z_t = (s_u/s)t_\alpha - (1/2)(s^2-s_u^2)\left(s^2T^{-2}\sum_{t=1}^{T} y_{t-1}^2\right)^{-1/2}, \tag{1.5}$$

where s_u^2 and s^2 are consistent estimates of σ_u^2 and σ^2 respectively. We will frequently refer to the term $(s^2-s_u^2)(T^{-2}\sum_{t=1}^{T} y_{t-1}^2)^{-1}$ as the serial correlation correction factor. The Z_α and Z_t statistics, hereafter referred to as the PP tests, are often used in situations where considerations of weakly dependent errors become relevant. The asymptotic distributions of these statistics are given by (1.2) and (1.3) respectively.[1].

Stock (1990) proposed a class of statistics which exploits the feature that a series converges with different rates of normalization under the null and the alternative hypotheses. We consider two such tests, hereafter referred to as the M tests. The first statistic is MZ_α, defined as:[2]

$$MZ_\alpha = (y_T^2 - T\, s^2)\Big(2T^{-1}\textstyle\sum_{t=1}^{T} y_t^2\Big)^{-1}. \tag{1.6}$$

The statistic can be rewritten as

$$MZ_\alpha = Z_\alpha + (T/2)(\hat{\alpha} - 1)^2. \tag{1.7}$$

For this reason, MZ_α can be seen as a modified version of Z_α. The term $T(\hat{\alpha} - 1)^2/2$ will subsequently be referred to as the modification factor. Under standard assumptions, the result that $\hat{\alpha}$ converges to one at rate T ensures that Z_α and MZ_α are asymptotically equivalent. The critical values of MZ_α are therefore the same as Z_α, namely, those of the normalized least squares estimator given by (1.2). The second statistic, MSB, is defined as:

$$MSB = \Big(T^{-2}\textstyle\sum y_{t-1}^2/s^2\Big)^{1/2}. \tag{1.8}$$

Noting that the sum of squares of an $I(1)$ series is $O_p(T^2)$ but that of an $I(0)$ series is $O_p(T)$, the MSB statistic effectively tests the null hypothesis that the former condition is true. Under the alternative hypothesis, the statistic tends to zero. Hence, the unit root hypothesis is rejected in favor of stationarity when MSB is smaller than some appropriate critical value. Note that MSB is bounded from below by zero, unlike the other tests. The statistic is related to Bhargava's (1986) R_1 statistic which is built upon the work of Sargan and Bhargava (1983). Critical values with y_t demeaned and detrended are provided by Stock (1990).

Note that MSB and the PP tests are related as follows:

$$Z_t = MSB \cdot Z_\alpha. \tag{1.9}$$

This suggests the relationship $MZ_t = MSB \cdot MZ_\alpha$ should hold. Hence, we can define a new modified PP test as:

$$MZ_t = Z_t + (1/2)\Big(\textstyle\sum y_{t-1}^2/s^2\Big)^{1/2}(\hat{\alpha} - 1)^2. \tag{1.10}$$

[1] If there are additional deterministic components (constant and trends) in regression (1.1), the Weiner process $W(r)$ in (1.2) and (1.3) should be replaced by its detrended counterpart.

[2] The statistics are more generally implemented with y demeaned or detrended. For the purpose of this analysis, there is no loss in assuming this simpler specification where no deterministic components are present.

Thus, each PP test has a modified counterpart.

1.2 Estimating σ^2 and σ_u^2.

Construction of the statistics defined in the previous subsection requires estimates of σ_u^2 and/or σ^2. We consider three choices of s^2 as an estimator for σ^2. First, we assume that σ^2 is known, and the estimator is denoted σ_T^2. Second, we let σ^2 be estimated by an autoregressive spectral density estimator, defined as

$$s_{AR}^2 = s_{ek}^2/(1-\hat{b}(1))^2,$$

$s_{ek}^2 = T^{-1}\sum_{t=k+1}^{T}\hat{e}_{tk}^2$, $\hat{b}(1) = \sum_{j=1}^{k}\hat{b}_j$, with $\hat{b}_j$ and $\{\hat{e}_{tk}\}$ obtained from a k^{th} order augmented autoregression in Δy_t:

$$\Delta y_t = b_0 y_{t-1} + \sum_{j=1}^{k} b_j \Delta y_{t-j} + e_{tk}. \tag{1.11}$$

For roots of u_t bounded away from the unit circle, consistency of the parameters in the augmented autoregression (1.11) has been shown by Said and Dickey (1984) to hold under the null hypothesis that $\alpha = 1$ if $k = o(T^{1/3})$. Consistency of s_{AR}^2 based upon (1.11) follows from Said and Dickey's results. This formulation, used in Stock (1990), differs from the usual regression model used to construct s_{AR}^2, namely,

$$\hat{u}_t = \sum_{j=1}^{k} b_j \hat{u}_{t-j} + e_{tk}, \tag{1.12}$$

as discussed in Priestley (1981). Consistency of s_{AR}^2 based on (1.12) was proved in Berk (1974) for cases where $0 < \sigma^2 < \infty$.

The regression model (1.11) replaces $\hat{u}_t$ in (1.12) by Δy_t and includes y_{t-1} as a regressor. Since $\hat{u}_t = y_t - \hat{\alpha} y_{t-1}$ in our context, implicit in the use of (1.12) is the assumption that $\hat{\alpha}$ is consistent for α. Estimates of the spectral density implied by (1.11) and (1.12) are asymptotically equivalent when $\hat{\alpha}$ is a consistent estimate of $\alpha = 1$. As we will see, the advantage of defining s_{AR}^2 according to (1.11) is that it is does not depend on $\hat{\alpha}$ (through $\hat{u}_t$), and therefore decouples the estimation of α from the estimation of σ^2. This permits an estimator of σ^2 that is bounded below by 0 under both the null and the alternative hypotheses.

The third estimator of σ^2 considered is a kernel estimator based on the sample autocovariances. This last estimator, which we denote s_{WA}^2, is of the form

$$s_{WA}^2 = T^{-1}\sum_{t=1}^{T}\hat{u}_t^2 + 2T^{-1}\sum_{k=1}^{M} w(k/M)\sum_{t=k+1}^{T}\hat{u}_t\hat{u}_{t-k}, \tag{1.13}$$

where $\hat{u}_t$ are the least squares residuals from the first order autoregression (1.1). Following Andrews (1991), $w(\cdot)$ is a real-valued kernel in the set:

$\aleph = \{w(\cdot) : \Re \rightarrow [-1,1], w(0) = 1, \;\; w(x) = w(-x) \;\; \forall x \in \Re, \;\; \int_{-\infty}^{\infty} w(x)dx < \infty$, $w(x)$ is continuous at 0 and at all but a finite number of other points$\}$.

Estimators of σ^2 in the class of $\aleph$ produce estimates, s^2_{WA}, which are consistent for σ^2 provided $M/T \rightarrow 0$ and $M \rightarrow \infty$ as $T \rightarrow \infty$. We will restrict our analysis to kernels that also satisfy $w(x) = 0$ for $|x| > 1$. In that case, M acts as a truncation lag parameter, and we also define

$$\psi = \int_0^1 w(x)dx. \tag{1.14}$$

Throughout this paper, we make use of the following estimator of σ^2_u :

$$s^2_u = T^{-1}\sum_{t=1}^{T} \hat{u}^2_t. \tag{1.15}$$

1.3 Finite Sample Properties of the Statistics.

One problem with the PP tests is that they suffer from noticeable size distortions when the root of the error process is close to the unit circle. The problem of overrejecting the unit root hypothesis is particularly serious when the moving average errors have large negative serial correlation, as noted by the authors in their original work. Simulations of Schwert (1989), Kim and Schmidt (1990), Hyslop (1991), and Dejong, Nankervis, Savin and Whiteman (1992) among others, confirmed this result. Although less severe, size problems also exist in models with autoregressive errors that have roots close to -1 or 1. In the former case, the PP tests reject the unit root hypothesis far too often, and in the latter case, not often enough.

To be precise about the issue at hand, Table 1 presents results based on 1000 replications of the DGP $y_t = y_{t-1} + u_t$, with $(1-\rho L)u_t = (1+\theta L)e_t$. The noise function, u_t, is a pure moving average process when $\rho = 0$, and a pure autoregressive process when $\theta = 0$. The regression is $y_t = \mu + \alpha\, y_{t-1} + v_t$. The size of the PP tests was evaluated at various sample sizes and the results for $T = 100$ and 500 are selected for discussion. In column one, we report results based on the Bartlett window using an automatic selection of the bandwidth as discussed in Andrews (1991). In column two, we present results for the autoregressive spectral density estimator formulated according to (1.11) in terms of Δy_t.[3] Results with σ^2 assumed known are reported in column three. For the sake of comparison, the size of the t_ρ statistic proposed by Dickey and Fuller (1979) and extended by Said and Dickey (1984) is

[3]The truncation lag is set according to the rule $k = (T/100)^{1/4}$.

given in the last column of Table 1.[4] Critical values are taken as the left tail 5 percentage point of the distribution given in Fuller (1976).

As we can see from the results, the PP tests are too liberal when θ is negative. Size distortions are noticeable even when θ is around -0.5, with the unit root hypothesis always being rejected as θ approaches -1. For autoregressive noise functions, the tests are too conservative when the residual autocorrelation is positive and too liberal when the residual autocorrelation is negative. The augmented Dickey-Fuller t_ρ test also has substantial size distortions in the negative moving-average case, although the problem is somewhat less severe than with the PP tests. As expected, t_ρ has good size properties with an autoregressive noise component.

Two explanations to these results seem possible. First, the PP tests may require s^2 to be a good estimator of σ^2. In the simulations, we have used a truncation lag of M and Bartlett weights of $1 - k/(M+1)$ to weigh the sample autocovariances at lag k. However, there is evidence[5] that the Bartlett window leads to estimates of σ^2 that are inferior to the Quadratic, the Parzen, and the Bohman windows. The inadequacy of s^2, which is not required in the construction of t_ρ, could be responsible for the behavior of the PP tests.

The choice of the kernel on the size of the PP tests was analyzed in Kim and Schmidt (1990) via simulations. These authors experimented with the Bohman, the Bartlett, and the Parzen windows and found the choice of the kernel not to make a significant difference as far as size distortions are concerned. Our own simulations also found the Quadratic window, reported to have good properties by Andrews (1991), incapable of resolving the size problem. The prewhitening procedure recently developed by Andrews and Monahan (1992) also does not provide improvements. More importantly, Table 1 suggests that replacing the kernel-based estimator of σ^2 by an autoregressive spectral density estimator will not reduce size distortions in the PP tests. Indeed, size problems persist even when σ^2 is assumed known.

The second explanation for the poor size is that the PP tests are based on large sample considerations. The recorded size distortions could be due to poor approximations provided by (1.2) and (1.3) for the finite sample properties of the statistics. However, we found the size problem to persist even with sample sizes as large as 500 or 1000, in accordance with the results of Schwert (1989).

Stock (1990) evaluated the properties of MZ_α and MSB using an autoregressive spectral density estimator. We extended his analysis and used alternative spectral density estimators

[4] The truncation lag is selected according to the Akaike Information Criteria.

[5] The exact error of fourteen kernel based spectral density estimators are analyzed in Perron and Ng (1993).

to construct the statistics. The results are presented in Table 2. An interesting result stands out. When the autoregressive spectral density estimator as defined in (1.11) is used to construct the statistics, the size distortions diminish dramatically and the exact sizes are relatively close to the nominal size of 5 percent. However, size distortions remain significant when σ^2 is estimated by sample autocovariances with Bartlett weights. Hyslop (1991) found the same results using the Parzen window, and we obtained similar results with a Quadratic window. The method used to estimate σ^2 evidently has important size implications for the M tests. As we recall, large size distortions are recorded for the PP tests regardless of the choice of s^2. This suggests an intricate interplay between the choice of s^2 and the properties of the M tests. These issues will be the subject of our analysis.

1.4 Preliminaries.

Our interest is in understanding the properties of the PP and the M tests when the root of the error process is close to the unit circle. A related problem was analyzed by Perron (1992) and Nabeya and Perron (1994) who studied the local asymptotic behavior of $\hat{\alpha}$ when θ is close to -1 and $|\rho|$ local to 1. The authors found that (1.2) indeed provides a poor guide to the distribution of $T(\hat{\alpha}-1)$ in the mentioned parameter space. Since the PP tests are built upon $T(\hat{\alpha}-1)$, one might also expect (1.2) and (1.3) to provide inadequate approximations for Z_α and Z_t. Some results on this issue were provided by Pantula (1991) who used a different local asymptotic framework to study the behavior of the PP tests when θ is local to -1. The author showed for the case when σ^2 is estimated by the Bartlett window that the distributions for Z_α and Z_t are unbounded and that the rate of divergence to $-\infty$ depends on the rate θ approaches -1 as well as the rate of increase of the truncation lag. However, it is unclear whether these results generalize to other estimators of σ^2, and whether the distributions of the statistics also have unusual properties when the autoregressive error has a root close to one.

This paper provides further understanding to the source of size distortions in unit root tests by extending previous work in two directions. First, we derive the local asymptotic distributions for the PP and the M tests for both autoregressive and moving average error processes with a root close to the unit circle. Our local asymptotic analysis is based on the following framework:

$$y_t = (1 + c/T)y_{t-1} + u_t, \tag{1.16}$$

where $\rho(L)u_t = \theta(L)e_t$, and $e_t \sim i.i.d(0, \sigma_e^2)$, $y_0 = e_0 = 0$. The series $\{y_t\}$ has an autoregressive root local to unity with non-centrality parameter c. Under the null hypothesis of a unit root, $c = 0$. Our attention will be restricted to the simple cases where $\{u_t\}$ is either a pure AR or a pure MA process with roots local to the boundary of -1 and/or 1.

Second, we analyze the dependence of the statistics on the choice of the spectral density estimator. The issue is relevant because as we shall see, the properties of s^2 are affected by those of $\hat{\alpha}$ via the estimated residuals. We have seen from Tables 1 and 2 that size distortions are much larger with a kernel based estimator of σ^2 than with the autoregression based alternative.

Our derivations will be based on a regression without a constant or a trend, but it is straightforward to generalize the results to encompass additional deterministic terms. In most cases, one can substitute $W(r)$ by a demeaned or a detrended Weiner process without altering any of the analytics.

2. The Nearly White Noise Nearly Integrated Case.

In this section, we analyze the properties of the tests for the case where a large negative moving average component is present. Following the framework used by Nabeya and Perron (1994), we specify the data generating process as:

$$y_t = (1 + c/T)y_{t-1} + u_t, \tag{2.1}$$

$$u_t = e_t + \theta_T e_{t-1}, \tag{2.2}$$

$$\theta_T = -1 + \delta/\sqrt{T}. \tag{2.3}$$

The process defined by (2.1) to (2.3) is an ARMA(1,1) with an autoregressive root local to unity and a moving average root local to -1. While the roots cancel and $\{y_t\}$ is a white noise process in the limit, it is nearly integrated in finite samples. Theorem 1 of Nabeya and Perron (1994) showed that for $y_0 = e_0 = 0$,

$$(\hat{\alpha} - 1) \Rightarrow -\left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{-1}, \tag{2.4}$$

where $J_c(r) = \int_0^r \exp((r-s)c)dW(s)$. Since the limit of $(\hat{\alpha} - 1)$ has a negative support, the normalized least squares statistic $T(\hat{\alpha} - 1)$ is unbounded and converges to $-\infty$. Note that $\hat{\alpha}$ is not a consistent estimate of α in this local asymptotic framework. Furthermore, it is easy to show that given

$$s_u^2 \Rightarrow \sigma_e^2\left(2 - (1 + \delta^2 \int_0^1 J_c(r)^2 dr)^{-1}\right), \tag{2.5}$$

s_u^2 is also not a consistent estimate of $\sigma_u^2 = 2\sigma_\epsilon^2$. Since s_u^2 is based on the estimated residuals, the inconsistency of $\hat{\alpha}$ directly affects the properties of s_u^2. Using (2.4) and (2.5) it is straightforward to show that

$$T^{-1/2}t_\alpha \Rightarrow -\left(1 + 2\delta^2 \int_0^1 J_c(r)^2 dr\right)^{-1/2}. \tag{2.6}$$

Thus, t_α also diverges to $-\infty$, albeit at a slower rate then the normalized least squares statistic. The following theorems, proved in the Appendix, characterize the asymptotic distributions of Z_α, Z_t, MZ_α and MSB for the different estimators of σ^2. We start with the case where σ^2 is assumed known, followed by the autoregressive spectral density estimator.

Theorem 2.1 : *Let $\{y_t\}$ be defined by (2.1) to (2.3) with $y_0 = e_0 = 0$, and $s^2 = \sigma_T^2 = 2\pi f_u(0) = \sigma_\epsilon^2\delta^2/T$. Let $e_\infty = \lim_{T\to\infty} e_T/\sigma_\epsilon$. Then as $T \to \infty$:*

a. $T^{-1}Z_\alpha \Rightarrow -(1/2)\left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{-2}$;

b. $T^{-1}Z_t \Rightarrow -\left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{-3/2} /(2\delta)$;

c. $MZ_\alpha \Rightarrow ((e_\infty + \delta J_c(1))^2 - \delta^2)\left(2(1 + \delta^2 \int_0^1 J_c(r)^2 dr)\right)^{-1}$;

d. $MSB \Rightarrow \left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{1/2} /\delta$.

Theorem 2.2 : *Let $\{y_t\}$ be defined by (2.1) to (2.3) with $y_0 = e_0 = 0$ and s_{AR}^2 be based upon OLS applied to (1.11). Furthermore, assume that $k/T \to 0$ and $k \to \infty$ as $T \to \infty$. The asymptotic distributions of Z_α, Z_t, MZ_α, and MSB are the same as those given in Theorem 2.1.*

Remark 1

- The spectral density at the zero frequency is zero in the limit for the present model, a situation ruled out by assumption in Berk (1974). We cannot appeal to his results or those of Said and Dickey (1984) to prove consistency of s_{AR}^2. However, we show in Perron and Ng (1994) that $\hat{b}(1)$ based upon (1.11) diverges to infinity, and that s_{ek}^2 converges to σ_ϵ^2. Thus, s_{AR}^2 tends to zero, the asymptotic value of σ_T^2. The equivalence of the results with Theorem 2.1 follows from the consistency of s_{AR}^2 for σ^2.

- The spectral density at frequency zero plays no role in the PP tests in the limit because it converges to 0 whereas s_u^2 is $O_p(1)$. However, the use of the rate T^{-2} to normalize the sample moments of y_{t-1}^2 is too strong for this nearly integrated nearly white noise process and causes the serial correlation correction factor to be $O_p(T)$. The correction factor therefore neither aggravates nor annihilates the explosive nature of $T(\hat{\alpha}-1)$ and the PP tests continue to diverge at rate T. Further inspection reveals that the limiting distribution for Z_α is that of $(T/2)(\hat{\alpha} - 1)^2$.

• Although the normalized least squares estimator, the correction for serial correlation, and the modification factor in MZ_α all explode at rate T, the distribution of MZ_α is bounded. This is because the explosive terms in the PP tests are offset by the modification factor asymptotically. As this last term is absent from the PP tests, their asymptotic equivalence with the M tests breaks down under the assumptions of Theorems 2.1 and 2.2.

• Size distortions associated with MZ_α are due to the discrepancies between (1.3) and the distribution implied by (1.6) evaluated at the null hypothesis of $c = 0$. The dramatically smaller size distortions reported in Table 1 for MZ_α suggest that while the approximation provided by (1.3) for MZ_α is not perfect, it is a substantial improvement compared to Z_α. The adequacy of the local asymptotic approximations will be analyzed in more detail in Section 5.

Theorem 2.3 : *Let $\{y_t\}$ be defined by (2.1) to (2.3) with $y_0 = \epsilon_0 = 0$, s^2_{WA} defined by (1.13) and ψ by (1.14). Assume that $M \to \infty$ and $M/T \to 0$ as $T \to \infty$. Then:*

a. $(MT)^{-1} Z_\alpha \Rightarrow -\psi\delta^2 \left(\int_0^1 J_c(r)^2 dr\right) \left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{-3}$;

b. $(MT)^{-1/2} Z_t \Rightarrow -\left(\psi(\delta^2/2) \int_0^1 J_c(r)^2 dr\right)^{1/2} \left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{-3/2}$;

c. $(MT)^{-1} MZ_\alpha \Rightarrow -\psi\delta^2 \left(\int_0^1 J_c(r)^2 dr\right) \left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{-3}$;

d. $(MT)^{1/2} MSB \Rightarrow \left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{3/2} \left(2\psi\delta^2 \int_0^1 J_c(r)^2 dr\right)^{-1/2}$.

Remark 2

• The PP tests in Theorem 2.3 with s^2_{WA} based on weighted estimated autocovariances explode to $-\infty$ at a rate of MT, faster than the rate T when σ^2 is presumed known or estimated by s^2_{AR}. With Bartlett weights of $1 - k/(M+1)$, $\psi = 1/2$. In general, different choices of the weighting function will only affect the magnitude of ψ but not the rate at which the statistics explode. This confirms the result of Kim and Schmidt (1990) that size distortions cannot be eliminated by suitable choice of the kernel.

An intuitive explanation of these results is as follows. The quantity s^2_{WA} is based upon a weighted sum of sample autocovariances which inherit inconsistency from $\hat{\alpha}$. Divergence arises not from the inconsistency of s^2_{WA} per se, but from the inappropriate use of the rate T^2 to normalize the sample moments of y^2_{t-1}. Each normalized autocovariance is therefore $O_p(T)$, and the explosive terms cumulate as the M lags of sample autocovariances are summed up. For this reason, the serial correlation correction factor is $O_p(MT)$ and accounts for the reported rate of approach to $-\infty$. As well, the modification factor in MZ_α, $(T/2)(\hat{\alpha}-1)^2$, is only $O_p(T)$, and is not strong enough to offset the normalized autocovariances which diverge at a faster rate. Thus, MZ_α also diverges to negative infinity and has the same asymptotic distribution as Z_α.

• To the extent that the local asymptotic distributions of the PP tests diverge to $-\infty$ at rate MT, the higher the order of the truncation lag, the faster the rate of divergence. Thus, although a long truncation lag is preferable when estimating the spectral density at frequency zero for a stationary series with a large negative moving average component, see Perron and Ng (1993), a large truncation lag is not optimal when there is a near common factor.

Remark 3

• Phillips and Ouliaris (1990) showed that the PP tests will be inconsistent against stationary alternatives under the standard asymptotic framework (i.e. θ fixed as T increases) if residuals under the null hypothesis, $\Delta y_t = u_t$, are used to construct s^2_{WA}. Theorem 2.3 shows that the use of the estimated residuals, $\{\hat{u}_t\}$, will lead to increasing size distortions in the present local asymptotic framework. Thus, estimators of the form (1.13) based on sample autocovariances are inadequate whether one uses the estimated residuals or the residuals under the null.

• In view of the inconsistency of $\hat{\alpha}$, the formulation of the autoregressive spectral density estimator based on Δy_t is not asymptotically equivalent to that based on $\hat{u}_t$. Specifically, (1.11) does not have a first order dependence on $\hat{\alpha}$ as compared to a regression based on $\hat{u}_t$. Simulations show that if s^2_{AR} is based upon (1.12), the statistics will behave much like those reported in Table 1 for the Bartlett window. Intuitively, the reason is that the moment matrix of regressors in (1.12) does not converge to the population moments, and estimates from (1.12) cannot be used to obtain a consistent estimate of σ^2. An important implication is that, among those estimators for σ^2 considered, the one that is adequate in both the standard and the local asymptotic framework is that based on the autoregression of the form (1.11) using first differences of $\{y_t\}$ instead of the estimated autoregression residuals.

• Although our discussion has focused on the properties of MZ_α, the intuition applies to MSB and MZ_t as well. For example, under Theorems 2.1 and 2.2, we have $T^{-2}\sum_{t=1}^{T} y_{t-1}^2$ tending to ∞ and σ_T^2 tending to zero at the same rate. The resulting distribution for MSB is therefore neither degenerate nor explosive. It can be also seen that MSB tends to zero in Theorem 2.3 because of the unusually slow rate of convergence of $\sum_{t=1}^{T} y_t^2$. Since the test rejects the null hypothesis of a unit root the closer the statistic is to zero, the probability that MSB rejects the null hypothesis converges to one.

• The rate of divergence of the PP tests depends on whether one uses s_{AR}^2 or s_{WA}^2. For the former, divergence is at rate T as noted under Remark 1. For the latter, the speed of divergence is MT. For this reason, size distortions reported in Table 1 are larger for s_{WA}^2 than for s_{AR}^2.

3. The Nearly Twice Integrated Case.

The aim of this section is to study the behavior of the statistics in the presence of autoregressive errors with a large positive coefficient. The data generating process is:

$$y_t = (1 + c/T)y_{t-1} + u_t, \tag{3.1}$$

$$u_t = (1 + \phi/T)u_{t-1} + e_t, \tag{3.2}$$

where $e_t \sim i.i.d.(0, \sigma_e^2)$. As $T \to \infty$, $y_t = 2\, y_{t-1} - y_{t-2} + e_t$, a process with two unit roots, hence the terminology "nearly twice integrated". Nabeya and Perron (1994) showed that

$$T(\hat{\alpha} - 1) \Rightarrow Q_c(J_\phi(1))^2 \left(2 \int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1}, \tag{3.3}$$

where $Q_c(J_\phi(r)) = \int_0^r \exp((r - v)c) J_\phi(v) dv$ and $J_\phi(v) = \int_0^v exp((v - s)\phi) dW(s)$. Note that unlike the previous model, $\hat{\alpha}$ is consistent, in the sense that $\hat{\alpha} - (1 + c/T) \to 0$. It can be verified that

$$T^{-1/2} t_\alpha \Rightarrow \left(Q_c(J_\phi(1))^2/2\right) \left(\lambda \int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1/2}, \tag{3.4}$$

where

$$\lambda = \int_0^1 J_\phi(r)^2 dr - \left(Q_c(J_\phi(1))^2 - 2c \int_0^1 Q_c(J_\phi(r))^2 dr\right)^2 \left(4 \int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1} \tag{3.5}$$

$$\text{and} \qquad T^{-1} s_u^2 = T^{-2} \sum \hat{u}_t^2 \Rightarrow \lambda \sigma_e^2. \tag{3.6}$$

Although inconsistency of $\hat{\alpha}$ is not an issue in this model, the rates of normalization for the relevant partial sums are still different from those in the usual asymptotic framework. In particular, s_u^2 is $O_p(T)$, and $\sum_{t=1}^T y_t^2$ is $O_p(T^4)$. The asymptotic properties of the PP tests in this local framework are direct consequences of these results. The following Theorems summarize the properties of the statistics for various ways of estimating σ^2.

Theorem 3.1 : *Let $\{y_t\}$ be a stochastic process given by (3.1) and (3.2), $T^{-2}s^2 = T^{-2}\sigma_T^2 = \sigma_e^2/\phi^2$. Then as $T \to \infty$ and if $\phi < 0$, we have:*

a. $Z_\alpha \Rightarrow \left(Q_c(J_\phi(1))^2 - 1/\phi^2\right)\left(2\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1}$;

b. $Z_t \Rightarrow (\phi/2)\left(Q_c(J_\phi(1))^2 - 1/\phi^2\right)\left(\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1/2}$;

c. $MZ_\alpha \Rightarrow \left(Q_c(J_\phi(1))^2 - 1/\phi^2\right)\left(2\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1}$;

d. $MSB \Rightarrow \left(\phi^2 \int_0^1 Q_c(J_\phi(r))^2 dr\right)^{1/2}$.

Theorem 3.2 : *Let $\{y_t\}$ be defined by (3.1) to (3.2). Let the estimator of σ^2 be the autoregressive spectral density estimator s_{AR}^2 defined by (1.11) with the truncation lag, k, chosen to be such that $k^3/T \to 0$ and $k \to \infty$ as $T \to \infty$. Let $T(\hat{b}(1) - b(1)) \to \eta$, with η defined in Perron and Ng (1994). We have*

a. $Z_\alpha \Rightarrow \left(Q_c(J_\phi(1))^2 - 1/(c+\phi+\eta)^2\right)\left(2\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1}$;

b. $Z_t \Rightarrow (1/2)(c+\phi+\eta)\left(Q_c(J_\phi(1))^2 - 1/(c+\phi+\eta)^2\right)\left(\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1/2}$;

c. $MZ_\alpha \Rightarrow \left(Q_c(J_\phi(1))^2 - 1/(c+\phi+\eta)^2\right)\left(2\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1}$;

d. $MSB \Rightarrow \left((c+\phi+\eta)^2 \int_0^1 Q_c(J_\phi(r))^2 dr\right)^{1/2}$.

Remark 4

• Said and Dickey's proof of consistency of the coefficients $\hat{b}_i$ in (1.11) requires that $1 - b(1)$ be bounded away from zero, a condition which is not satisfied in the limit since the autoregressive coefficient for $\{u_t\}$ converges to 1. Indeed, $\sigma_T^2 \to \infty$ for the data generating process in question. However, in this limiting case, Δy_t is an integrated process and the coefficients on Δy_{t-i} are order T consistent (see Park and Phillips (1988) for the case of a finite order autoregression). In consequence, s_{AR}^2 also tends to ∞ and is equivalent to σ_T^2 asymptotically.

• Even though s_{AR}^2 and σ_T^2 both tend to ∞, the limiting distributions of the statistics are not the same for the two cases. The results of Theorems 3.1 and 3.2 differ in that ϕ is replaced by $c + \phi + \eta$. This is because $T^{-2}s_{AR}^2 \to \sigma_\epsilon^2/(c+\phi+\eta)^2$, whereas the true the spectral density at frequency zero satisfies $T^{-2}\sigma_T^2 = \sigma_\epsilon^2/\phi^2$. Thus, the limiting distributions of the statistics based on s_{AR}^2 contain the variable η even under the null hypothesis that $c = 0$.

• Unlike in the previous model where the serial correlation correction factor is dominated by s_u^2, it is s^2 that dominates in this model. This is because s_u^2 diverges to ∞ at a slower rate than s^2 by a factor of T. When normalized by the sample moments of y_{t-1}^2, the effect of s_u^2 vanishes completely.

• The modification factor in MZ_α can be written as $(2T)^{-1}[T(\hat{\alpha}-1)]^2$. In view of (3.3), this expression has a probability limit of zero, and is the reason why Z_α and MZ_α have the same limiting distribution. This result also confirms that given the consistency of $\hat{\alpha}$, the asymptotic equivalence of Z_α and MZ_α holds. Thus, under the assumptions of this model, the modifications to the PP tests are vacuous.

Theorem 3.3 : *Let $\{y_t\}$ be defined by (3.1) and (3.2) and let σ^2 be estimated by s^2_{WA} as defined by (1.13) with truncation lag M. Let λ be defined by (3.5), $Q_c(J_\phi(r))$ by (3.3) and ψ by (1.14). If $M \to \infty$ and $M/T \to 0$ as $T \to \infty$, we have:*

a. $Z_\alpha \Rightarrow Q_c(J_\phi(1))^2 \left(2\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1}$;

b. $(T/M)^{-1/2} Z_t \Rightarrow (1/2)\left(Q_c(J_\phi(1))^2\right)\left(2\psi\lambda\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1/2}$;

c. $MZ_\alpha \Rightarrow Q_c(J_\phi(1))^2 \left(2\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{-1}$;

d. $(T/M)^{-1/2} MSB \Rightarrow \left(\int_0^1 Q_c(J_\phi(r))^2 dr\right)^{1/2} (2\psi\lambda)^{-1/2}$.

Remark 5

• Given that u_t is nearly integrated, its autocovariances die off very slowly. One might conjecture that a truncation lag chosen according to the rule of $o(T^{1/4})$ as suggested in Phillips (1987) might not be appropriate in this situation. However, the results given in Theorem 3.3 hold under very flexible conditions on M, requiring only that $M/T \to 0$ and $M \to \infty$ as $T \to \infty$. Yet, the proof of Theorem 3.3 shows that even when a wider choice of the truncation lag is allowed, Z_α and MZ_α will still have the same limiting distribution given in (3.3) for $T(\hat{\alpha}-1)$. Evidently, neither the correction for serial correlation, nor the modification made to MZ_α has any effect on the statistics asymptotically. This is in spite of the result that $(s^2_{WA} - s^2_u)$ diverges to infinity at rate MT.

To understand these results, note that $T^{-2}\sum_{t=1}^T y_{t-1}$ diverges at rate T^2. This is faster than the rate at which $(s^2_{WA} - s^2_u)$ diverges since M increases at a rate slower than T by assumption. The serial correlation correction factor therefore tends to zero, and Z_α has the same limiting distribution as $T(\hat{\alpha}-1)$. Since the modification factor has a limit of zero as well, the asymptotic distribution of MZ_α is also that of the normalized least squares estimator. Thus, neither the choice of the kernel function nor the truncation lag will change the asymptotic properties of the statistics under the assumptions of Theorem 3.3.

Remark 6

• Since most of the statistics are bounded under the parameterization of this section, the discrepancies between the approximate and the exact distributions are not as large as in the model considered in Section 2, and the size distortions are accordingly smaller. The size distortions for Z_α observed in Tables 1 and 2 when ρ is close to 1 can be traced to the fact that the local asymptotic distribution of $T(\hat{\alpha} - 1)$ is non-negative, but the critical values based on (1.2) are negative. Hence, the $T(\hat{\alpha} - 1)$ statistic will yield a zero rate of rejection of the null hypothesis against a sequence of stationary local alternatives if critical values from (1.2) are used. Nabeya and Perron (1994) provide a detailed explanation for why $T(\hat{\alpha} - 1)$ is undersized when there are close to two unit roots.

When the statistic is adjusted for the presence of serial correlation, Z_α and MZ_α are being shifted to the left of $T(\hat{\alpha} - 1)$ by a quantity that depends on the choice of s^2. The left shifting effect is nil in the case of s^2_{WA} as the results of Theorem 3.3 suggest. Since the statistics behave in the limit as $T(\hat{\alpha} - 1)$, they too are undersized when there are close to two unit roots in the *DGP*.

When σ^2 is known or is estimated by s^2_{AR}, the location of Z_α and MZ_α relative to $T(\hat{\alpha}-1)$ depends on the magnitude of $1/\phi$ and $1/(c+\phi+\eta)$ respectively. The closer is ϕ or $c+\phi+\eta$ to zero, the further will the local asymptotic distributions of Z_α and MZ_α lie to the left of that of $T(\hat{\alpha}-1)$. In those cases, one can expect the critical values provided by (1.2) to reject the null hypothesis of one unit root too often. For certain parameterizations of the data generating process, there is a possibility that the statistics will reject the null hypothesis of one unit root in favour of stationarity even when there are nearly two unit roots.

4. The Nearly Integrated Seasonal Model.

The aim of this section is to study the behavior of the statistics in the presence of autoregressive errors with large negative coefficients. Consider the following data generating process:

$$y_t = (1 + c/T)y_{t-1} + u_t = \alpha_T \ y_{t-1} + u_t; \tag{4.1}$$

$$u_t = -(1 + \phi/T)u_{t-1} + e_t = \rho_T \ u_{t-1} + e_t. \tag{4.2}$$

It is easy to see that the model can be written as $y_t = [(1 + c/T) - (1 + \phi/T)]y_{t-1} + (1 + (c+\phi)/T)y_{t-2} + e_t$. As $T \to \infty$, the process becomes a seasonal model of period two with a root on the unit circle. That is, $y_t = y_{t-2} + e_t$. As shown in Nabeya and Perron (1994),

$$\hat{\alpha} \Rightarrow 1 - \left(2\int_0^1 B(r)^2 dr\right)\left(\int_0^1 (C(r)^2 + B(r)^2)dr\right)^{-1}, \tag{4.3}$$

where $A(r) = (\phi - c)[Q_c(J_{\phi,1}(r)) - Q_c(J_{\phi,2}(r))] + 2\ J_{c,1}(r)$, $B(r) = J_{\phi,1}(r) - J_{\phi,2}(r)$, $C(r) = A(r) - B(r)$, $J_{c,1}(s) = \int_0^s \exp((s-v)c)dW_1(v)$, $J_{\phi,i}(s) = \int_0^s \exp((s-v)\phi)dW_i(v)$, $Q_c(J_{\phi,i}(r)) =$

$\int_0^r \exp((r-s)c)J_{\phi,i}(s)ds$ for $i=1,2$, $W_1(r)$ and $W_2(r)$ being independent Weiner processes. As in the moving average model of Section 2, $\hat{\alpha}$ is not a consistent estimate of α. The support of the limiting distribution of $(\hat{\alpha}-1)$ is restricted to the interval $[-2,0]$ since $\hat{\alpha}$ is of the form $(a-b)/(a+b)$. Hence, the statistic$T(\hat{\alpha}-1)$ is unbounded and diverges to $-\infty$ as T increases, just as in the first model. Since, in the limit, the error process has an autoregressive root on the unit circle, it also has a property in common with the model of Section 3, namely, that the limiting variance of $\{u_t\}$ is infinite. Specifically, define

$$\lambda_1 = (1/2)\left(\int_0^1 B(r)^2dr \int_0^1 C(r)^2dr\right)\left(\int_0^1 C(r)^2 + B(r)^2dr\right)^{-1}. \tag{4.4}$$

Then

$$T^{-1}s_u^2 \Rightarrow \lambda_1\sigma_\epsilon^2. \tag{4.5}$$

For future reference, it is also convenient to define

$$\lambda_2 = \left((\int_0^1 B(r)^2dr)/(\int_0^1 C(r)^2 + B(r)^2dr)\right)^2\left(\int_0^1 C(r)^2 - B(r)^2dr\right). \tag{4.6}$$

Straightforward calculations show that

$$T^{-1/2}t_\alpha \Rightarrow -(\int_0^1 B(r)^2dr/\int_0^1 C(r)^2dr)^{1/2}. \tag{4.7}$$

The following two theorems characterize the distributions of the various statistics when σ^2 is presumed known and when estimated by the autoregressive spectral density estimator.

Theorem 4.1 : *Let $\{y_t\}$ be a stochastic process generated by (4.1) and (4.2) with $\phi < 0$, and $\sigma_T^2 = 2\pi f_u(0) = \sigma_\epsilon^2/(2+\phi/T)^2$. As $T \to \infty$, we have:*

a. $T^{-1}Z_\alpha \Rightarrow -2\left(\int_0^1 B(r)^2dr\right)^2\left(\int_0^1 C(r)^2 + B(r)^2dr\right)^{-2}$;

b. $T^{-1}Z_t \Rightarrow -\sqrt{2}\left(\int_0^1 B(r)^2dr\right)^2\left(\int_0^1 C(r)^2 + B(r)^2dr\right)^{-3/2}$;

c. $MZ_\alpha \Rightarrow \left(\frac{A(1)^2}{2}-1\right)\left(\int_0^1 C(r)^2 + B(r)^2dr\right)^{-1}$;

d. $MSB \Rightarrow \left(\frac{1}{2}\int_0^1 C(r)^2 + B(r)^2dr\right)^{\frac{1}{2}}$.

Theorem 4.2 : *Let $\{y_t\}$ be defined by (4.1) to (4.2) with $\phi < 0$. Let the estimator of σ^2 be the autoregressive density estimator s_{AR}^2 defined by (1.11) with the truncation lag k chosen to be such that $k^3/T \to 0$ and $k \to \infty$ as $T \to \infty$. The statistics Z_α, Z_t, MZ_α, and MSB have the same limiting distributions as given in Theorem 4.1 for the case where σ^2 is known.*

Remark 7

- The autoregressive representation of this model also has a unit root, a case ruled out by Said and Dickey (1984). However, we show in Perron and Ng (1994) that s^2_{AR} converges to $\sigma^2_\epsilon/4$, the same limiting value as $\sigma^2_T = \sigma^2_\epsilon/(2 + \phi/T)^2$. The asymptotic equivalence of the results in the two theorems follows.

- The asymptotic equivalence between MZ_α and Z_α breaks down in this model, just as in the nearly integrated, nearly white noise case. In both models, the two PP tests diverge but the M tests do not. Furthermore, $\hat{\alpha}$ is not a consistent estimate of α in either case, and $T^{-1}\sum_{t=1}^T y_{t-1}u_t \Rightarrow -1/2\sigma^2_u$ in both cases. Thus, the analysis of Section 2 applies: the PP tests are driven by $(T/2)(\hat{\alpha}-1)^2$ in the limit, but this explosive term is being offset by the modification factor in the M tests. The main difference between the two models is that in the negative MA case, $\sum_{t=1}^T y^2_{t-1}$ and $\sum_{t=1}^T \hat{u}^2_t$ are both $O_p(T)$, but in the negative AR case, they are both $O_p(T^2)$. In each case, the usual rates of normalization of T^2 and T for some of these sample moments are inappropriate.

Theorem 4.3 : *Let $\{y_t\}$ be defined by (4.1) and (4.2) and ψ by (1.14). Let s^2_{IVA} be defined by (1.13) with truncation lag M. If $M \to \infty$ and $M/T \to 0$, then as $T \to \infty$,*

a. $(MT)^{-1}Z_\alpha \Rightarrow (-4\psi\lambda_2)\left(\int_0^1 C(r)^2 + B(r)^2 dr\right)^{-1}$;

b. $(MT)^{-1/2}Z_t \Rightarrow -(2\psi\lambda_2)^{1/2}\left(\int_0^1 C(r)^2 + B(r)^2\right)^{-1/2}$;

c. $(MT)^{-1}MZ_\alpha \Rightarrow (-4\psi\lambda_2)\left(\int_0^1 C(r)^2 + B(r)^2 dr\right)^{-1}$;

d. $(MT)^{1/2}MSB \Rightarrow \left(\int_0^1 C(r)^2 + B(r)^2\right)^{1/2}(8\psi\lambda_2)^{-1/2}$.

Remark 8

- As in the nearly twice integrated model where the autoregressive root of u_t lies close to the unit circle, the autocovariances of u_t in the nearly seasonally integrated model also die off very slowly and a longer truncation lag would seem necessary. The results of Theorem 4.3 suggest that the statistics will still be divergent for any choice of the truncation lag. Since $\hat{\alpha}$ is not a consistent estimate of α, the sample autocovariances constructed on the basis of $\hat{u}_t$ are also inconsistent estimates of the true autocovariances. However, size distortions are smaller in this model with negative autocorrelation in the residuals than in the earlier model with negative moving average errors. This is because for a ρ and a θ of the same size, the implied value of ϕ is larger than the implied value of δ for a given T. In this sense, a ρ of -0.8 is further away from the boundary than a θ of -0.8. The simulations reflect the different rates of normalization on the sequence of local alternatives in the two models.

5. Size and Power.

The theoretical results derived in the previous sections suggest that the modified statistics are likely to be more robust to substantial serial correlation in the errors. To assess the local asymptotic properties of the statistics, we simulate the local asymptotic critical values for each model assuming σ^2 is estimated by s_{AR}^2. More precisely, for a given set of non-centrality parameters (δ in the first model, and ϕ in models two and three), integrals of the Weiner processes that appear in local asymptotic distributions in Theorems 2.2, 3.2, 4.2 are approximated by partial sums based upon 1000 draws of $N(0,1)$ errors. The distributions for MZ_t as implied by MZ_α and MSB are also tabulated. Percentage points are obtained via 10000 simulations of each local asymptotic distribution. These critical values are then used to construct size-adjusted power functions evaluated for values of the non-centrality parameter, c, between -15 and 5. For T=250, this corresponds to values of α between .94 and 1.02.

For the nearly integrated, nearly white noise model, we present results for values of δ between 1 and 20. This implies, for $T = 250$, that θ varies approximately between $-.94$ and -0.21. The results are presented in Figure 1. The top panel is the asymptotic size of the statistics when critical values from the standard asymptotic distributions are used.[6] For small values of δ, the use of standard critical values is associated with an under-rejection of the unit root hypothesis, but the approximation improves as δ increases. The bottom three panels are the size adjusted local asymptotic power for MZ_α, MSB, and MZ_t respectively. In general, local asymptotic power is higher against explosive alternatives (with $c > 0$) than against stationary alternatives ($c < 0$). For values of δ close to zero, MZ_α and MZ_t apparently has no local asymptotic power. This can be traced to the fact that the distributions of these statistics are independent of c when δ approaches zero. However, such is not the case with MSB since its power function is independent of the value of δ.

For the nearly twice integrated model, values of ϕ are chosen to be between -5 and -100, corresponding to values of ρ between .98 and .60 for T=250. We see from the top panel of Figure 2 that for extremely small values of ϕ, the use of standard asymptotic critical values will imply an over- rejection of the null hypothesis of one unit root. This issue was discussed earlier under Remark 6, where we note that the distribution of MZ_α will be shifted to the left of $T(\hat{\alpha}-1)$ by an amount that depends on ϕ. For small ϕ, the left shifting effect is large,

---

[6] For MZ_α and MZ_t, the lower five percentage points are -8.1 and -1.95 as given in Fuller (1976). For MSB, the case without a constant was not provided by Stock (1990). Approximating the distribution by $(\int_0^1 W(r)^2 dr)^{1/2}$ gives critical values for the lower and upper five percentage points of .23 and 1.28 respectively.

and explains why the tests over-reject a unit root when standard critical values are used. All three statistics have rather similar local asymptotic power; power is low when ϕ is small but increases monotonically as ϕ increases. This is to be expected since the data contain a unit root even under the local alternative that $c < 0$. Hence, the statistics are more powerful against explosive alternatives than against stationary ones.

Values of ϕ are also chosen to be between -1 and -50 for the nearly seasonally integrated model, implying values of ρ between -.99 and -.50 for T=100. The results are shown in Figure 3. As in the negative moving average case, the use of standard critical values implies tests that tend to under reject the unit root hypothesis. However, in this negative autoregressive case, MZ_α and MZ_t do not suffer from power loss at small values of ϕ. All three statistics have comparable power. As in the nearly twice integrated model, there is a strong dependence of power on the values of ϕ.

Our local asymptotic simulations suggest that except when values of the non-centrality parameters are extremely small, the local asymptotic power of the tests are good. Power increases with $|c|$. More importantly, the exact size of the tests based on standard critical values is usually within reasonable range of the nominal size unless the non-centrality parameters are very small. This is a useful result since it suggests that critical values appropriate for error processes with roots bounded away from one can also be used in these special cases.

It remains to evaluate the finite sample power of the statistics using the standard critical values and when σ^2 has to be estimated. To this end, we simulate the power of the tests using the autoregressive spectral density estimator for T=100 and T=500. We also report the power of the augmented Dickey-Fuller test, t_ρ. It is important to point out that the results reported in Table 3 are unadjusted for size distortions. The seemingly high power reported for t_ρ is inflated in the negative moving average case and should be interpreted with caution because of severe size distortions (see Table 1). With this in mind, we highlight several features of the results that are noteworthy.

First, MZ_α tends to be more powerful than MZ_t. This is consistent with a finding of Phillips and Ouliaris (1990) that the t statistic is generally less powerful than the normalized least squares estimator. Second, the power of MZ_α and MSB matches, and sometimes outperforms, that of t_ρ in finite samples except in the negative AR case, where the power discrepancies can be traced to the fact the the M tests are undersized.

Third, comparing the power of MZ_t with that of t_ρ after adjusting for size, the former has more power in models with positive serial correlation. For large negative AR errors, t_ρ is noticeably more powerful, while for large negative MA errors, the power of MZ_t is significantly higher. Although the limiting distribution of MZ_t and t_ρ are both given by

(1.3), the two statistics have rather different finite samples properties.

Fourth, comparing the results in (a) and (b) of Table 3, we see that the rate at which the power of the M tests increases is model dependent. The M tests have lowest power when there is a large negative AR root in the noise function and when the sample size is small, but power increases rapidly as T increases. Although the power of these tests is much higher in small samples in the negative MA case, it increases only slowly with T. These results are due to the fact that $\theta \to -1$ at rate $\sqrt{T}$ but that the autoregressive root approaches the boundary of one at rate T. The distinction between the null and the alternative hypotheses sharpens more rapidly in the AR models as T increases.

A result of particular importance concerns the power of the M tests when there are negative MA errors. This is the parameter space for which most tests have size problems. The M tests have much better size, and the power of MZ_α and MSB is higher than the size adjusted power of t_ρ. This finding that the much reduced size distortions are associated with good power suggests that the M tests, in particular, MZ_α and MSB, can be very useful in empirical work.

6. Conclusions.

When the root of the error process is close to the unit circle, many of the commonly used unit root tests are known to have distorted sizes. The problem arises because many of the relevant partial sums have non-standard rates of normalization when the root of the error process is close to the unit circle. However, simple modifications which have negligible effects in a standard asymptotic framework can lead to significantly more accurate exact sizes in the local asymptotic framework analyzed. The proviso is that the modifications have to be used in conjunction with a consistent estimate of the spectral density at frequency zero. While kernel based spectral density estimators do not satisfy this criteria and tend to aggravate the size problem, the autoregressive spectral density estimator formulated on the basis of the an augmented autoregression serves this purpose. When appropriately implemented, the modified statistics have rather robust properties and are useful tests for the presence of a unit root. The statistics will also be useful in cointegration analysis when serial correlation in the noise function is often encountered.

It is important to put into perspective the properties of the modified statistics vis-á-vis the general issue of distinguishing between processes with unit roots and stationary ones. As discussed in Campbell and Perron (1991) these two types of processes are observationally equivalent in the sense that for any stationary process, there will exist a unit root process which approximates it arbitrarily well and vice-versa. To be concrete about the implication of this result, consider the case of MA(1) errors with a negative coefficient θ. The near-observational equivalence implies than when using unit root tests with asymptotic critical values, there will exist values of θ in the range $(-1, x)$ for some $-1 < x < 0$, say, such that liberal size distortions will surface. The value of x will depend on the sample size and the test used, but it will always approach -1 as the sample size increases. In other words, the range over which size distortions occur will diminish, and this is true of any test. The problem, as shown in previous simulations including some presented in this paper, is that the rate at which the range shrinks can be very slow.

From a practical point of view, the problem is that for conventional tests (e.g, the Phillips-Perron or Dickey-Fuller tests) and sample sizes commonly encountered, this value of x where size distortions start to be important is too far away from -1 (e.g, somewhere around -.4 when T=100). This has been the cause of some concern in the literature because this range includes parameter values (e.g, between -.8 and -.4) which are of practical relevance[7] and for which we would rather not classify unit root processes with such moving average coefficients as stationary processes.

The class of modified statistics discussed in this paper can be viewed as tests with a much smaller range of size distortions (e.g, between -1 and -.9) for any given common sample size. This can be useful in practice because classifying unit root processes with values of θ in this range as stationary is likely to be of less concern. It is important to note that this improvement in size is achieved while retaining reasonable power.

The above justifications for using the modified statistics are valid insofar as the aim of testing for unit roots is to classify as precisely as possible whether a process is difference or trend stationary. There are, however, instances when the objective of the analysis is otherwise and using the modified statistics may not be appropriate. Suppose the aim of unit root tests is to decide which restrictions to make in a forecasting exercise. As reported in Campbell and Perron (1991), near-stationary unit processes are better forecast using stationary models, while near-integrated stationary processes are better forecast using integrated models. To the extent that imposing a false restriction may help reduce the mean squared error in this context, it is desirable to misclassify trend stationary processes as difference stationary and vice-versa, and one would rather use the conventional Dickey-Fuller or Phillips-Perron statistics to test for unit roots. Of course, in such cases, the "optimal" value of x is highly dependent on the overall objectives of the analysis of which unit root tests is just an important first step. On this issue, more work remains to be done.

[7]See Schwert (1987) for some empirical examples.

References

Andrews, D. (1991), Heteroskedastic and Autocorrelation Consistent Matrix Estimation, *Econometrica* **59**, 817-854.

Andrews, D. W. K. and Monahan, J. (1992), An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator, *Econometrica* **60**, 953–966.

Berk, K. N. (1974), Consistent Autoregressive Spectral Estimates, *The Annals of Statistics* **2**, 489–502.

Bhargava, A. (1986), On the Theory of Testing for Unit Roots in Observed Time Series, *Review of Economic Studies* **53**, 369–384.

Campbell, J. Y. and Perron, P. (1991), Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots, *NBER Macroeconomic Annual. MIT Press.* pp. 141–201.

Dejong, D. N., Nankervis, J. C., Savin, N. E. and Whiteman, C. H. (1992), Integration versus Trend Stationarity in Time Series, *Econometrica* **60**, 423–434.

Dickey, D. A. and Fuller, W. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* **74**, 427–431.

Fuller, R. (1976), *Introduction to Statistical Time Series*, John Wiley, New York.

Hyslop, D. (1991), The Performance of Phillips-Perron type Unit Root Tests using the Spectral Density Estimators, manuscript, Princeton University.

Kim, K. and Schmidt, P. (1990), Some Evidence on the Accuracy of Phillips-Perron Tests Using Alternative Estimates of Nuisance Parameters, *Economic Letters* **34**, 345-350.

Nabeya, S. and Perron, P. (1994), Local Asymptotic Distributions Related to the AR1 Model with Dependent Errors, *Journal of Econometrics* **62**, 229–264.

Newey, W. and West, K. (1987), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* **55**, 703–708.

Pantula, S. G. (1991), Asymptotic Distributions of Unit Root Tests When the Process is Nearly Stationary, *Journal of Business and Economic Statistics* **9**, 63–71.

Park, J. Y. and Phillips, P. C. B. (1988), Regressions with Integrated Processes, Part I, *Econometric Theory* **4**, 468–498.

Perron, P. (1992), The Adequacy of Asymptotic Approximations in the Near-Integrated Autoregressive Model with Dependent Errors, manuscript, Université de Montréal.

Perron, P. and Ng, S. (1993), The Exact Error in Estimating the Spectral Density at the Origin, manuscript, Université de Montréal.

Perron, P. and Ng, S. (1994), Properties of the Autoregressive Spectral Density Estimator at Frequency Zero in ARIMA Processes, manuscript in Preparation, Université de Montréal.

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

Phillips, P. C. B. and Ouliaris, S. (1990), Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica* **58**, 165–193.

Phillips, P. C. B. and Perron, P. (1988), Testing for a Unit Root in Time Series Regression, *Biometrika* **75**, 335–346.

Priestley, M. B. (1981), *Spectral Analysis and Time Series*, Vol. 1, Academic Press, New York.

Said, S. E. and Dickey, D. A. (1984), Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika* **71**, 599–607.

Sargan, J. D. and Bhargava, A. (1983), Testing for Residuals from Least Squares Regression Being Generated by Gaussian Random Walk, *Econometrica* **51**, 153–174.

Schwert, G. W. (1987), Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data, *Journal of Monetary Economics* **20**, 73–103.

Schwert, G. W. (1989), Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economic Statistics* **7**, 147–160.

Stock, J. H. (1990), A Class of Tests for Integration and Cointegration, manuscript, Harvard University.

White, J. (1958), The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case, *Annals of Mathematical Statistics* **29**, 1188–1197.

Appendix A: Mathematical Results

As a matter of notation, we shall let C_i denote (not necessarily the same) constants throughout this appendix. For each model, we start with a series of lemmas that consider the limit of the relevant sample moments.

2. Proofs of results in Section 2

The following Lemma is taken from Nabeya and Perron (1994):

Lemma 2.1 : *Let $\{y_t\}$ be generated by (2.1)-(2.3), and define $X_t = (1 + c/T)X_{t-1} + e_t$, $a_T = (1 - \delta/\sqrt{T})(1 - c/T)$, $b_T = 1 - (1 - c/T)(1 - \delta/\sqrt{T})$ with $a_T \to 1$ and $T^{1/2}b_T \to \delta$ as $T \to \infty$. We have*

a) $y_t = a_T e_t + b_T X_t$;

b) $\sum_1^T X_{t-1}e_t = O_p(T)$;

c) $T^{-2}\sum_1^T X_t^2 \Rightarrow \sigma_e^2 \int_0^1 J_c(r)^2 dr$.

Lemma 2.2 : *Let $\{y_t\}$ be generated according to (2.1) to (2.3) and $J_c(r)$ be defined as in (2.4). Let $e_\infty = \lim_{T\to\infty} e_T/\sigma_e$. Then as $T \to \infty$,*

a) $T^{-1}\sum_1^T y_{t-1}^2 \Rightarrow \sigma_e^2 + \sigma_e^2\delta^2 \int_0^1 J_c(r)^2 dr$;

b) $T^{-1}\sum_1^T y_{t-1}u_t \Rightarrow -\sigma_e^2$;

c) $y_T \Rightarrow e_\infty + \delta\sigma_e J_c(1)$;

d) $T^{-1}\sum_1^T u_t^2 \Rightarrow 2\ \sigma_e^2$.

The proofs of (a) and (b) are given in Nabeya and Perron (1994). Part (d) is obvious since $T^{-1}\sum_1^T e_t^2 \to \sigma_e^2$ and $\sum_1^T e_t e_{t-j}$ is $O_p(T^{1/2})$. Part (c) follows from the definition of y_t in Lemma 2.1.

To prove (2.5), write $\sum_{t=1}^T \hat{u}_t^2 = \sum_{t=1}^T (u_t - (\hat{\alpha} - \alpha)y_{t-1})^2$. Also note that $(\hat{\alpha} - \alpha) = \sum_1^T y_{t-1}u_t / \sum_1^T y_{t-1}^2$. Then $T^{-1}\sum_1^T \hat{u}_t^2 = T^{-1}\sum_1^T u_t^2 - (T^{-1}\sum_1^T y_{t-1}u_t)^2\ (T^{-1}\sum_1^T y_{t-1}^2)^{-1}$, and the result of (2.5) follows from Lemma 2.2. The following Lemma is a straightforward generalization of results in Fuller (1976) p. 374-376.

Lemma 2.3 : *Let $X_t = (1+c/T)X_{t-1} + e_t$, $e_t = \sum_{i=0}^\infty \omega_i v_{t-i}$, $v_t \sim iid(0, \sigma_v^2)$ and $\sum_{i=0}^\infty |w_i| < \infty$. Then*

a) $|E(e_t e_{t-k})| \le C_1 \lambda^{|k|}$ for some $0 < \lambda < 1$; b) $|E(X_t e_s)| \le C_2$;

c) $|E(X_t X_s)| \le TC_3$ for $t, s, \le T$; d) $Var(\sum_{t=1}^T X_t e_s) = O(T^2)$;

e) $E(\sum_{t=1}^T X_t e_s)^2 = O(T^2)$.

The following Lemma will also be useful.

Lemma 2.4 : *Let $\{y_t\}$ be generated by (2.1) to (2.3). Then for $i, j = 1, \dots, k$,*

a) $T^{-1}\sum_{t=k+1}^T y_{t-1}u_{t-i} \Rightarrow \begin{cases} \sigma_e^2 \text{ if } i = 1, \\ 0 \text{ otherwise;} \end{cases}$

b) $T^{-1}\sum_{t=k+1}^T y_{t-j}y_{t-i} \Rightarrow \begin{cases} \sigma_e^2(1 + \delta^2 \int_0^1 J_c(r)^2 dr) \text{ if } i = j, \\ \sigma_e^2\delta^2 \int_0^1 J_c(r)^2 dr \text{ if } i \neq j; \end{cases}$

c) $T^{-1}\sum_{t=k+1}^T u_{t-i}u_{t-j} \Rightarrow \begin{cases} 2\sigma_e^2 \text{ if } i = j, \\ -\sigma_e^2 \text{ if } |i - j| = 1, \\ 0 \text{ otherwise.} \end{cases}$

Proof. To prove part (a), note that

$$T^{-1}\sum_{t=k+1}^{T} y_{t-1}u_{t-i} = T^{-1}\sum_{t=k+1}^{T}(a_T e_{t-1} + b_T X_{t-1})(e_{t-i} + \theta_T e_{t-i-1})$$
$$= a_T T^{-1}\sum_{t=k+1}^{T} e_{t-1}e_{t-i} + \theta_T a_T T^{-1}\sum_{t=k+1}^{T} e_{t-1}e_{t-i-1}$$
$$+T^{1/2}b_T T^{-3/2}\sum_{t=k+1}^{T} X_{t-1}e_{t-i} + T^{1/2}b_T\theta_T T^{-3/2}\sum_{t=k+1}^{T} X_{t-1}e_{t-i-1}.$$

Since $a_T \to 1$ and $T^{-1}\sum_{t=k+1}^{T} e_{t-1}e_{t-i} \to \sigma_e^2$ if $i = 1$ and 0 otherwise (provided $k/T \to 0$), it remains to show that the last three terms converge to zero. The result is straightforward for the second since $\theta_T \to -1$, $a_T \to 1$, and $T^{-1}\sum_{t=k+1}^{T} e_{t-1}e_{t-i-1} \to 0$ for $i = 1,\ldots,k$. Now for the third term,

$$E[(T^{-3/2}\sum_{t=k+1}^{T} X_{t-1}e_{t-i})^2] = T^{-3}Var(\sum_{t=k+1}^{T} X_{t-1}e_{t-i}) + T^{-3}[E(\sum_{t=k+1}^{T} X_{t-1}e_{t-i})]^2$$

and is $O(T^{-1})$ if $k/T \to 0$ using Lemma 2.3 (b,d). The fourth term converges to zero using similar arguments. To prove part (b), note that by Lemma 2.1,

$$T^{-1}\sum_{t=k+1}^{T} y_{t-i}y_{t-j} = a_T^2 T^{-1}\sum_{t=k+1}^{T} e_{t-j}e_{t-i} + a_T T^{1/2}b_T T^{-3/2}\sum_{t=k+1}^{T} e_{t-j}X_{t-i}$$
$$+ a_T T^{1/2}b_T T^{-3/2}\sum_{t=k+1}^{T} e_{t-i}X_{t-j} + Tb_T^2 T^{-2}\sum_{t=k+1}^{T} X_{t-i}X_{t-j}.$$

For the first term, we have $T^{-1}\sum_{t=k+1}^{T} e_{t-i}e_{t-j} \to \sigma_e^2$ if $i = j$ and 0 otherwise provided $k/T \to 0$. By Lemma 2.3, $T^{-3}E[(\sum_{t=k+1}^{T} e_{t-j}X_{t-i})^2] = T^{-3}O[(T-k)^2] = O(T^{-1})$ provided $k/T \to 0$. Since $a_T \to 1$ and $T^{1/2}b_T \to \delta$, the second and third terms converge to zero. For the fourth term, define the *cadlag* process on $D[0,1]$, with $X_T(r) = X_t$, $\frac{t-1}{T} \le r < \frac{t}{T}$. Note that $X_T(r - \frac{i}{T}) = X_{t-j}$. We have

$$T^{-2}\sum_{t=k+1}^{T} X_{t-i}X_{t-j} = \int_{k/T}^{1} T^{-1/2}X_T(r - j/T)T^{-1/2}X_T(r - i/T)dr \Rightarrow \sigma_e^2\int_0^1 J_c(r)^2 dr$$

since $k/T \to 0$ (hence $j/T, i/T \to 0$) and $T^{-1/2}X_T(r) \Rightarrow \sigma_e J_c(r)$. Collecting the terms, we have the results stated. To prove part (c), note that

$$T^{-1}\sum_{t=k+1}^{T} u_{t-i}u_{t-j} = T^{-1}\sum_{t=k+1}^{T} e_{t-i}e_{t-j} + \theta_T T^{-1}\sum_{t=k+1}^{T} e_{t-i}e_{t-j-1}$$
$$+ \theta_T T^{-1}\sum_{t=k+1}^{T} e_{t-i-1}e_{t-j} + \theta_T^2 T^{-1}\sum_{t=k+1}^{T} e_{t-i-1}e_{t-j-1}.$$

The first term converges to σ_e^2 if $i = j$ and 0 otherwise. Since $\theta_T \to -1$, the second term converges to $-\sigma_e^2$ if $i = j+1$ and 0 otherwise. Similarly, the third term converges to $-\sigma_e^2$ if $i = j-1$ and 0 otherwise. The last term converges to σ_e^2 if $i = j$. Part (c) follows upon collecting terms.

Proof of Theorem 2.1
It is convenient to write $T^{-1}Z_\alpha = (\hat{\alpha} - 1) - (1/2)(s^2 - s_u^2)/\ T^{-1}\sum_1^T y_{t-1}^2$. Given (2.2) and (2.3), $s^2 = \sigma_T^2 = \sigma_e^2(1+\theta_T)^2 = \sigma_e^2\delta^2/T$, the correction factor is:

$$(s^2 - s_u^2)/(T^{-1}\sum_1^T y_{t-1}^2) = (\delta^2\sigma_e^2/T)/(T^{-1}\sum_1^T y_{t-1}^2) - s_u^2/(T^{-1}\sum_1^T y_{t-1}^2). \qquad \text{(A2.1)}$$

It follows that,

$$T^{-1}Z_\alpha = (\hat{\alpha} - 1) + s_u^2/(2T^{-1}\sum_1^T y_{t-1}^2) + o_p(1)$$
$$\Rightarrow -1/(1 + \delta^2 \int_0^1 J_c(r)^2 dr) + (1 + 2\delta^2 \int_0^1 J_c(r)^2 dr)/(2(1 + \delta^2 \int_0^1 J_c(r)^2 dr)^2)$$

Part (a) follows upon simplification. Part (c) follows from (a) and (b) of Lemma 2.2 and the definition of s^2. To prove part (d), note that given the definition of σ^2, $MSB = [T^{-1}\sum_1^T y_{t-1}^2/(\delta^2\sigma_\epsilon^2)]^{1/2}$, and the result follows from (a) of Lemma (2.2). Given that $Z_t = MSB \cdot Z_\alpha$, part (b) follows from parts (a) and (d).

Proof of Theorem 2.2

Lemma 2.5 : *Let $\{y_t\}$ be generated by (2.1) to (2.3) and s_{AR}^2 be obtained by applying OLS to (1.11). Then $s_{AR}^2 \to 0$ provided $k/T \to 0$ and $k \to \infty$ as $T \to \infty$.*

The proof of the Lemma is given in Perron and Ng (1994). Given that s_{AR}^2 and σ_T^2 have the same asymptotic limit, the results for Theorem 2.2 are the same as those for Theorem 2.1.

Proof of Theorem 2.3

We start with a Lemma concerning the limiting distribution of $M^{-1}s_{WA}^2$.

Lemma 2.6 : *Let $\{y_t\}$ be generated by (2.1) to (2.3) and s_{WA}^2 be defined by (1.13), where $\hat{u}_t$ are OLS residuals obtained from (1.1). Let ψ be defined by (1.14) and $J_c(r)$ by (2.4). Then*

$$M^{-1}s_{WA}^2 \Rightarrow \left(2\sigma_\epsilon^2\delta^2\psi \int_0^1 J_c(r)^2 dr\right)\left(1 + \delta^2 \int_0^1 J_c(r)^2 dr\right)^{-2}. \quad \text{(A2.2)}$$

Proof. We first note that since $s_u^2 = O_p(1)$, the limit of $M^{-1}(s_{WA}^2 - s_u^2)/2$ is the same as the limit of $M^{-1}s_{WA}^2/2$. We have

$$\begin{aligned} M^{-1}(s_{WA}^2 - s_u^2)/2 &= M^{-1}T^{-1}\sum_{k=1}^M w_k \sum_{t=k+1}^T \hat{u}_t\hat{u}_{t-k} \\ &= M^{-1}T^{-1}\sum_{k=1}^M w_k \sum_{t=k+1}^T \{u_t u_{t-k} - (\hat{\alpha} - \alpha)y_{t-1}u_{t-k} \\ &\quad - (\hat{\alpha} - \alpha)y_{t-k-1}u_t + (\hat{\alpha} - \alpha)^2 y_{t-1}y_{t-k-1}\} \end{aligned} \quad \text{(A2.3)}$$

The following Lemma presents the limit for each term:

Lemma 2.7 : *Let $M \to \infty$ and $M/T \to 0$ as $T \to \infty$. Then*
a) $M^{-1}T^{-1}\sum_{k=1}^M w_k \sum_{t=k+1}^T u_t u_{t-k} \to 0$;
b) $M^{-1}T^{-1}\sum_{k=1}^M w_k \sum_{t=k+1}^T y_{t-1}u_{t-k} \to 0$;
c) $M^{-1}T^{-1}\sum_{k=1}^M w_k \sum_{t=k+1}^T y_{t-k-1}u_t \to 0$;
d) $M^{-1}T^{-1}\sum_{k=1}^M w_k \sum_{t=k+1}^T y_{t-1}y_{t-k-1} \Rightarrow \sigma_\epsilon^2\psi\delta^2 \int_0^1 J_c(r)^2 dr$.

Proof. a) Note that

$$T^{-1}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} u_t u_{t-k} = T^{-1} w_1 \sum_{t=2}^{T} u_t u_{t-1} + T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} u_t u_{t-k}. \quad (A2.4)$$

Consider the first term in (A2.4). Using the fact that $u_t = e_t - (1-\delta/\sqrt{T})e_{t-1}$,

$$w_1 T^{-1}\sum_{t=2}^{T} u_t u_{t-1} = w_1 T^{-1}\sum_{t=2}^{T} e_t e_{t-1} - w_1(1-\delta/\sqrt{T})T^{-1}\sum_{t=2}^{T} e_{t-1}^2$$
$$- w_1(1-\delta/\sqrt{T})T^{-1}\sum_{t=2}^{T} e_t e_{t-2} + w_1(1-\delta/\sqrt{T})^2 T^{-1}\sum_{t=2}^{T} e_{t-1}e_{t-2}.$$

Since $\{e_t\}$ is *i.i.d.*, all terms vanish except the second, and the expression converges to $-w_1\sigma_e^2$ and to zero upon normalization by M^{-1}. To show that the second term in (A2.4) vanishes, note that

$$T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} u_t u_{t-k}$$
$$= T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} e_t e_{t-k} - T^{-1}(1-\delta/\sqrt{T})\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} e_{t-1}e_{t-k}$$
$$- T^{-1}(1-\delta/\sqrt{T})\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} e_t e_{t-k-1} + T^{-1}(1-\delta/\sqrt{T})^2\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} e_{t-1}e_{t-k-1}. \quad (A2.5)$$

Now consider the first element. Define $\hat{s}_e^2 = T^{-1}\sum_{t=1}^{T} e_t^2 + 2T^{-1}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} e_t e_{t-k}$. We have

$$T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} e_t e_{t-k} = (\hat{s}_e^2 - T^{-1}\sum_{t=1}^{T} e_t^2)/2 - T^{-1} w_1 \sum_{t=2}^{T} e_t e_{t-1}.$$

This expression vanishes since $\hat{s}_e^2 \to \sigma_e^2$ (provided $M/T \to 0$ as $T\to\infty$), $T^{-1}\sum_{t=1}^{T} e_t^2 \to \sigma_e^2$, and $T^{-1}\sum_{t=2}^{T} e_t e_{t-1} \to 0$ given that $\{e_t\} \sim$ *i.i.d.*. Similar arguments apply to show that the remaining terms of (A2.5) also converge to zero. To prove part (b), note that

$$M^{-1}T^{-1}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}u_{t-k}$$
$$= M^{-1}a_T w_1 T^{-1}\sum_{t=2}^{T} e_{t-1}^2 - M^{-1}w_1 a_T(1-\delta/\sqrt{T})T^{-1}\sum_{t=2}^{T} e_{t-1}e_{t-2}$$
$$+ w_1 M^{-1}T^{1/2}b_T T^{-3/2}(\sum_{t=2}^{T} X_{t-1}e_{t-1} - (1-\delta/\sqrt{T})\sum_{t=2}^{T} X_{t-1}e_{t-2})$$
$$+ M^{-1}T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T}(a_T e_{t-1} + b_T X_{t-1})(e_{t-k} - (1-\delta/\sqrt{T})e_{t-k-1}).$$

Since $a_T \to 1$ and $T^{-1}\sum_{t=2}^{T} e_{t-1}^2 \to \sigma_e^2$, $a_T w_1 T^{-1}\sum_{t=2}^{T} e_{t-1}^2$ converges to $w_1\sigma_e^2$ and the first term vanishes upon normalization by M^{-1}. The next three terms vanish using $T^{1/2}b_T \to \delta$, $T^{-1}\sum e_{t-1}e_{t-2} \to 0$, and Lemmas 2.1 and 2.2. It remains to show that the last term vanishes. We have

$$M^{-1}T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T}(a_T e_{t-1} + b_T X_{t-1})(e_{t-k} - (1-\delta/\sqrt{T})e_{t-k-1})$$
$$= M^{-1}\Big(a_T T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} e_{t-1}e_{t-k} - a_T(1-\delta/\sqrt{T})T^{-1}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} e_{t-1}e_{t-k-1}$$
$$+ T^{1/2}b_T T^{-3/2}(\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} X_{t-1}e_{t-k} - (1-\delta/\sqrt{T})\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} X_{t-1}e_{t-k-1})\Big). \quad (A2.6)$$

The first two terms of (A2.6) converge to zero using arguments similar to those in part (a). Consider the third term (the behavior of the fourth is similar). Since $T^{1/2}b_T \to \delta$, we consider

$$M^{-1}T^{-3/2}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} X_{t-1}e_{t-k}$$
$$= M^{-1}T^{-3/2}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} \alpha^k X_{t-k-1}e_{t-k} + M^{-1}T^{-3/2}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T} z_{t,k}, \tag{A2.7}$$

where $z_{t,k} = e_{t-k}\sum_{i=1}^{k}\alpha^{i-1}e_{t-i}$. Since $X_{t-k-1}e_{t-k} = (1/2\alpha)(X_{t-k}^2 - \alpha^2 X_{t-k-1}^2 - e_{t-k}^2)$ and $X_0 = 0$ by assumption, the first term of (A2.7) simplifies to

$$(1/2\alpha)M^{-1}T^{-3/2}\sum_{k=2}^{M}\alpha^k w_k (X_{T-k}^2 - (\alpha^2 - 1)\sum_{t=k+1}^{T} X_{t-k-1}^2 - \sum_{t=k+1}^{T} e_{t-k}^2).$$

It is easy to see that the third term vanishes. For the first term, define the process on $D[0,1]$ as $X_T(s) = X_{[Ts]} = X_{j-1}$, $(j-1)/T \le s < j/T$ and $X_T(1) = X_T$. Now

$$\tfrac{1}{2\alpha}M^{-1}T^{-3/2}\sum_{k=2}^{M}\alpha^k w_k X_{T-k}^2 = \tfrac{1}{2\alpha}T^{-1/2}\int_0^1 \exp(\tfrac{c[Ms]}{T})w(\tfrac{[Ms]}{M})T^{-1}X_T^2(\tfrac{[T-s-1]}{T})\,ds + o_p(1),$$

which converges to 0 since $T^{-1/2}X_T(s) \Rightarrow \sigma_e J_c(s)$. Similar arguments using the facts that $T(\alpha^2-1) \to 2c$ and $T^{-2}\sum_{t=k+1}^{T} X_{t-k-1}^2 \Rightarrow \sigma_e^2\int_0^1 J_c(r)^2 dr$ show that the second term vanishes. Thus, the first term of (A2.7) vanishes. It remains to show that the second term of (A2.7) converge to zero. Since $E[z_{t,k}] = \alpha^{k-1}\sigma_e^2$, we can write the expression as

$$M^{-1}T^{-3/2}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T}(z_{t,k} - E[z_{t,k}]) + M^{-1}T^{-3/2}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T}(\alpha^{k-1}\sigma_e^2)$$
$$= M^{-1}T^{-3/2}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T}(z_{t,k} - E[z_{t,k}]) + o(1).$$

It can be shown that $Var(\sum_{t=k+1}^{T} z_{t,k}) = 2\sigma_e^4\alpha^{2(k-1)}(T-k) + (T-k)\sigma_e^4(\alpha^{2k}-1)/(\alpha^2-1)$. Hence using an argument as in Newey and West (1987),

$$P\{|M^{-1}T^{-3/2}\sum_{k=2}^{M} w_k \sum_{t=k+1}^{T}(z_{t,k} - E[z_{t,k}])| > \epsilon\}$$
$$\le \sum_{k=2}^{M} P\{|M^{-1}T^{-3/2}\sum_{t=k+1}^{T}(z_{t,k} - E[z_{t,k}])| > \epsilon/CM\} \text{ since } w_k \le C$$
$$\le \sum_{k=2}^{M}[M^{-2}T^{-3}(T-k)(C^2M^2/\epsilon^2)\sigma_e^4\{2\alpha^{2(k-1)} + \tfrac{\alpha^{2k}-1}{\alpha^2-1}\}] \text{ by Cauchy Schwartz inequality}$$
$$\le \frac{M}{T}\frac{2\sigma_e^4}{T}\frac{(T/M)(\alpha^{2M}-1)}{T(1-\alpha^2)}C^2/\epsilon^2 + \frac{M}{T}\frac{\sigma_e^4}{T(\alpha^2-1)}\left(\frac{(T/M)}{T}\frac{(\alpha^{2M}-1)}{(\alpha^2-1)} - 1\right)C^2/\epsilon^2 \to 0$$

since $T(\alpha^2-1) \to 2c$ and $(T/M)(\alpha^{2M}-1) \to 2c$ as $M/T \to 0$. This completes the proof for part (b). The proof of part (c) is analogous. To prove part (d), note that

$$M^{-1}T^{-1}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}y_{t-k-1} = M^{-1}T^{-1}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} a_T^2 e_{t-1}e_{t-k-1}$$
$$+ M^{-1}a_T T^{1/2}b_T T^{-3/2}(\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} X_{t-1}e_{t-k-1} + \sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} e_{t-1}X_{t-k-1})$$
$$+ M^{-1}Tb_T^2 T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} X_{t-1}X_{t-k-1}.$$

The first three terms converge to zero using arguments similar to those in parts (a) and (b). We therefore concentrate on the fourth term, which we write as:

$$M^{-1}Tb_T^2T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T}\alpha^k X_{t-k-1}^2 + M^{-1}Tb_T^2T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} X_{t-k-1}(\sum_{i=0}^{k-1}\alpha^i e_{t-i-1}).$$

The second term converges to zero by Lemma 2.1. Consider

$$T b_T^2 M^{-1} T^{-2} \sum_{k=1}^{M} w_k \sum_{t=1}^{T-k-1} \alpha^k X_{t-1}^2$$
$$= T b_T^2 \sum_{k=1}^{M} \int_{\frac{k-1}{M}}^{\frac{k}{M}} w(\tfrac{\lfloor Ms \rfloor}{M}) \sum_{t=1}^{T-k-1} \int_{\frac{t-1}{T}}^{\frac{t}{T}} \exp(\tfrac{ck}{T}) T^{-1} X_T^2(r) dr ds$$
$$= T b_T^2 \int_0^1 w(\tfrac{\lfloor Ms \rfloor}{M}) \int_0^{1-s\frac{M}{T}} \exp(cs\tfrac{M}{T}) T^{-1} X_T^2(r) dr\, ds$$
$$\Rightarrow \delta^2 \int_0^1 w(s)\sigma_e^2 \int_0^1 J_c(r)^2 dr\, ds = \delta^2 \psi \sigma_e^2 \int_0^1 J_c(r)^2 dr,$$

provided $\frac{M}{T} \to 0$. Using (A2.3), Lemma 2.7 and the result that $(\hat{\alpha}-\alpha) \Rightarrow -(1+\delta^2 \int_0^1 J_c(r)^2 dr)^{-1}$ gives the result stated in Lemma 2.6. Theorem 2.3 follows from Lemmas 2.1, 2.2 and 2.6.

3. Proofs of results in Section 3.

Parts (a) to (d) of the following Lemma is proved in Nabeya and Perron (1994), and part (e) is given in Perron and Ng (1994).

Lemma 3.1 : *Let $\{y_t\}$ be a process given by (3.1) and (3.2) with $J_\phi(r)$ and $Q_c(J_\phi(r))$ as defined in (3.3). As $T \to \infty$:*
a) $T^{-3/2} y_T \Rightarrow \sigma_e\, Q_c(J_\phi(1))$;
b) $T^{-4} \sum_1^T y_t^2 \Rightarrow \sigma_e^2 \int_0^1 Q_c(J_\phi(r))^2 dr$;
c) $T^{-3} \sum_1^T y_{t-1} u_t \Rightarrow (\sigma_e^2/2)\{Q_c(J_\phi(1))^2 - 2c \int_0^1 Q_c(J_\phi(r))^2 dr\}$;
d) $T^{-2} \sum_1^T u_t^2 \Rightarrow \sigma_e^2 \int_0^1 J_\phi(r)^2 dr$;
e) $T^{-1} \sum_{t=k+1}^T \Delta y_{t-1} e_t \Rightarrow \sigma_e^2 (c \int_0^1 Q_c(J_\phi(r)) dW(r) + \int_0^1 J_\phi(r) dW(r))$ if $k/T \to 0$ as $T \to \infty$.

The expressions (3.4) and (3.6) follow from Lemma 3.1. Theorem 3.1 uses this Lemma and the definition $T^{-2}s^2 = T^{-2}\sigma_T^2 = \sigma_e^2/\phi^2$.

Proof of Theorem 3.2:

Lemma 3.2 : *Let $\{y_t\}$ be a process given by (3.1) and (3.2) with $J_\phi(r)$ and $Q_c(J_\phi(r))$ as defined in (3.3). Let s_{AR}^2 be obtained by applying OLS to (1.11) with $k \to \infty$ and $k = o(T^{1/3})$ and let $T(\hat{b}(1) - b(1)) \to \eta$ with the random variable η defined as in Perron and Ng (1994). Then $T^{-2} s_{AR}^2 \to \sigma_e^2/(c + \phi + \eta)^2$.*

The proof of the Lemma is given in Perron and Ng (1994), and the results of Theorem 3.2 follow arguments analogous to those used in Theorem 3.1, with $(c + \phi + \eta)$ replacing ϕ.

Proof of Theorem 3.3:
Since $T^{-1}s_u^2 \Rightarrow \lambda\sigma_e^2$, (see (3.6)), the limit of $M^{-1}T^{-1}s_{WA}^2/2$ is the same as the limit of $M^{-1}T^{-1}(s_{WA}^2 - s_u^2)/2$, which we write as

$$M^{-1}T^{-1}\frac{s_{WA}^2 - s_u^2}{2} = M^{-1}T^{-2} \sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} \hat{u}_t \hat{u}_{t-k}$$
$$= M^{-1}T^{-2} \sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} u_t u_{t-k} \qquad \text{(A3.1)}$$
$$- \; T(\hat{\alpha} - \alpha) M^{-1} T^{-3} (\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1} u_{t-k} + \sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-k-1} u_t)$$
$$+ \; T^2(\hat{\alpha} - \alpha)^2 M^{-1} T^{-4} \sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1} y_{t-k-1}.$$

We note from (3.3) that $T(\hat{\alpha}-\alpha)=O_p(1)$. The next Lemma characterizes the limit of each term in (A3.1).

Lemma 3.3 : *Let $\{y_t\}$ be generated by (3.1) to (3.3) and let $\frac{M}{T}\to 0$ and $M\to\infty$ as $T\to\infty$. Let $\psi=\int_0^1 w(s)ds$, $Q_c(J_\phi(r))$ and $J_\phi(v)$ as defined in (3.3). Then*
a) $M^{-1}T^{-2}\sum_{k=1}^M w_k\sum_{t=k+1}^T u_t u_{t-k}\Rightarrow \psi\sigma_e^2\int_0^1 J_\phi(r)^2dr$;
b) $M^{-1}T^{-3}\sum_{k=1}^M w_k\sum_{t=k+1}^T y_{t-1}u_{t-k}\Rightarrow \frac{\psi}{2}\sigma_e^2\left(Q_c(J_\phi(1))^2-2c\int_0^1 Q_c(J_\phi(r))^2dr\right)$;
c) $M^{-1}T^{-3}\sum_{k=1}^M w_k\sum_{t=k+1}^T y_{t-k-1}u_t\Rightarrow \frac{\psi}{2}\sigma_e^2\left(Q_c(J_\phi(1))^2-2c\int_0^1 Q_c(J_\phi(r))^2dr\right)$;
d) $M^{-1}T^{-4}\sum_{k=1}^M w_k\sum_{t=k+1}^T y_{t-1}y_{t-k-1}\Rightarrow \psi\sigma_e^2\int_0^1 Q_c(J_\phi(r))^2dr$.

Proof. To prove part (a), note that since $u_t=\rho u_{t-1}+e_t$, where $\rho=(1+\phi/T)$,

$$M^{-1}T^{-2}\sum_{k=1}^M w_k\sum_{t=k+1}^T u_t u_{t-k}=M^{-1}T^{-2}\sum_{k=1}^M w_k\sum_{t=k+1}^T \rho^k u_{t-k}^2 + M^{-1}T^{-2}\sum_{k=1}^M w_k\sum_{t=k+1}^T z_{k,t} \qquad (A3.2)$$

where $z_{k,t}=u_{t-k}(\sum_{i=0}^{k-1}\rho^i e_{t-i})$. Consider the first term and let $U_T(s)=u_{[Ts]}=u_{j-1}$, $\frac{j-1}{T}\le s<\frac{j}{T}$, and note that $T^{-1/2}U_T(s)\Rightarrow\sigma_e J_\phi(s)$. Then

$$\begin{aligned}&M^{-1}T^{-2}\sum_{k=1}^M w_k\sum_{t=k+1}^T \rho^k u_{t-k}^2\\
&=\sum_{k=1}^M\int_{\frac{k-1}{M}}^{\frac{k}{M}} w(\tfrac{[sM]}{M})\sum_{t=k+1}^T\int_{\frac{t-1}{T}}^{\frac{t}{T}}\exp(\phi s\tfrac{M}{T})T^{-1}U_T(r)^2dr\,ds\\
&=\int_0^1 w(\tfrac{[sM]}{M})\int_{\frac{sM}{T}}^1\exp(\phi s\tfrac{M}{T})T^{-1}U_T(r)^2dr\,ds\\
&\Rightarrow \sigma_e^2\int_0^1 w(s)\int_0^1 J_\phi(r)^2dr\,ds\equiv\psi\sigma_e^2\int_0^1 J_\phi(r)^2dr,\end{aligned}$$

provided $M/T\to 0$ and using Lemma 3.1. The second term in (A3.2) converges to zero since $\sum_{t=k+1}^T e_t e_{t-k}$ is $O_p(\sqrt{T})$ and $M/T\to 0$ by assumption. The proofs to parts (b), (c), and (d) follow analogously using the results of Lemma 3.1. Combining (3.3), Lemma 3.3 and (A3.1), we have:

Lemma 3.4 : *Let s_{WA}^2 be defined by (1.13) with $M\to\infty$ and $M/T\to 0$ as $T\to\infty$. Let $\psi=\int_0^1 w(s)ds$ and λ be defined by (3.5). Then*

$$(MT)^{-1}s_{WA}^2\Rightarrow 2\sigma_e^2\psi\lambda.$$

The results of Theorem 3.3 follow directly from Lemmas 3.1 and 3.4.

4. Proofs of Results in Section 4

The following Lemma is proved in Nabeya and Perron (1994).

Lemma 4.1 : *Let $\{y_t\}$ be generated by (4.1) and (4.2). Using the definitions following (4.3), we have, as $T \to \infty$:*
a. $T^{-2}\sum_{t=1}^{T} y_{t-1}^2 \Rightarrow (\sigma_e^2/8)\int_0^1 (C(r)^2 + B(r)^2)dr$;
b. $T^{-2}\sum_{t=1}^{T} y_{t-1}u_t \Rightarrow -(\sigma_e^2/4)\int_0^1 B(r)^2 dr$;
c. $T^{-1}y_T^2 \Rightarrow (\sigma_e^2/8)A(1)^2$;
d. $T^{-2}\sum_{t=1}^{T} u_t^2 \Rightarrow (\sigma_e^2/2)\int_0^1 B(r)^2 dr$.

To prove (4.5), we simply apply Lemma 4.1 to the definition of $T^{-1}s_u^2$. The results in Theorem (4.1) use the fact that $s^2 = \sigma_T^2 = \sigma_e^2/(2+\phi/T)^2 \to \sigma_e^2/4$. Since s_u^2 is $O_p(T)$ and s^2 is $O_p(1)$, $(s^2 - s_u^2)$ is dominated by s_u^2. Part (a) follows from (4.3), (4.5), and Lemma 4.1 (a). Parts (c) and (d) follow directly from Lemma 4.1 and the definition of s^2. Part (b) is a direct implication of (1.9) and can be proved directly using the relevant parts of the Lemma.

Proof of Theorem 4.2:

Lemma 4.2 : *Let $\{y_t\}$ be generated by (4.1) and (4.2). Let s_{AR}^2 be obtained by applying OLS to (1.11) with $k = o(T^{1/3})$. Then $s_{AR}^2 \to \sigma_e^2/4$.*

The proof of the Lemma is given in Perron and Ng (1994), and the results of the Theorem are obvious in view of Theorem 4.1.

Proof of Theorem 4.3:
We begin by considering the limit of $M^{-1}T^{-1}(s_{WA}^2 - s_u^2)/2$, which we write as

$$M^{-1}T^{-1}(s_{WA}^2 - s_u^2)/2 = M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} u_t u_{t-k}$$
$$- (\hat{\alpha}-\alpha)M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}u_{t-k} - (\hat{\alpha}-\alpha)M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-k-1}u_t$$
$$+ (\hat{\alpha}-\alpha)^2 M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}y_{t-k-1}. \qquad \text{(A4.1)}$$

The next Lemma characterizes the limit of each term in (A4.1).

Lemma 4.3 : *Let $\{y_t\}$ be generated by (4.1) and (4.2) and let $\frac{M}{T} \to 0$ and $M \to \infty$ as $T \to \infty$ with $\psi = \int_0^1 w(x)dx$, we have*
a) $M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} u_t u_{t-k} \Rightarrow 0$;
b) $M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}u_{t-k} \Rightarrow 0$;
c) $M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-k-1}u_t \Rightarrow 0$;
d) $M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}y_{t-k-1} \Rightarrow \psi\frac{\sigma_e^2}{8}\int_0^1 (C(r)^2 - B(r)^2)dr$.

Proof. To prove part (a), note that since $u_t = \rho u_{t-1} + e_t$, where $\rho = -(1+\phi/T)$,

$$M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} u_t u_{t-k} = M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} \rho^k u_{t-k}^2 \qquad \text{(A4.2)}$$
$$+ M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} z_{k,t}$$

where $z_{k,t} = u_{t-k}(\sum_{i=0}^{k-1} \rho^i e_{t-i})$. It is straightforward to show that the second term converges to zero. Consider the first term. We have,

$$M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} \rho^k u_{t-k}^2 = M^{-1}T^{-2}\sum_{k=1}^{M} \rho^k w_k \sum_{t=k+1}^{T} u_{t-k}^2$$

$$= M^{-1}T^{-2}\left(\sum_{j=1}^{M/2} \rho^{2j} w_{2j} \sum_{t=2j+1}^{T} u_{t-2j}^2 - \sum_{j=1}^{M/2}(-\rho)^{2j-1} w_{2j-1} \sum_{t=2j}^{T} u_{t-2j+1}^2\right)$$

$$= M^{-1}T^{-2}\left(\sum_{j=1}^{M/2} \exp(\tfrac{2j\phi}{T}) w_{2j} \sum_{t=2j+1}^{T} u_{t-2j}^2 - \sum_{j=1}^{M/2} \exp(\tfrac{(2j-1)\phi}{T}) w_{2j-1} \sum_{t=2j}^{T} u_{t-2j+1}^2\right) + o_p(1).$$

Define $w_M^*(r) = w_{2j}$, $\frac{2(j-1)}{M} \leq r < \frac{2j}{M}$ and let $U_T(s) = u_{[Ts]} = u_{j-1}$ for $\frac{j-1}{T} \leq s < \frac{j}{T}$. Now rewrite the first term (for sum over even terms) as:

$$\tfrac{1}{2}(\tfrac{M}{2})^{-1}\sum_{j=1}^{M/2} \exp(\tfrac{2j\phi}{T}) w_{2j}\left(T^{-2}\sum_{t=2j+1}^{T} u_{t-2j}^2\right)$$

$$= \tfrac{1}{2}\sum_{j=1}^{M/2}\int_{2(j-1)/M}^{2j/M} \exp(\tfrac{2j\phi}{T}) w_{2j}\, ds \sum_{t=2j+1}^{T}\int_{t-1/T}^{t/T} T^{-1}U_T(r)^2 dr$$

$$= \tfrac{1}{2}\int_0^1 \exp(\tfrac{2[sM/2]\phi}{T}) w_M^*(s)\, ds \int_{\frac{sM/2}{T}}^{1} T^{-1}U_T(r)^2 dr$$

$$\Rightarrow \tfrac{1}{2}\int_0^1 w(s)ds\, \tfrac{\sigma_\epsilon^2}{2}\int_0^1 B(r)^2 dr = \tfrac{\sigma_\epsilon^2}{4}\psi \int_0^1 B(r)^2 dr.$$

The second term (for sum over odd terms) can be shown to converge to the same functional, giving the result as stated in (a). Similar arguments can be used to show that the odd and even terms in parts (b) and (c) cancel. To show part (d),

$$M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}y_{t-k-1} = M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} \alpha^k y_{t-k-1}^2$$
$$+ \; M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-k-1}(\sum_{i=1}^{k}\alpha^{i-1}u_{t-i}). \qquad (A4.3)$$

Using Lemma 4.1 and provided $M/T \to 0$ as $T \to \infty$, we have, for the first term of (A4.3):

$$M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} \alpha^k y_{t-k-1}^2 \Rightarrow \psi\tfrac{\sigma_\epsilon^2}{8}\int_0^1 (C(r)^2 + B(r)^2)dr. \qquad (A4.4)$$

The second term of (A4.3) can be written as

$$M^{-1}T^{-2}\sum_{k=1}^{M} w_k\alpha^{k-1}\sum_{t=k+1}^{T} y_{t-k-1}u_{t-k} + M^{-1}T^{-2}\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-k-1}(\sum_{i=1}^{k-1}\alpha^{i-1}u_{t-i}).$$

It is straightforward to show that the second term vanishes. For the first term,

$$M^{-1}T^{-2}\sum_{k=1}^{M} w_k\alpha^{k-1}\sum_{t=k+1}^{T} y_{t-k-1}u_{t-k} \Rightarrow -\tfrac{\psi\sigma_\epsilon^2}{4}\int_0^1 B(r)^2 dr. \qquad (A4.5)$$

The result of (d) follows by combining (A4.4) and (A4.5). Therefore

$$M^{-1}T^{-2}(\hat{\alpha}-\alpha)^2\sum_{k=1}^{M} w_k \sum_{t=k+1}^{T} y_{t-1}y_{t-k-1}$$

$$\Rightarrow (\psi\sigma_\epsilon^2/2)\left((\int_0^1 B(r)^2dr)(\int_0^1 C(r)^2 + B(r)^2dr)^{-1}\right)^2\left(\int_0^1 (C(r)^2 - B(r)^2)dr\right) \equiv \psi\sigma_\epsilon^2\lambda_2/2.$$

Combining (A4.1), Lemma 4.3, (4.5) and (4.3), we have

Lemma 4.4 : *Let s_{WA}^2 be defined by (1.13), λ_2 be defined as in (4.6), let $M \to \infty$ with $M/T \to 0$ as $T \to \infty$. Then*

$$M^{-1}T^{-1}s_{WA}^2 \Rightarrow \sigma_\epsilon^2\psi\lambda_2.$$

The results of Theorem 4.3 follow directly from Lemmas 4.1 and 4.4.

Table 1: Exact Size of the Phillips–Perron Tests; 5% Nominal Size.

Estimate of σ^2		$Z(\alpha)$			$Z(t)$			ADF(t)
		S^2_{wa}	S^2_{ar}	σ^2	S^2_{wa}	S^2_{ar}	σ^2	
a) T = 100								
i.i.d. errors		.026	.046	.049	.043	.058	.058	.057
MA(1)	−.800	.992	.794	.689	.993	.964	.982	.553
	−.500	.444	.146	.118	.451	.295	.269	.159
	−.200	.104	.054	.062	.123	.082	.089	.094
	.200	.022	.055	.048	.033	.061	.052	.062
	.500	.013	.065	.061	.027	.063	.058	.068
	.800	.028	.026	.049	.022	.030	.045	.062
AR(1)	−.800	.773	.307	.295	.726	.554	.552	.058
	−.500	.255	.089	.087	.273	.166	.154	.061
	−.200	.104	.071	.066	.108	.097	.092	.089
	.200	.024	.049	.050	.032	.050	.052	.057
	.500	.007	.067	.073	.029	.070	.076	.056
	.800	.002	.074	.132	.019	.068	.111	.057
b) T = 500								
i.i.d. errors		.051	.073	.055	.048	.066	.054	.047
MA(1)	−.800	.971	.631	.623	.967	.858	.869	.352
	−.500	.345	.085	.078	.332	.135	.119	.085
	−.200	.104	.063	.063	.098	.073	.070	.059
	.200	.036	.062	.047	.039	.055	.052	.062
	.500	.038	.054	.052	.035	.047	.044	.049
	.800	.038	.058	.062	.047	.062	.059	.058
AR(1)	−.800	.653	.170	.172	.638	.276	.274	.051
	−.500	.195	.073	.069	.184	.082	.081	.049
	−.200	.083	.050	.060	.080	.055	.062	.066
	.200	.041	.057	.059	.045	.057	.062	.061
	.500	.021	.064	.057	.033	.056	.055	.062
	.800	.013	.065	.080	.024	.053	.067	.043

Table 2: Exact Size of the Modified Tests; 5% Nominal Size.

Estimate of σ^2	MZ(α)			MSB			MZ(t)		
	S^2_{wa}	S^2_{ar}	σ^2	S^2_{wa}	S^2_{ar}	σ^2	S^2_{wa}	S^2_{ar}	σ^2
a) T = 100									
i.i.d. errors	.012	.039	.039	.021	.049	.042	.026	.036	.036
MA(1) −.800	.981	.077	.000	.988	.088	.000	.977	.057	.001
−.500	.352	.021	.012	.395	.030	.001	.326	.026	.009
−.200	.069	.030	.031	.089	.038	.042	.071	.033	.040
.200	.012	.043	.034	.017	.054	.050	.027	.051	.038
.500	.013	.055	.047	.026	.066	.063	.019	.051	.050
.800	.009	.019	.038	.013	.027	.046	.018	.028	.041
AR(1) −.800	.676	.017	.002	.709	.020	.003	.649	.017	.007
−.500	.195	.024	.013	.228	.033	.021	.177	.023	.014
−.200	.057	.043	.035	.075	.055	.042	.064	.045	.039
.200	.013	.038	.034	.017	.052	.046	.023	.043	.040
.500	.004	.054	.057	.007	.067	.073	.024	.064	.069
.800	.001	.062	.122	.001	.086	.141	.018	.066	.110
b) T = 500									
i.i.d. errors	.037	.061	.043	.052	.069	.054	.044	.060	.048
MA(1) −.800	.961	.010	.000	.970	.012	.000	.950	.010	.005
−.500	.304	.042	.021	.336	.049	.032	.285	.060	.040
−.200	.072	.055	.047	.100	.063	.055	.088	.063	.053
.200	.025	.052	.034	.034	.063	.047	.039	.051	.047
.500	.028	.045	.040	.038	.059	.057	.034	.046	.043
.800	.024	.048	.044	.038	.058	.062	.046	.059	.057
AR(1) −.800	.624	.027	.030	.653	.036	.041	.594	.038	.041
−.500	.169	.040	.035	.192	.055	.045	.150	.048	.047
−.200	.059	.029	.037	.083	.042	.048	.073	.049	.052
.200	.027	.042	.046	.042	.066	.064	.044	.053	.058
.500	.013	.048	.039	.025	.057	.064	.032	.055	.055
.800	.009	.046	.063	.009	.060	.081	.023	.053	.064

Table 3: Power of the Modified Statistics Using the Autoregressive Spectral Density Estimator.
(5% one-tailed tests)

	MZ(α)			MSB			MZ(t)			ADF(t)		
α	.95	.90	.85	.95	.90	.85	.95	.90	.85	.95	.90	.85
a) T = 100												
i.i.d. errors	.224	.409	.613	.258	.449	.665	.155	.302	.496	.144	.339	.618
MA(1)–.800	.386	.701	.884	.423	.739	.902	.279	.576	.811	.881	.976	.996
–.500	.135	.338	.558	.159	.389	.606	.083	.242	.431	.408	.684	.868
–.200	.174	.392	.598	.201	.429	.648	.118	.278	.463	.259	.526	.762
.200	.233	.423	.617	.261	.468	.664	.160	.306	.493	.140	.302	.536
.500	.191	.397	.603	.232	.437	.645	.137	.283	.476	.164	.341	.529
.800	.122	.263	.443	.146	.313	.488	.087	.177	.336	.148	.288	.428
AR(1) –.800	.065	.182	.309	.083	.203	.350	.035	.114	.206	.138	.336	.617
–.500	.144	.335	.536	.169	.375	.583	.084	.237	.397	.137	.322	.602
–.200	.197	.401	.608	.233	.449	.651	.127	.289	.476	.212	.473	.602
.200	.200	.425	.629	.232	.473	.666	.148	.315	.490	.112	.276	.696
.500	.223	.413	.585	.252	.471	.630	.164	.307	.462	.115	.266	.430
.800	.226	.360	.472	.271	.407	.526	.171	.290	.369	.117	.198	.274
b) T = 500												
i.i.d. errors	.893	.985	.994	.916	.988	.997	.814	.969	.987	.961	1.000	1.000
MA(1)–.800	.521	.779	.923	.559	.809	.937	.391	.716	.890	1.000	1.000	1.000
–.500	.842	.957	.982	.877	.967	.988	.775	.912	.959	.964	1.000	1.000
–.200	.885	.980	.997	.916	.988	.999	.810	.961	.988	.965	1.000	1.000
.200	.884	.988	.998	.916	.991	1.000	.899	.970	.986	.945	1.000	1.000
.500	.893	.983	.996	.919	.988	.998	.812	.959	.981	.902	.999	1.000
.800	.867	.982	.997	.893	.987	.997	.774	.963	.983	1.000	1.000	1.000
AR(1) –.800	.944	1.000	1.000	.966	1.000	1.000	.836	1.000	1.000	.945	1.000	1.000
–.500	.970	1.000	1.000	.983	1.000	1.000	.918	1.000	1.000	.957	1.000	1.000
–.200	.981	1.000	1.000	.987	1.000	1.000	.930	1.000	1.000	.958	1.000	1.000
.200	.964	1.000	1.000	.980	1.000	1.000	.929	1.000	1.000	.929	1.000	1.000
.500	.971	1.000	1.000	.980	1.000	1.000	.904	1.000	1.000	.920	1.000	1.000
.800	.925	1.000	1.000	.954	1.000	1.000	.829	1.000	1.000	.815	1.000	1.000

Figure 1: Nearly Integrated, Nearly White Noise Model

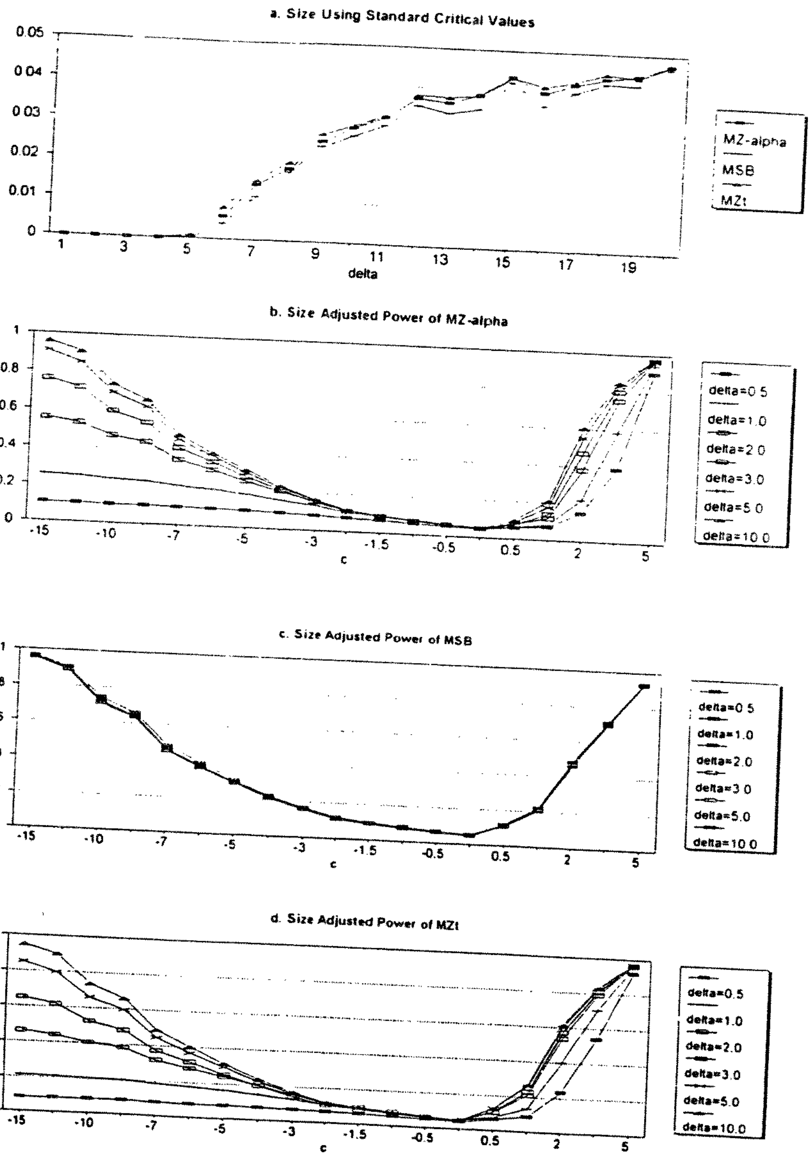
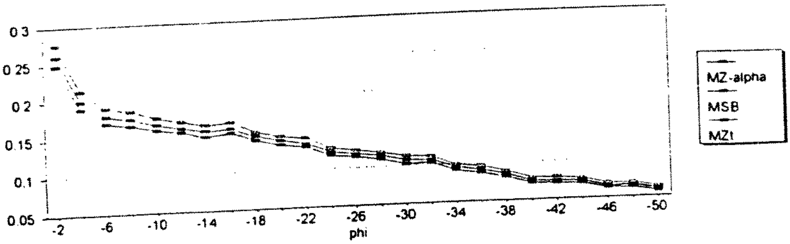
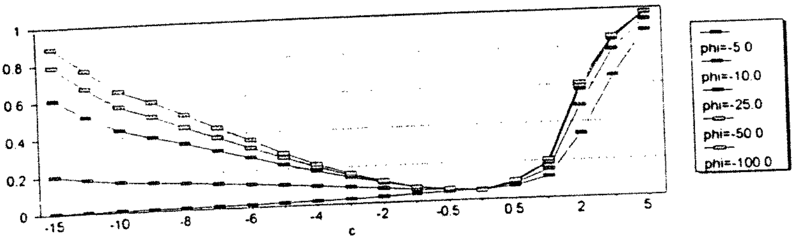


Figure 2: Nearly Twice Integrated Model

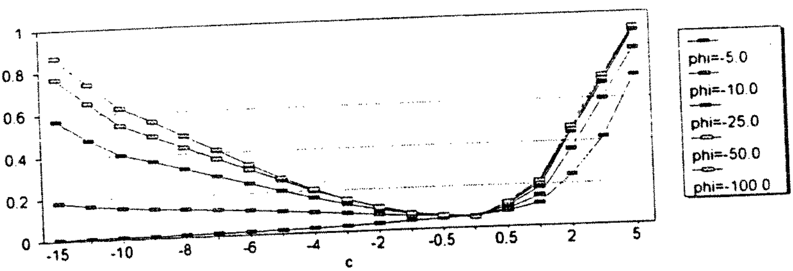
a. Size Using Standard Critical Values



b. Size Adjusted Power of MZ-alpha



c. Size Adjusted Power of MSB



d. Size Adjusted Power of MZt

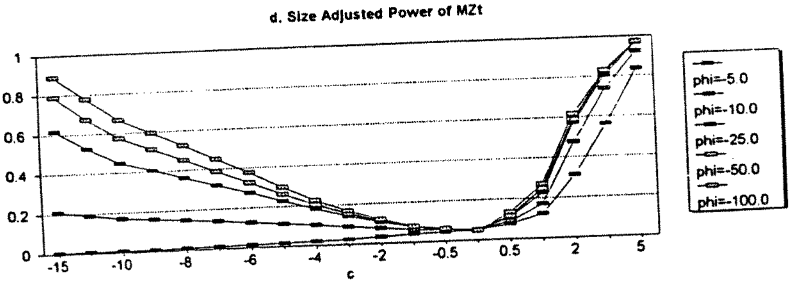
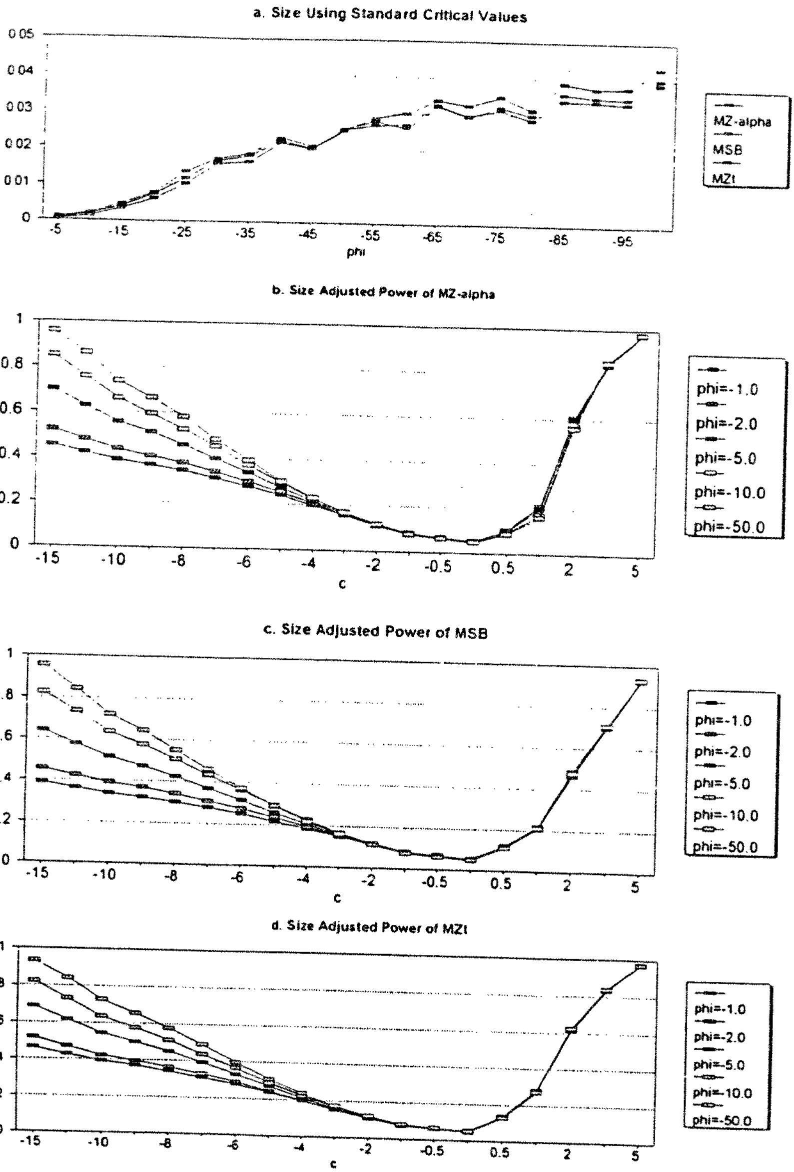


Figure 3: Nearly Seasonally Integrated Model



Université de Montréal
Département de sciences économiques
Centre de documentation
C.P. 6128, succursale Centre-ville
Montréal (Québec)
H3C 3J7

Cahiers de recherche (Discussion papers)
1993 à aujourd'hui (1993 to date)

Si vous désirez obtenir un exemplaire, vous n'avez qu'à faire parvenir votre demande et votre paiement (5 $ l'unité) à l'adresse ci-haut mentionnée. / To obtain a copy ($ 5 each), please send your request and prepayment to the above-mentioned address.

9301 : Mercenier, Jean, "Nonuniqueness of Solutions in Applied General-Equilibrium Models with Scale Economies and Imperfect Competition : A Theoretical *Curiosum*?", 26 pages.
9302 : Lemieux, Thomas, "Unions and Wage Inequality in Canada and in the United States", 66 pages.
9303 : Lemieux, Thomas, "Estimating the Effects of Unions on Wage Inequality in a Two-Sector Model with Comparative Advantage and Non-Random Selection", 54 pages.
9304 : Harchaoui, Tarek H., "Time-Varying Risks and Returns : Evidence from Mining Industries Data", 20 pages.
9305 : Lévy-Garboua, Louis et Claude Montmarquette, "Une étude économétrique de la demande de théâtre sur données individuelles", 40 pages.
9306 : Montmarquette, Claude, Rachel Houle et Sophie Mahseredjian, "The Determinants of University Dropouts : A Longitudinal Analysis", 17 pages.
9307 : Gaudry, Marc, Benedikt Mandel et Werner Rothengatter, "A Disagregate Box-Cox Logit Mode Choice Model of Intercity Passenger Travel in Germany", 17 pages.
9308 : Fortin, Nicole M., "Borrowing Constraints and Female Labor Supply : Nonparametric and Parametric Evidence of the Impact of Mortgage Lending Rules", 38 pages.
9309 : Dionne, Georges, Robert Gagné, François Gagnon et Charles Vanasse, "Debt, Moral Hazard and Airline Safety : an Empirical Evidence", 34 pages.
9310 : Dionne, Georges, Anne Gibbens et Pierre St-Michel, "An Economic Analysis of Insurance Fraud", 40 pages.
9311 : Gaudry, Marc, "Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes", 26 pages.
9312 : Laferrière, Richard et Marc Gaudry, "Testing the Linear Inverse Power Transformation Logit Mode Choice Model", 29 pages.

9313 : Kollmann, Robert, "Fiscal Policy, Technology Shocks and the US Trade Balance Deficit", 38 pages.
9314 : Ghysels, Eric, "A Time Series Model with Periodic Stochastic Regime Switching", 54 pages.
9315 : Allard, Marie, Camille Bronsard et Lise Salvas-Bronsard, "C*-Conjugate Expectations and Duality", 22 pages.
9316 : Dudley, Leonard et Claude Montmarquette, "Government Size and Economic Convergence", 28 pages.
9317 : Bronsard, Camille, "L'histoire de l'économie mathématique racontée à Juliette", 17 pages.
9318 : Tremblay, Rodrigue, "The Quest for Competitiveness and Export-Led Growth", 16 pages.
9319 : Proulx, Pierre-Paul, "L'ALÉNA", 12 pages.
9320 : Proulx, Pierre-Paul, "Le Québec dans l'ALÉNA", 28 pages.
9321 : Dionne, Georges, Denise Desjardins, Claire Laberge-Nadeau et Urs Magg, "Medical Conditions, Risk Exposure and Truck Drivers' Accidents : an Analysis with Count Data Regression Models", 20 pages.
9322 : Ghysels, Eric, "Seasonal Adjustment and other Data Transformations", 28 pages.
9323 : Dufour, Jean-Marie et David Tessier, "On the Relationship between Impulse Response Analysis, Innovation Accounting and Granger Causality", 12 pages.
9324 : Dufour, Jean-Marie et Eric Renault, "Causalités à court et à long terme dans les modèles VAR et ARIMA multivariés", 68 pages.
9325 : Ghysels, Eric et Alastair Hall, "On Periodic Time Series and Testing the Unit Root Hypothesis", 36 pages.
9326 : Campbell, Bryan et Jean-Marie Dufour, "Exact Nonparametric Orthogonality and Random Walk Tests", 28 pages.
9327 : Proulx, Pierre-Paul, "Quebec in North America : from a Borderlands to a Borderless Economy : an Examination of its Trade Flows with the U.S.A. at the National and Regional Levels", 24 pages.
9328 : Proulx, Pierre-Paul, "L'ALÉNA, le Québec et la mutation de son espace économique", 36 pages.
9329 : Sprumont, Yves, "Strategyproof Collective Choice in Economic and Political Environments", 48 pages.
9330 : Cardia, Emanuela et Steve Ambler, "The Cyclical Behaviour of Wages and Profits under Imperfect Competition", 24 pages.
9331 : Arcand, Jean-Louis L. et Elise S. Brezis, "Disequilibrium Dynamics During the Great Depression", 64 pages.
9332 : Beaudry, Paul et Michel Poitevin, "Contract Renegotiation : a Simple Framework and Implications for Organization Theory", 48 pages.
9333 : Dagenais, Marcel G. et Jean-Marie Dufour, "Pitfalls of Rescaling Regression Models with Box-Cox Transformations", 16 pages.
9334 : Bonomo, Marco et René Garcia, "Disappointment Aversion as a Solution to the Equity Premium and the Risk-Free Rate Puzzles", 40 pages.

9335 : Ghysels, Eric et Offer Lieberman, "Dynamic Regression and Filtered Data Series : A Laplace Approximation to the Effects of Filtering in Small Samples", 24 pages.

9336 : Kollmann, Robert, "The Duration of Unemployment as a Signal : Implications for Labor Market Equilibrium", 19 pages.

9337 : Kollmann, Robert, "Fertility, Consumption and Bequests in a Model with Non-Dynastic Parental Altruism", 22 pages.

9401 : Mercenier, Jean et Bernardin Akitoby, "On Intertemporal General-Equilibrium Reallocation Effects of Europe's Move to a Single Market", janvier 1994, 41 pages.

9402 : Gauthier, Céline et Michel Poitevin, "Using Ex Ante Payments in Self-Enforcing Risk-Sharing Contracts", février 1994, 38 pages.

9403 : Ghysels, Eric et Joanna Jasiak, "Stochastic Volatility and Time Deformation : an Application of Trading Volume and Leverage Effects", février 1994, 37 pages.

9404 : Dagenais, Marcel G. et Denyse L. Dagenais, "GMM Estimators for Linear Regression Models with Errors in the Variables", avril 1994, 33 pages.

9405 : Bronsard, C., Fabienne Rosenwald et Lise Salvas-Bronsard, "Evidence on Corporate Private Debt Finance and the Term Structure of Interest Rates", avril 1994, 42 pages.

9406 : Dinardo, John, Nicole M. Fortin et Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992 : A Semiparametric Approach", avril 1994, 73 pages.

9407 : Campbell, Bryan et Jean-Marie Dufour, "Exact Nonparametric Tests of Orthogonality and Random Walk in the Presence of a Drift Parameter", avril 1994, 32 pages.

9408 : Bollerslev, Tim et Eric Ghysels, "Periodic Autoregressive Conditional Heteroskedasticity", mai 1994, 29 pages.

9409 : Cardia, Emanuela, "The Effects of Government Financial Policies : Can We Assume Ricardian Equivalence?", mai 1994, 42 pages.

9410 : Kollmann, Robert, "Hidden Unemployment : A Search Theoretic Interpretation", mai 1994, 9 pages.

9411 : Kollmann, Robert, "The Correlation of Productivity Growth Across Regions and Industries in the US", juin 1994, 14 pages.

9412 : Gaudry, Marc, Benedikt Mandel et Werner Rothengatter, "Introducing Spatial Competition through an Autoregressive Contiguous Distributed (AR-C-D) Process in Intercity Generation-Distribution Models within a Quasi-Direct Format (QDF)", juin 1994, 64 pages.

9413 : Gaudry, Marc et Alexandre Le Leyzour, "Improving a Fragile Linear Logit Model Specified for High Speed Rail Demand Analysis in the Quebec-Windsor Corridor of Canada", août 1994, 39 pages.

9414 : Lewis, Tracy et Michel Poitevin, "Disclosure of Information in Regulatory Proceedings", juillet 1994, 38 pages.

9415 : Ambler, Steve, Emanuela Cardia et Jeannine Farazli, "Export Promotion and Growth", août 1994, 41 pages.

9416 : Ghysels, Eric et Haldun Sarlan, "On the Analysis of Business Cycles Through the Spectrum of Chronologies", août 1994, 37 pages.

9417 : Martel, Jocelyn et Timothy C.G. Fisher, "The Creditors' Financial Reorganization Decision : New Evidence from Canadian Data", août 1994, 21 pages.

9418 : Cannings, Kathy, Claude Montmarquette et Sophie Mahseredjian, "Entrance Quotas and Admission to Medical Schools : A Sequential Probit Model", septembre 1994, 26 pages.

9419 : Cannings, Kathy, Claude Montmarquette et Sophie Mahseredjian, "Major Choices : Undergraduate Concentrations and the Probability of Graduation", septembre 1994, 26 pages.

9420 : Nabeya, Seiji et Pierre Perron, "Approximations to Some Exact Distributions in the First Order Autoregressive Model with Dependent Errors", septembre 1994, 40 pages.

9421 : Perron, Pierre, "Further Evidence on Breaking Trend Functions in Macroeconomic Variables", octobre 1994, 50 pages.

9422 : Vogelsang, Timothy J. et Pierre Perron, "Additional Tests for a Unit Root Allowing for a Break in the Trend Function at an Unknown Time", novembre 1994, 57 pages.

9423 : Ng, Serena et Pierre Perron, "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag", décembre 1994, 41 pages.

9424 : Perron, Pierre, "The Adequacy of Asymptotic Approximations in the Near-Integrated Autoregressive Model with Dependent Errors", décembre 1994, 37 pages.

9425 : Ghysels, Eric et Pierre Perron, "The Effect of Linear Filters on Dynamic Time Series with Structural Change", décembre 1994, 35 pages.

9426 : Boyer, Marcel, Jean-Jacques Laffont, Philippe Mahenc et Michel Moreaux, "Sequential Location Equilibria Under Incomplete Information", décembre 1994, 38 pages.

9427 : Perron, Pierre et Serena NG, "Useful Modifications to Some Unit Root Tests with Dependent Errors and their Local Asymptotic Properties", décembre 1994, 41 pages.