Rochester Center for Economic Research

Unit Roots in Macroeconometrics: A Survey

Ogaki, Masao

Working Paper No. 364
October 1993

University of Rochester

Unit Roots in Macroeconometrics: A Survey

Masao Ogaki

Rochester Center for Economic Research
Working Paper No. 364

Forthcoming in *Monetary and Economic Studies*

UNIT ROOTS IN MACROECONOMETRICS: A SURVEY

Masao Ogaki

Department of Economics

University of Rochester

March, 1993

This paper is a revision and extension of part of my paper entitled "Applications of Cointegration to Models with Optimizing Agents." I thank Mark Dwyer fir research assitance. Financial support by the Bank of Japan and NSF Grant SES-9213930 is gratefully acknowledged.

Abstrct

This paper provides a selective survey of the recent literature of unit root econometrics. Since the seminal work of Nelson and Plosser (1982) was published, much theoretical and empirical research has been done in the area of unit root nonstationarity. Nelson and Plosser found that the null hypothesis of unit root nonstationarity was not rejected for many macroeconomic series. When a linear combination of unit root nonstationary variables is stationary, they are said to be cointegrated. Recent developments in estimation method for coitegrated systems allow researchers to estimate structural parameters and make inferences without exogeneity assumptions.

Table of Contents

1. Introduction

Since the seminal work of Nelson and Plosser (1982) was published, much theoretical and empirical research has been done in the area of unit root nonstationarity. Nelson and Plosser found that the null hypothesis of unit root nonstationarity was not rejected for many macroeconomic series. When one or more variables of interest are unit root nonstationary, standard asymptotic distribution theory does not apply to the econometric system involving these variables. The spurious regresssion results disucssed in Section 3 are concrete examples of this type of problem.

When a variable is unit root nonstationary, it has a stochastic trend. If linear combinations of two or more unit root nonstationary variables do not contain stochastic trends, then these variables are said to be cointegrated. Then the cointegrating vector, that eliminates the stochastic trends, can be estimated consistently by regressions without the use of instrumental variables, even when no variables are exogenous. If the cointegrating vector includes structural parameters, then the econometrician can estiamte these structural parameters without making exogeneity assumptions.[1]

The rest of this paper is organized as follows. In Section 2, univariate unit root econometrics is discussed. It begins with definitions of basic concepts such as stationarity, difference stationarity, and trend stationarity. Then a decomposition of a difference stationary variable into a determinstic trend, a stochastic trend, and a stationary component is

---

[1]Stock and Watson (1988b), Deibold and Nerlove (1990), Cambell and Perron (1991), and Watson (1992) are examples of surveys for unit root econometrics.

discussed. Spurious regression results, tests for the null of differnce stationarity, and tests for the null of stationarity are reviewed.

Section 3 reviews multivariate unit root econometrics. Cointegration, stochastic cointegration, and the deterministic cointegration restriction are defined. Then some estimators for cointegrating vectors are described. Tests for the null of no cointegration the null of cointegratio as well as tests for the number of cointegrating vectors are presented.

Section 4 discusses how cointegration may be combined with standard econometric methods that assume stationarity. This is accomplished in the context of the Generalized Method of Moments, which includes many estimation methods as special cases.

2. Unit Root Nonstationarity

This section deals with a time series of a scalar random variable.

2.1. Definitions

Consider a stochastic process, $\{x_t\colon t=\ldots,-2,-1,0,1,2,\ldots\}$, which is a sequence of random variables. If the joint distributions of $\{x_t,x_{t+1},\ldots,x_{t+\tau}\}$ are the same as $\{x_{t+k},x_{t+k+1},\ldots,x_{t+k+\tau}\}$, then x_t is (strictly) stationary. If x_t has finite second moments, and if $E(x_t) = E(x_{t+\tau})$ and $E(x_t x_{t-\tau})=E(x_{t+k}x_{t+k-\tau})$ for all t, τ, k, then x_t is said to be covariance stationary. If x_t is stationary and has finite second moments, then x_t is covariance stationary.

Many macroeconomic variables tend to grow over time, so that their distributions shift upward over time. Hence they are not stationary. However, there are many possible forms of nonstationarity, and it is not clear which form of nonstationarity is appropriate in representing macroeconomic variables. It may be reasonable to assume that the growth rate or the first difference of (natural) log of a variable is stationary for many macroeconomic variables. Let us now assume that the first difference of x_t ($\Delta x_t = x_t - x_{t-1}$) is stationary. Then x_t is either difference stationary or trend stationary. If x_t is stationary after removing a deterministic time trend, then x_t is said to be trend stationary. Because Δx_t is assumed to be stationary, x_t has a linear time trend when x_t is trend stationary:

$$(2.1) \qquad x_t = \theta + \mu t + \varepsilon_t,$$

where ε_t is stationary with mean zero. If Δx_t is stationary and if Δx_t has a positive long-run variance, then x_t is said to be (first) difference stationary. Alternatively, x_t is said to be unit root nonstationary or integrated of order one. The long-run variance of a stationary variable y_t is defined by

$$(2.2) \qquad \omega^2 = \sum_{\tau=-\infty}^{\infty} \mathrm{E}\{[y_t - \mathrm{E}(y_t)][y_{t-\tau} - \mathrm{E}(y_t)]\}.$$

The trend stationary process is also stationary after taking the first difference but its first difference has a zero long-run variance.

A special case of a difference stationary process is a random walk. If $\mathrm{E}(x_{t+1}|x_t, x_{t-1}, x_{t-2},...) = x_t$ and if $\mathrm{E}((\Delta x_{t+1})^2|x_t, x_{t-1}, x_{t-2},...)$ is constant over time, then x_t is a random walk. If Δx_t is a random walk, then Δx_t does not have serial correlation. In general, if x_t is difference stationary, then Δx_t has nonzero serial correlation.

2.2. Decompostions

It is often convenient to decompose a difference stationary process into components representing a determinstic trend, a stochastic trend, and a stationary component.

Let x_t be a difference stationary process:

(2.3) $$x_t - x_{t-1} = \mu + \varepsilon_t$$

for t≥1 where ε_t is stationary with mean zero. Here μ is called a drift, which is the mean of Δx_t. Then

(2.4) $$x_t = \mu + x_{t-1} + \varepsilon_t = 2\mu + x_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

$$= 3\mu + x_{t-3} + \varepsilon_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} = \dots = \mu t + x_0 + \sum_{\tau=1}^{t} \varepsilon_\tau.$$

Hence

(2.5) $$x_t = \mu t + x_t^0$$

where x_t^0 is

(2.6) $$x_t^0 = x_0 + \sum_{\tau=1}^{t} \varepsilon_t.$$

Relation (2.6) decomposes the difference stationary process x_t into a deterministic trend arising from drift μ, and the difference stationary process without drift x_t^0.

Let us now consider the Beveridge-Nelson (1981) decomposition, which further decompose x_t^0 into a random walk component and a stationary component. Because Δx_t^0 is covariance stationary, it has a Wold representation:

(2.7) $$(1\text{-L})x_t^0 = A(L)\nu_t,$$

where L is the lag operator, $A(L)=\sum_{\tau=0}^{\infty} A_\tau L^\tau$, and $\nu_t = x_t^0 - E(x_t^0 | x_{t-1}^0, x_{t-2}^0, \dots)$. Here $E(\cdot | x_{t-1}^0, x_{t-2}^0, \dots)$ is the linear projection operator. Then

(2.8) $x_t^0 = z_t + c_t$,

where

(2.9) $z_t = z_{t-1} + A(1)v_t$,

is the random walk component or a stochastic trend, and

(2.10) $c_t = -\{(\sum_{\tau=1}^{\infty} A_\tau)v_t + (\sum_{\tau=2}^{\infty} A_\tau)v_{t-1} + (\sum_{\tau=3}^{\infty} A_\tau)v_{t-2} + ...\}$

is the stationary component of x_t. Thus a difference stationary process x_t is decomposed into a deterministic trend, a stochastic trend, and a stationary component.

The variance of the random walk component, Var(Δz_t), is equal to $A(1)^2$Var(v_t), which in turn is equal to the long-run variance of Δx_t and 2π times the spectral density of Δx_t at frequency zero. If the long-run variance is zero, then $x_t = \mu t + c_t$, and x_t is trend stationary.

Cochrane (1988), among others, uses Var(Δz_t)/Var(Δx_t) as a measure of persistence of x_t. This measure is zero for trend stationary x_t and is one for a random walk. Cochrane estimates Var(Δz_t) by 1/k times the variance of k-differences of x_t for a large enough k. Cochrane's estimator is essentially the same as the Barlett estimator, which was advocated by Newey and West (1984) in a different context. Any estiamtor of the long-run variance or the spectral density at frequency zero can be used for the purpose of estimating Cochrane's measure of persistence.

2.3. Spurious Regressions

One reason why macroeconomists need to be careful about unit root nonstationary variables is that standard regression theory can be very misleading when variables in a regression are difference stationary.

For example, suppose that y_t is a random walk and x_t is a random walk which is independent of y_t. Granger and Newbold (1974) found that the standard Wald test for the hypothesis that the coefficient on x_t is zero tended to be large (compared with standard critical values) in ordinary least squares (OLS) regressions of y_t onto x_t in their Monte Carlo experiments. Later, Phillips (1986) showed that the Wald test diverges to infinity as the sample size is increased. In a regression with two independent difference stationary variables without drift, the random walk components will dominate the stationary components at least asymptotically. Hence these spurious regression results imply that the absolute value of the t-ratio of the regressor tends to be larger than the critical valued implied by the standard statistical theory that assumes stationarity. An econometrician who ignores unit root nonstationarity issues tend to spuriously conclude that two independent difference staitonary variables are related.

Another example of the spurious regression results is in Durlauf and Phillips (1988). When a difference stationary x_t without drift is regressed onto a constant and a linear time trend, the Wald test statistic for the hypothesis that a coeficient for the linear trend is zero diverges to infinity as the sample size increases.

2.4. Near Observational Equivalence

Most of the tests that will be described in sections 2.4 and 2.5 below seek to discriminate between difference stationary and trend stationary processes. In the finite samples that we observe, there is a conceptual difficulty with this task. In finite samples, any difference stationary process can be approximated arbitrary well by a series of trend stationary processes. This can be done by driving the dominant autoregressive root of trend stationary processes to one from below. After all, it is very difficult to discriminate between the dominant autoregressive root of 0.999 and that of one. This type of problem exists for virtually any hypothesis testing. What is special about hypothesis testing for unit root nonstationarity is that the opposit is also true: any trend stationary process can be approximated arbitrary well by a series of difference stationary processes. This can be done by driving the long-run variance of the first difference of diffrence staitonary processes to zero. Some authors call this problem the near observational equivalence problem (see, e.g., Blough (1988), Campbell and Perron (1991), Christiano and Eichenbaum (1990), Cochrane (1988)).

2.5. Tests for the Null of Difference Stationarity

This section explains Dickey-Fuller (1979), Said-Dickey (1984), Phillips-Perron (1988), and Park's (1990a) tests for the null of difference stationarity. More recent work to improve small sample properties of tests includes Kahn and Ogaki (1990), Elliott, Rothenburg, and Stock (1992), and Hansen (1993).

2.5.1. Dickey-Fuller Tests

Dickey and Fuller (1979) propose to test for the null of a unit root in an AR(1) model:[2]

---

[2]It should be noted that Dickey and Fuller's (1981) joint tests with determinstic terms can have significantly lower power than Dickey and Fuller's (1979) one-tailed single unit root tests as explained by Parks (1989).

(2.11) $$x_t = \theta + \mu t + \alpha x_{t-1} + \varepsilon_t.$$

where ε_t is NID. One of their test is based on $T(\hat{\alpha}-1)$, where T is the sample size and $\hat{\alpha}$ is the OLS estimator for α in (2.11) and another test is based on the t-ratio for the hypothesis $\alpha=1$. These test statistics do not have standard distributions. Depending on whether or not a constant and a linear time trend are included, distributions of these tests under the null are different.[3] Fuller (1976, Tables 8.5.1 and 8.5.2) tabulates critical values for Dickey-Fuller tests.

Whether or not a constant and a linear time trend should be included in the regression depends on what type of alternative is appropriate. If the alternative hypothesis is that x_t is stationary with mean zero, then no deterministic terms should be included. This alternative is not appropriate for most of the macroeconomic time series. If the alternative hypothesis is that x_t is stationary with unknown mean, then a constant should be included. This alternative is appropriate for the time series which exhibit a consistent tendency to grow (or shrink) over time. If the alternative is that x_t is trend stationary, then a constant and a linear time trend should be included. This alternative is appropriate for the time series which exhibit a consistent tendency to grow (or shrink) over time. When these test statistics are negative and greater than the appropriate critical value in absolute value, then the null of a unit root is rejected in favor of one of these alternatives.

[3]If the data are demeaned prior to the regression, then the test statistics have the same distributions as those from the regression with a constant in (2.11). If the data are detrended prior to the regression, then the test statistics have the same distributions as those from the regression with a constant and a linear time trend.

Dickey-Fuller tests assume that the econometrician knows the order of AR. The following tests treat the case of unknown order of AR (or even more general cases).

2.5.2. Said-Dickey Test

Said and Dickey (1984) extend the Dickey-Fuller's t-ratio test to the case where the order of AR is unknown. Consider a regression

$$(2.11) \qquad \Delta x_t = \theta + \mu t + \rho x_{t-1} + \beta_1 \Delta x_{t-1} + \dots + \beta_p \Delta x_{t-p} + \nu_t.$$

When the order of the AR, p, is increased as the sample size at a certain rate, the t-ratio for the hypothesis $\rho=0$ has the same asymptotic distribution as Dickey-Fuller t-ratio test. Some authors call this test the augmented Dickey-Fuller (ADF) test while others reserve the word ADF for the corresponding cointegration test. A constant and a linear time trend are included or excluded according to the appropariate alternative hypothesis as before.

In many applications, the Said-Dickey test results are very sensitive to the choice of the order of the AR, p. Campbell and Perron (1991) propose to start with a reasonably large value of p that is chosen a priori and decreases p until the coefficient on the last included lag is significant.

2.5.3. Phillips-Perron Tests

Phillips (1987) and Phillips and Perron (1988) use a nonparametric method to correct for serial correlation of ε_t. Their modification of the Dickey-Fuller $T(\hat{\alpha}-1)$ test is called $Z(\alpha)$ test, while their modification of the Dickey-Fuller t-ratio test is called Z(t) test. These corrections are based on a nonparametric estimate of the long run variance of ε_t. See

Section 3.2.3 below for a discussion of nonparametric estimation methods. Phillips-Perron tests are constructed so that they have the same asymptotic distributions as corresponding Dickey-Fuller tests.

An advantage of the Phillips-Perron tests over the Said-Dickey test is that they tend to be more powerful as shown in Monte Carlo experiments of Phillips and Perron. A drawback of the Phillips-Perron tests is that they are subject to more severe size distortions than the Sadi-Dickey test (see Monte Carlo results of Phillips and Perron(1988) and Schwert (1989)). Size distortion exists when the actual size of a test in small samples is very different from the size of the test indicated by asymptotic theory. Such differences are due to approximations involved in the asymptotic theory.

2.5.4. Park's J Tests

Park's (1990a) J tests based on a variable addition method are originally proposed by Park and Choi (1988). These tests are based on spurious regression results. Consider a regression

$$(2.12) \qquad x_t = \sum_{\tau=0}^{p} \mu_\tau t^\tau + \sum_{\tau=p+1}^{q} \mu_\tau t^\tau + \eta_t.$$

Here the maintained hypothes is that x_t possesses the deterministic time polynominals up to the order of p (typically, p is zero or one). The additional time polynomials are spurious time trends. Let F(p,q) be the standard Wald test statistic (without any correction for serial correlation of η_t) for the null hypothesis $\mu_{p+1}=...=\mu_q=0$. Under the null hypothesis that η_t is unit root nonstationary, spurious regression results imply that F(p,q) explodes but F(p,q)/T has an asymptotic distribution. The J(p,q) test is defined as F(p,q)/T. The null hypothesis of the difference staitonarity is rejected against the alternative of trend stationarity when

J(p,q) is *small* because J(p,q) converges to zero under the alternative hypothesis of trend stationarity. Part of Park and Choi's table of critical values for J tests are reproduced in Table 1 for convenience.

(Table 1 around here)

The J(p,q) tests do not require the estimation of the long-run variance of η_t, and thus have an advantage over the Said-Dickey and Phillips-Perron tests in that neither the order of autoregression nor the lag truncation number needs to be chosen. Park and Choi's Monte Carlo experiments show that J tests have relatively stable sizes and are not dominated by Said-Dickey and Phillips-Perron tests in terms of size-adjusted power.

2.6. Tests for the Null of Stationarity

In some cases, it is useful to test the null of stationarity (or trend stationarity) rather than the null of difference stationarity. For example, if an econometrician plans to apply econometric theory that assumes stationarity, a natural procedure is to test the null of stationarity rather than test the null of difference stationarity. Tests for the null of stationarity will also lead to tests for the null of cointegration as will be discussed in Section 3.4. However, most of the tests in the unit root literature take the null of a unit root rather than the null of stationarity. Only recently, Fukushige, Hatanaka, and Koto (1990), Kahn and Ogaki (1992), Kwiatkowski, Phillips, Schmidt, and Shin (1992), Birens and Guo (1993), and Choi and Ahn (1993) among others have developed tests for the null of stationarity.

Park's (1990a) G tests for the null of stationarity were first developed by Park and Choi (1988). These tests, which have been used in empirical work by several researchers, are based on the same spurious regression results as Park's J tests. With the notations in Section 2.4.4, $G(p,q)=F(p,q)\hat{\sigma}^2/\hat{\omega}^2$, where $\hat{\sigma}^2=(1/T)\sum_{t=1}^{T}\hat{\eta}_t^2$, $\hat{\omega}^2$ is an estimate of the long-run variance of η_t, and $\hat{\eta}_t$ is the esimated residual in regression (2.12). Under the null that x_t is stationary after removing the maintained deterministic time terms of time polynominal of order p, G(p,q) test has asymptotic chi-square distribution with the degree of freedom q-p. Under the alternative hypothesis that x_t is difference stationary (after removing the maintained deterministic terms), the G(p,q) statistic diverges to infinity. This is due to the spurious regression result that time polynonimals tend to mimic a stochastic trend.

Unlike Park's J tests, Park's G tests require estimation of the long-run variance. Kahn and Ogaki's (1992) Monte Carlo experiments on Park's G tests suggest that it is advisable to use relatively small q when the sample size is small and not to use prewhitening method discussed in Section 3.2.3.

3. Cointegration

When the stochastic trends of two or more difference stationary variables are eliminated by forming a linear combination of these variables, the variables are said to be cointegrated in the terminology of Engle and Granger (1987). What is striking about cointegration is that a cointegrating vector that eliminates the stochastic trends can be estimated consistently by regressions without using instrumental variables, even when no variables are exogenous.

3.1. Definitions

Let Z_t be a $n\times1$ vector of difference stationary random variables with ΔZ_t being stationary. If there exists a nonzero vector of real numbers α such that $\alpha' Z_t$ is stationary, then Z_t is said to be cointegrated with a cointegrating vector α. If α is a cointegrating vector, $b\alpha$ is also a cointegrating vector for any real number b. It is often convenient to normalize one element by one. Suppose that the first elment of α is nonzero, then partition Z_t by $Z_t=(y_t, X_t')$ and normalize α by $\alpha=(1,-\gamma)$. Here y_t is a difference stationary process, X_t is a vector difference stationary process, and γ is a normalized cointegrating vector.

There may exist more than one linearly independent cointegrating vectors. Suppose that there are r linear independent cointegrating vectors ($0\leq r\leq n-1$). Then Z_t possesses only $n-r$ common stochastic trends (see Stock and Watson (1988a)).

We now introduce the notions of stochastic cointegration and the determinstic cointegration restriction, as defined by Ogaki and Park (1989).[4] Consider a vector difference stationary process X_t with drift:

(3.1) $$X_t - X_{t-1} = \mu_x + \nu_t$$

for $t\geq1$ where μ_x is a $(n-1)$-dimensional vector of real numbers where ν_t is stationary with mean zero. As in (2.5), recursive substitution in (3.1) yields

(3.2) $$X_t = \mu_x t + X_t^0$$

---

[4]West (1988) consider estimation under the determinstic cointegration restriction for the special case of one regressor. Hansen (1992) and Park (1992) consider the deterministic cointegration restriction under more general cases.

where X_t^0 is difference stationary without drift. Relation (3.2) decomposes the difference stationary process X_t into deterministic trends arising from drift μ_x and the difference stationary process without drift, X_t^0. Suppose that y_t is a scalar difference stationary process with drift μ_y. Similarly, decompose y_t into a deterministic trend $\mu_y t$ and a difference stationary process without drift y_t^0 as in (2.7):

(3.3) $\quad y_t = \mu_y t + y_t^0.$

Difference stationary processes y_t and X_t are said to be *stochastically cointegrated* with a *normalized cointegrating vector* γ when there exists a *(n-1)*-dimensional vector γ such that $y_t^0 - \gamma' X_t^0$ is stationary. Stochastic cointegration only requires that stochastic trend components of the series are cointegrated. We may then write $y_t^0 - \gamma' X_t^0 = \theta_c + \varepsilon_t$, where ε_t is stationary with mean zero. Then by (3.2) and (3.3),

(3.4) $\quad y_t = \theta_c + \mu_c t + \gamma' X_t + \varepsilon_t$

where

(3.5) $\quad \mu_c = \mu_y - \gamma' \mu_x$.

Nest, suppose that a vector γ^* satisfies

(3.6) $\quad \mu_y = \gamma^{*\prime} \mu_x.$

Then $Y_t - \gamma^{*\prime} X_t$ does not possess any deterministic trend, and Y_t and X_t are *cotrended* with a *normalized cotrending vector* γ^*. If $n>2$ and if one of the components of μ_x is nonzero, there are infinitely many cotrending vectors. Consider an extra restriction that the normalized cointegrating vector γ is a cotrending vector. This restriction, which we call the *deterministic cointegration restriction*, requires that the cointegrating vector eliminates both the stochastic and deterministic trends. In this case,

(3.7) $$y_t = \theta_c + \gamma' X_t + \varepsilon_t.$$

3.2. Estimation

When y_t and X_t are cointegrated, an OLS regression (3.4) (or (3.7) either if the determistic cointegration restriction is satisfied or if the drift terms are known to be zero) is called a cointegrating regression. The OLS estimator is consistent (see Phillips and Durlauf (1986) and Stock (1987)), but is asymptotically biased. It also has a nonstandard distribution, which make statistical inference very difficult. For example, the OLS standard errors calculated in the standard econometric packages for OLS are not very meaningful for cointegrating regressions. Many efficient estimation methods that solve all or part of these problems have been developed by Phillips and Hansen (1990), Phillips and Loretan (1988), Saikkonen (1991), and Stock and Watson (1993) among others. In the following, we focus on Johansen's (1988, 1991) Maximum Likelihood Estimation and Park's (1992) Canonical Cointegrating Regressions.

3.2.1. Error Correction and Johansen's Maximum Likelihood Estimation

Johansen's (1988, 1991) maximum likelihood (ML) estimation is based on an error correction representation:

(3.8) $$\Delta Z_t = \mu + \lambda\alpha' Z_{t-1} + \beta_1 \Delta Z_{t-1} + \beta_2 \Delta Z_{t-2} + \ldots + \beta_p \Delta Z_{t-p} + \varepsilon_t,$$

where Z_t and ε_t are $n\times 1$ vectors of random variables, λ and α are $n\times r$ matrices of real numbers, and β_i's are $n\times n$ matrices of real numbers. The first term $\lambda\alpha' Z_{t-1}$ is called an error correction term.[5] Engle and Granger (1987) show that first difference stationary Z_t has a possibly infinite order error correction representation with a nonzero λ under general regularity conditions if Z_t is cointegrated with r linear independent cointegrating vectors. The columns of α are these cointegrating vectors.

It should be noted that Johansen's assumption that the error correction representation of finite order can be very restrictive in some applications. For example, Gregory, Pagan, and Smith (192) shows that linear quadratic economic models with adjustment costs imply moving average terms in the error correction representation. Phillips's (1991) ML estimation method may be useful in these circumstances.

Johansen makes an additional assumption that ε_t is normaly distributed and derive maximum likelihood estimator for α. In his procedure, all parameters are jointly estimated and his estimators are asymptotically efficient. Another way to estimate an error correction representation is to use Engle and Granger's (1987) two step estimation method. In the first step, cointegrating vectors are estimated. For example, if there is only one linear independent cointegrating vector, it can be estimated by an OLS. Other efficient estimators may be used in this first step. Then the rest of the parameters in the error correction representation are estimated in the second step. Because cointegrating vector estimators converge faster than $\sqrt{T}$, the first step estimation does not affect asymptotic distributions of the second step estimators. In the second step, only stationary variables

---

[5]Johansen uses an error correction term $\lambda\alpha' Z_{t-p}$ instead of more conventional $\lambda\alpha' Z_{t-1}$. Hoever, these two representations can be shown to be equivalent.

are involved, so the standard econometric theory can be used.

3.2.2. Canonical Cointegrating Regression

Johansen's maximum likelihood estimation makes a parametric correction for long-run correlation of ΔX_t and ε_t. Another way to obtain an efficient estimator is to utilize a nonparametric estimate of the long-run covariance parameters. Both Phillips and Hansen's (1990) and Park (1992) employ such covariance estimatees. Here, attention is confined to Park's Canonical Cointegration Regressions (CCR).

Consider a cointegrated system

(3.9) $$y_t = X_t'\gamma + \varepsilon_t$$

(3.10) $$\Delta X_t = \nu_t,$$

where y_t and X_t are difference stationary, and ε_t and ν_t are stationary with zero mean.[6] Here y_t is a scalar and X_t is a $(n\text{-}1)\times 1$ random vector. Let

(3.11) $$w_t = (\varepsilon_t, \nu_t').$$

Define $\Phi(i) = E(w_t w_{t-i}')$, $\Sigma=\Phi(0)$, $\Gamma = \sum_{i=0}^{\infty}\Phi(i)$, and $\Omega = \sum_{i=-\infty}^{\infty}\Phi(i)$. Here Ω is the matrix version of (2.2) and is called the long run variance (or covariance) matrix of w_t. Partition Ω as

(3.12) $$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \begin{matrix} 1 \\ n\text{-}1 \end{matrix}$$

and partition Γ comformably. Define

(3.13) $$\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$$

and $\Gamma_2 = (\Gamma_{12}', \Gamma_{22}')'$. The CCR procedure assumes that Ω_{22} is positive definite, implying that X_t is not itself cointegrated (see, e.g., Phiilips

[6]There are no deterministic terms in (1) and (2), but it is simple to analyze effects of deterministic terms on the asympotic distributions in the CCR system.

(1986) and Engle and Granger (1987)). This assumption assures that $(1,-\gamma)$ is the unique cointegrating vector (up to a scale factor).[7]

The OLS estimator in (3.9) is super-consistent in that the estimator converge to β at the rate of T (smaple size) even when $\Delta x(t)$ and $u(t)$ are correlated. The OLS estimator, however, is not asymptotically efficient. Consider transformations

(3.12) $$y_t^* = y_t + \Pi_y' w_t$$

(3.13) $$X_t^* = X_t + \Pi_x' w_t.$$

Because w_t is stationary, y_t^* and X_t^* are cointegrated with the same cointegrating vector $(1, -\beta)$ as y_t and X_t for any Π_y and Π_x. The idea of the CCR is to choose Π_y and Π_x, so that the OLS estimator is asymptotically efficient when y_t^* is regressed on X_t^*.[8] This requires

(3.14) $$\Pi_y = \Sigma^{-1}\Gamma_2\gamma + (0,\Omega_{12}\Omega_{22}^{-1})'$$

(3.15) $$\Pi_x = \Sigma^{-1}\Gamma_2.$$

In practice, long-run covariance parameters in these formulas are estimated, and estimated Π_y and Π_x are used to transform y_t and X_t. As long as these parameters are estimated consistently, the resultant CCR estimator

---

[7]For many applications, it is natural to assume that $\Delta^{-1}\varepsilon_t$ is not cointegrated with X_t. This assumption implies that $\Omega_{11.2}$ is positive. Park (1992) calls cointegration between y_t and X_t singular when $\Omega_{11.2}$ is zero. For the singular models, either a different CCR procedure described by Park is necessary (the removable singularity case) or the CCR procedure is not applicable (the essential singularity case).

[8]Under gemeral conditions, a sequence of functions $(1/T)^{1/2}\sum_{t=1}^{T} w(t)$ converges in distribution to a vector Brownian motion B with covariance matrix Ω. The OLS estimator converges in distribution to ????

is asymptotically efficient.

Here we have considered a single regression. If there are many cointegrating regressions with disturbances with nonzero long-run covariances in an econmetric system of interest, then asymptotically it is more efficient to apply seemingly unrelated regressions. Park and Ogaki (1990a) develop a method of Seemingly Unrelated Canonical Cointegrating Regressions (SUCCR) for this case. In the SUCCR, transformations of y_t and X_t that are slightly different from (3.14) and (3.15) are applied in each regression. After transforming variables, the standard seemingly unrelated regression method is applied to the transformed variables.

3.2.3. Estimation of Long-Run Covariance Parameters

In order to use efficient estimators for cointegrating vectors based on nonparametric correction such as CCR estimators, it is necessary to estimate long-run covariance parameters Ω and Γ. In many applications of cointegration, the order of serial correlation is unknown. Let $\Phi(\tau)=E(w_t w'_{t-\tau})$,

$$(3.16) \qquad \Phi_T(\tau) = \frac{1}{T}\sum_{t=\tau+1}^{T} \hat{w}_t \hat{w}_{t-\tau}' \qquad \text{for } \tau\geq 0,$$

and $\Phi_T(\tau)=\Phi_T(-\tau)'$ for $\tau<0$, where $\hat{w}_t$ is constructed from a consistent estimate of the cointegrating vector. Many estimators for Ω in the literature have the form

$$(3.17) \qquad \Omega_T = \frac{T}{T-p}\sum_{\tau=-T+1}^{T-1} k\left(\frac{\tau}{S_T}\right) \Phi_T(\tau),$$

where $k(\cdot)$ is a real-valued kernel, and S_T is a band-width parameter. The factor T/(T-p) is a small sample degrees of freedom adjustment. See Andrews (1991) for examples of kernels. Similary, Γ is estimated by

(3.18) $$\Gamma_T = \frac{T}{T-p} \sum_{\tau=0}^{T-1} k(\frac{\tau}{S_T})\, \Phi_T(\tau).$$

One important problem is how to choose the bandwidth parameter S_T. Andrews (1991) provides formulas for optimal choice of the bandwidth parameter for a variety of kernels. These formulas include unknown parameters and Andrews proposes automatic bandwidth estimators in which these unknown parameters are estimated from the data. The first step is to use a parametric approximation to estimate the law of motion of the disturbance w_t. The second step is to calculate the parameters for the optimal bandwidth parameter from the estimated law of motion. In his Monte Carlo simulations, Andrew uses a AR(1) parameterization for each term of the disturbance. This seems to work well in the models he considers. More recently, Newey and West (1992) provided an alternative method to choose the bandwidth parameter.

Monte Carlo experiments by Newey and West (1992) show the choice of kernel is less important than the choice of the bandwidth parameter for the purpose of more accurate inference. The Bartlett kernel recommended by Newey and West (1987) has been used by many applied researchers. Andrews (1991) recommends quadratic spectral (QS) kernel that has certain asymptotic optimal properties.

Andrews and Monahan (1992) propose a VAR prewhitening method to estimate Ω. Their Monte Carlo experiments, in the context of inference in systems with stationary variables, show that the VAR prewhitening improves small sample properties of estimators of Ω substantially. The intuition behind this is that the estimators of the form (3.18) only take care of MA components of w_t and cannot handle the AR components well in small samples.

Park and Ogaki (1990) extend the VAR prewhitening method to estimation of Γ, so that it can be applied to cointegrating regressions. The first step in the VAR prewhitening method is to run a VAR:

(3.19) $$w_t = A_1 w_{t-1} + A_2 w_{t-2} + \dots + A_k w_{t-k} + e_t.$$

Note that the model (3.19) need not be a true model in any sense. Then the estimated VAR is used to form an estimate e_t and estimators of the form (3.17) and (3.18) are applied to the estimated e_t to estimate the long-run variance of e_t, Ω^* and the paramter Γ for e_t, Γ^*. The estimator based on the QS kernel with the automatic bandwidth parameter can be used to e_t for example. Then the sample counterpart of the formulas

(3.20) $$\Omega = [I - \sum_{i=1}^{k} A]^{-1}\Omega^*[I - \sum_{i=1}^{k} A']^{-1}$$

(3.21) $$\Gamma = \Phi(0) + [I-\sum_{i=1}^{k} A]^{-1}(\Gamma^*-E(e_t e_t'))(I-\sum_{i=1}^{k} A')^{-1} + [I-\sum_{i=1}^{k} A]^{-1}\sum_{j=0}^{k-1}\sum_{i=j+1}^{k} A_i\Phi(-i)$$

are used to form estimates of Ω and Γ.[9]

Monte Carlo experiments in Park and Ogaki (1990) show that the VAR prewhitening improves small sample properties of CCR estimators substantially.

3.3. Tests for the Null of No Cointegration

Many tests for cointegration apply unit root tests to the residuals of a cointegrating regression. When tests for the null hypothesis of unit root nonstatioarity are applied to residuals, the null of no cointegration is

[9]See Park and Ogaki (1992) for a derivation of (3.21).

tested against the alternative of cointegration. It should be noted that the asymptotic distributions of these tests generally depend on the number of the variables in the cointegrating regression.

Engle and Granger's (1987) augmented Dickey-Fuller (ADF) test applies the Said-Dickey test to the residual from cointegrating regressions. Asymptotic properties of the ADF test is studied in Phillips and Ouliaris (1990). These authors and MacKinnon (1990) tabulate critical values from Monte Carlo simulations. Note that these critical values assume the OLS is used for the cointegrating regression, so that the efficient estimation methods discussed in Section 3.2 above shoud not be used for this test. Just as the Said-Dickey test, the ADF test may be sensitive to the choice of the order of the order of AR.

Phillips and Ouliaris also study asymptotic properties of tests for cointegration obtained by applying the Phillips-Perron test to OLS cointegrating regression residuals. Asymptotic critical values are reported by Phillips and Ouliaris. This test requires an estimate of the long run variance of the residual.

Park's (1990a) I(p,q) test basically applies his J(p,q) test to OLS cointegrating regression residuals. This test was originally developed by Park, Ouliaris, and Choi (1988). The I(p,q) test is computed by adding spurious time trends as additional regressors in the cointegrating regression:

$$y_t = \sum_{\tau=0}^{p} \mu_\tau t^\tau + \sum_{\tau=p+1}^{q} \mu_\tau t^\tau + \gamma' X_t + \varepsilon_t. \tag{3.22}$$

Here, time polynomials up to the order of p represent maintained trends, while higher order time polynomials are spurious trends. Part of Park, Ourliaris, and Choi's table of critical values for I(p,q) tests are reproduced here in Table 2. This test has an advantage over ADF and Phillips-Ouliaris tests in that neither the order of AR or the bandwidth parameter needs to be chosen.

(Table 2 around here)

3.4. Tests for the Null of Cointegration

When an economic model implies cointegration, it is often more appealing to test for the null of cointegration, so that an econometrician can control the probability of rejecting a valid economic model. Phillips and Ouliaris (1990) discussed why it was hard to develop tests for the null of cointegration. More recently, Fukushige, Hatanaka, and Koto (1990), Hansen (1992b), and Kwiatowski, Phillips, Schmidt, and Shin (1992), among others, have developed tests for the null of cointegration.

Park's (1990a) H(p,q) test is computed by applying the CCR to (3.22). Thus, this test essentially applies Park's G(p,q) test to CCR residuals. A similar test was originally developed by Park, Ouriaris, and Choi (1988), where G(p,q) tests were applied to OLS residuals, and their tests have nonstandard distributions. In contrast, Park's H(p,q) tests have asymptotic chi-squre distributions with p-q degrees of freedom. The H(p,q) test is computed by applying CCR to (3.22). Under the alternative of no cointegration, the H(p,q) statistic diverges to infinity because spurious trends try to mimick the stochastic trend left in the residual. Therefore, this test is consistent.

In many applications, it is appropriate to model each variable in the econometric system as first difference sationary with drift. Beacus of drift, each vairable possess a linear deterministic trend as well as a stochastic trend in Section 3.1. In this case, H(1,q) statistics test the null hypothesis of stochastic cointegration. The H(0,1) test can be considered as a test for the deterministic cointegration restriction because the restriction implies that the cointegrating vector that eliminates the stochastic trends also eliminates the liear deterministic trends.

3.5. Tests for the Number of Cointegrating Vectors

Johansen's (1988, 1991) likelihood ratio tests and Stock and Watson's (1988a) tests for common trends are often used to determine the number of cointegrating vectors in a system. These tests take the null hypothesis that a $n\times 1$ vector process Z_t has $r\geq 0$ linear independent cointegrating vectors (or it has n-r common stochastic trends) against the alternative that it has $k>r$ linear independent cointegrating vectors (or it has n-k common stochastic trends). Hence if $r=0$, these statistics test the null hypothesis of no cointegration against the alternative of cointegration.

Podivinsky's (1990) Monte Carlo results suggest that there can be severe size distortion problem with Johansen's tests when the sample size is small. For example, when there is no cointegrating vector in the data generation process and when asymptotic critical values are used, Podivinsky finds a tendency for the test with the null hypothesis of $r=0$ to overreject and the test with the null hypothesis of $r\leq 1$ to underreject.

3.6. How Should an Estimation Method be Chosen?

There exist many estimation and testing methods for cointegration. It is advisable for an applied researcher to try at least two methods and check sensitivity of empirical results. When the researcher to chooses a main method to be used, the following considerations naturally come to mind.

3.6.1. Are Short-Run Dynamics of Interest?

If, in addition to coinetrating vectors, the short-run dynamics are of interest, then it seems (at least conceptually) natural to estimate short-run dynamics and cointegrating vectors simultaneously. For example, this can be done by applying Johansen's ML method to estimate an error correction model.

On the other hand, the researcher is often interested in the cointegrating vector but not in short-run dynamics (see, e.g., Atkeson and Ogaki (1991), Clarida (1993ab), and Ogaki (1992a)). In such cases, it is desirable to avoid making unnecessary assumptions about short-run dynamics. An estimation method that uses a nonparametric method to estimate long-run covariance parameters such as CCR is natural in these circumstances.

3.6.2. The Number of the Cointegrating Vectors

In some empirical applications, the researcher may have many economic variables and may not have any guidance from economic models about which variables may be cointegrated. In such applications, tests for the number of cointegrating vectors are useful. It should be noted, however, that these tests may not have very good small sample properties because of the near observational equivalence problem discussed in Section 2.4. For this reason, it is desirable to use economic models to give some a priori information about which variables should be cointegrated.

In some applications, an economic model implies that there exist two or more linearly independnet cointegrating vectors. In this case of multiple cointegrating vectors in a cointegrating regression, neither OLS nor CCR can be used to identify cointegrating vectors. Tests for the null of cointegration based on CCR discussed above also assume that there is only one cointegrating vector and hence cannot be used. However, it is sometimes possible to use a priori information from economic models to handle multiple cointegrating vectors with the CCR methodology.[10] Johansen's ML method has an advantage that it allows multiple cointegrating vectors. However, as pointed out by Park (1991) and Pagan (1992) among others, cointegrating vectors may not be identified even by the Johansen's ML method.

3.6.1. Small Sample Properties

It is known that Johansen's ML estimates and test results can be very sensitive to the choice of the order of autoregression in empirical applications (see, e.g., Stock and Watson (1993)). Therefore, it is important to check sensitivity of empirical results with respect to the order of autoregression when Johansen's method is used. This may be related to the fact that Johansen's estimator for a normlized cointegrating vector has very large mean square error when the sample size is small (see Park and Ogaki (1991b)). Gonzalo (1991) also reports this even though he emphasizes that Johansen's estimator has good small sample properties when the sample size is increased. Podvinsky's (1990) result that Johansen's likelihood ratio tests have severe size distortion problems in some circumsances discussed in Section 3.5 may be to these observations.

Park and Ogaki (1991b) find that the CCR estimator typically has

---

[10] See Kakkar and Ogaki (1993) for an example of an empirical application.

smaller mean square errors than Johansen's ML estimator when the prewhitening method is used. Han and Ogaki (1991) find that Park's tests for the null of cointegration have reasonable small sample properties.

To improve small sample properties of CCR estimators, iterations on the estimation of the long-run covariance parameters are recommended. In empirical applications of CCR, OLS is typically used as an initial estimator. Since OLS coincides with CCR when there is no correlation between the disturbance term and the first difference of the regressors at all leads and lags, the initial OLS may be called the first stage CCR. The second stage CCR is obtained from the long-run covariace parameters calculated from the first stage CCR estimates. The third stage CCR is obtained from the long-run covariance parameters calculated from the second stage CCR estimates, and so on. Park and Ogaki (1991b) report that the small sample properties of the third stage CCR estimator is typically better than those of the second stage CCR estimator. On the other hand, the fourth stage CCR estimator sometimes had significantly larger mean square error. For Park's tests for the null of cointegration to be consistent, it is necessary to bound both the eigenvalues of the VAR prewhitening coefficient matrices and the bandwidth parameter estimate. For example, while using the first order VAR for prewhitening, Han and Ogaki bounds the singular values of the VAR coefficient matrix by 0.99 and the bandwidth parameter by the square root of the sample size. When the variables are cointegrated, the CCR estimators have better small sample properties without these bounds. Because of this, Han and Ogaki recommend reporting the third stage CCR estimates without the bounds imposed and the fourth stage CCR test results with the bounds imposed.

4. Generalized Method of Moments and Unit Roots

When difference stationary variables are involved in the econometric system, standard econometric methods that assume stationarity are not applicable because of spurious regresssion problems. Hence econometricians detrend data by taking growth rates of variables, for example. However, by detrending data, the econmetrician loses the information contained in stochastic and deterministic trends. It is thus natural to seek a method to combine standard econometric methods and cointegrating regressions. Estimating an error correction representation explained in Section 3.2.1 is an example of such a method in vector autoregressions. Let us now consider this problem in the context of Hansen's (1982) Generalized Method of Moments (GMM) estimation. This is particularly useful because many estimators can be considered special cases of GMM.[11]

Let $\{X_t\!:\!t=1,2,...\}$ be a collection of random vectors X_t's, β_0 be a p-dimensional vector of the parameters to be estimated, and $f(X_t,\beta)$ a q-dimensional vector of functions (q≥p). We refer to $u_t=f(X_t,\beta_0)$ as the disturbance of GMM. Consider the (unconditional) moment restrictions

$$E(f(X_t,\ \beta_0)) = 0. \tag{4.1}$$

Suppose that a law of large numbers can be applied to $f(X_t,\beta)$ for all admissible β, so that the sample mean of $f(X_t,\beta)$ converges to its population mean:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f(X_t,\ \beta) = E(f(X_t,\ \beta)) \tag{4.2}$$

with probability one (or in other words, almost surely). The basic idea of

---

[11]See Ogaki (1993) for a survey of GMM.

GMM estimation is to mimic the moment restrictions (2.1) by minimizing a quadratic form of the sample means

$$(4.3)\qquad J_T(\beta) = \{\frac{1}{T}\sum_{t=1}^{T} f(X_t,\ \beta)\}' W_T\{\frac{1}{T}\sum_{t=1}^{T} f(X_t,\ \beta)\}$$

with respect to β; where W_T is a positive semidefinite matrix, which satisfies

$$(4.4)\qquad \lim_{T\to\infty} W_T = W_0.$$

with probability one for a positive definite matrix W_0. The matrices W_T and W_0 are both referred to as the distance or weighting matrix. The GMM estimator, β_T, is the solution of the minimization problem (4.3). Under fairly general regularity conditions, the GMM estimator β_T is a consistent estimator for arbitrary distance matrices.

Let Ω be the long run variance of the GMM disturbance u_t. Then the optimal choice of the distance matrix is $W_0=\Omega^{-1}$, in the sense that this distance matrix yields the most efficient GMM estimator. Under suitable regularity conditions, $\sqrt{T}(\beta_T-\beta_0)$ approximately has a normal distribution with mean zero and the covariance matrix $\{D'\Omega^{-1}D\}^{-1}$ in large samples, where $D=E(\partial f(X_t,\beta)/\partial\beta')$. Moreover, with this choice of the distance matrix, $TJ_T(\beta_T)$, has an asymptotic chi-square distribution with q-p degrees of freedom. This test is sometimes called Hansen's J test and is used to test if the moment condtions (4.1) are satisfied by the data.

The asymptotic theory of GMM does not make strong distributional assumptions, such as that the varibles are normaly distributed. However, Hansen assumes that X_t is stationary. Hence if variables are difference stationary, the econometrician needs to transform the variables to induce stationarity. One such transformation is to take the first difference of a variable, or to take the growth rate of the variables if the log of the variable is difference stationary. But it may not be possible to take growth rates of all variables for some functions in $f(X_t,\beta)$ while retaining moment condtions. In such cases, it may be possible to use cointegrating relationships to induce staitonarity by taking linear combinations of variables. In empirical applications of Eichenbaum and Hansen (1990) and Eichenbaum, Hansen, and Singleton (1988), their economic models imply some variables are cointegrated with a known cointegrating vector. They use this coinegration relationship to induce stationarity for the equations involving the first order condition that equate the relative price and the marginal rate of substitution.

In Cooley and Ogaki (1991), their economic model implies a cointegration relationship, but the cointegrating vector is not known. They employ a two-step procedure. In the first step, they estimate the cointegrating vector, using a cointegrating regression. In the second step, they plug in estimates from the first step into GMM functions $f(X_t,\beta)$. This two step procedure is similar to Engle and Granger's two step procedure for error correction model discussed in Section 3.2.1 above. Asymptotic distributions of GMM estimators in the second step are not affected by the first step estimation because cointegrating regression estimators converge at a faster rate than $\sqrt{T}$.

5. Concluding Remarks

This paper has surveyed the recent leterature on unit roots. An important related issue left out from this survey is a model of breaks in determinstic trends proposed by Perron (1989). Macroeconomic time series may be modeled as stationary fluctuations around determinstic trends that change over time. Even in this case, an idea similar to cointegration can be applied if economic variables have common deterministic trends. Because no interated variables are involved, such relationships may be called cotrending as in Chapman and Ogaki (1993). Under certain conditions, cotrending relationships can be used to estimate cotrending vectors and test economic models.

References

Andrews, D.W.K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica* 59, 817-858.

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

Atkeson, A., and M. Ogaki (1991): "Wealth-Varying Intertemporal Elasticity of Substitution: Evidence from Panel and Aggregate Data." Rochester Center for Economic Research Working Paper No.303. Rochester, New York: University of Rochester.

Beveridge, S. and C.R. Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Comonents with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economy* 7, 151-74.

Bierens, H.J. and S. Guo (1993): "Testing Stationarity and Trend Stationarity Against the Unit Root Hypothesis," *Econometric Review*, forthcoming.

Blough, S.R. (1988): "On the Imposisibility of Testing for Unit Roots and Cointegration in Finite Samples. Working Paper No. 211, Johns Hopkins University.

Campbell, J.Y. and P. Perron (1991): "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots" in *NBER Macroeconomics Annual 1991*, ed. by O.J. Blanchard and S. Fishcer, Cambridge, MA: MIT Press.

Chapman, D.A. and M. Ogaki (1993): "Cotrending and the Stationarity of the Real Interest Rate," *Economics Letters*, forthcoming.

Choi, I. and By.C. Ahn (1993): "Testing the Null of Stationarity for Multiple Time Series," manuscript, Ohio State University.

Christiano, L.J., and M. Eichenbaum (1990): "Unit Roots in Real GNP - Do We Know and Do We Care?" *Carnegie-Rochester Conference on Public Policy* 32, 7-62.

Clarida, R. (1993a): "Cointegration, Aggregate Consumption, and the Demand for Imports: A Structural Econometric Investigation," *American Economic Review*, forthcoming.

——— (1993b): "Permanent Income, Import Prices, and the Demand for Imported Consumer Durables: A Structural Econometric Investigation," manuscript, Columbia University.

Cochrane, J.H. (1988): "How Big is the Random Walk in GNP?" *Journal of Political Economy*, 96, 893-920.

Cooley, T.F., and M. Ogaki (1991): "A Time Series Analysis of Real Wages, Consumption, and Asset Returns under Optimal Labor Contracting: A Cointegration-Euler Equation Approach." Rochester Center for Economic Research Working Paper No.285, Rochester, New York: University of Rochester.

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

—— (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica* 49, 1057-1072.

Durlauf, S.N., and P.C.B. Phillips (1988): "Trends versus Random Walks in Time Series Analysis," *Econometrica* 56, 1333-1354.

Elliott, G. T.J. Rothenburg, and J.H. Stock (1992): "Efficient Tests for an Autoregressive Unit Root," manuscript, Harvard University.

Engle, R.F., and C.W.J. Granger (1987): "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251-276.

Fukushige, M. M. Hatanaka, and Y. Koto (1990): "A Test on the Attainability of Equilibrium and Its Applications to Interanational Capital Markets," invited paper at the 6th World Congress of Econometric Society August 1990.

Gonzalo, J. (1989): "Comparison of Five Alternative Methods of Estimating Long Run Equilibrium Relationships," UCSD Working Paper No. 89-55, San Diego: University of California at San Diego.

Granger, C.W.J. and P. Newbold (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics* 2, 111-120.

Gregory, A. W., A. Pagan, and G.W. Smith (1992): "Estimating Linear Quadratic Models with Integrated Processes," manuscript, Queens University and Australian National University.

Han, H., and M. Ogaki (1991): "Consumption, Income, and Cointegration: Further Analysis." Rochester Center for Economic Research Working Paper No.305. Rochester, New York: University of Rochester.

Hansen, B. (1992a): "Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends," *Journal of Econometrics*, 53, 87-121.

—— (1992b): "Tests for Parameter Instability in Regressions with I(1) Processes," *Journal of Business & Economic Statistics* 10, 321-335.

—— (1993): "Testing for Unit Roots Using Covariates," manuscript, University of Rochester.

Hansen, L.P., and K.J. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269-1286.

Johansen, S. (1988): "Statistical Analysis of Cointegrating Vectors." *Journal of Economic Dynamics and Control* 12, 231-254.

—— (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* 59, 1551-80.

Kakkar, V. and M. Ogaki (1993): "Real Exchange Rates and Nontradables," manuscript, University of Rochester.

Kahn, J.A. and Ogaki, M. (1990): "A Chi-Square Test for a Unit Root," *Economics Letters* 34, 37-42.

—— (1991): "A Consistent Test for the Null of Stationarity against the Alternative of a Unit Root." *Economics Letters* 39, 7-11.

King, R.G., C.I. Plosser, J. H. Stock, and M.W. Watson (1991): "Stochastic Trends and Economic Fluctuations." *American Economic Review* 81, 819-40.

Kocherlakota, N. (1990): "On Tests of Representative Consumer Asset Pricing Models, *Journal of Monetary Economics* 26, 285-304.

Kwiatkowski, D. P.C.B. Phillips, P. Schmidt, and Y. Shin (1992): "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root?," *Journal of Econometrics* 54, 159-78.

MacKinnon, J.G. (1990): "Critical Values for Cointegration Tests." in *Long-Run Economic Relationships: Readings in Cointegration* ed. by R.F. Engle and C.W.J. Granger, Oxford: Oxford University Press, 267-276.

Nelson, C. R. and C.I. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* 10, 139-162.

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

Newey, W.K. and K.D. West (1992): "Automatic Lag Selection in Covariance Matrix Estimation," Social Systems Research Insistute Working Paper No. 9220, Madison: University of Wisconsin.

Ogaki, M. (1988): *Learning about Preferences from Time Trends*, unpublished Ph.D. dissertation. Chicago: University of Chicago.

—— (1989): "Information in Deterministic Trends about Preferences." Manuscript. Rochester, New York: University of Rochester.

—— (1990): "Aggregation of Intratemporal Preferences under Complete Markets" Rochester Center for Economic Research Working Paper No. 249, Rochester, New York: University of Rochester.

—— (1992): "Engel's Law and Cointegration," *Journal of Political Economy*, 100, 1027-46.

—— (1993a): "Generalized Method of Moments: Econometric Applications," forthcoming in *Handbook of Statistics: Vol. 11: Econometrics*, ed. by G.S. Maddala, C.R. Rao, and H.D. Vinod. Amsterdam, Elsevier.

—— (1993b): "GMM: A User's Guide," Rochester Center for Economic Research Working Paper No. 348, Rochester, New York: University of Rochester.

—— (1993c): "CCR: A User's Guide," Rochester Center for Economic Research Working Paper No. 349, Rochester, New York: University of Rochester.

Ogaki, M. and J. Y. Park (1989): "A Cointegration Approach to Estimating Preference Parameters," Rochester Center for Economic Research Working Paper No. 209, Rochester, New York: University of Rochester.

Pagan, A. R. "Three Econometric Methodologies: An Update," manuscript, Australian National University.

Park, J.Y. (1989): "On the Joint Test of a Unit Root and Time Trend," manuscript, Cornell University.

—— (1990a): "Testing for Unit Roots and Cointegration by Variable Addition," *Advances in Econometrics* 8, 107-133.

—— (1990b): "Maximum Likelihood Estimation of Simultaneous Cointegrated Models," manuscript, Cornell University.

—— (1992): "Canonical Cointegrating Regressions," *Econometrica* 60, 119-43.

Park, J.Y., and Choi, B. (1988): "A New Approach to Testing for a Unit Root," CAE Working Paper No. 88-23, Cornell University.

Park, J.Y., and Ogaki, M. (1991a): "Seemingly Unrelated Canonical Cointegrating Regressions," Rochester Center for Economic Research Working Paper No. 280, Rochester, New York: University of Rochester.

—— (1991b): "Inference in Cointegrated Models Using VAR Prewhitening to Estimate Shortrun Dynamics." Rochester Center for Economic Research Working Paper No.281, Rochester, New York: University of Rochester.

Park, J.Y., S. Ouliaris, and B. Choi (1988): "Spurious Regressions and Tests for Cointegration." CAE Working Paper No. 88-07, Cornell University.

Perron, P. and J.Y. Campbell (1992): "A Note on Johansen's Cointegration Procedure When Trends Are Present," manuscript, Princeton University.

Phillips, P.C.B. (1986): "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics* 33, 311-340.

—— (1987): "Time Series Regression with a Unit Root," *Econometrica* 55, 277-301.

—— (1991): "Optimal Inference in Cointegrated Systems," *Econometrica* 59, 283-306.

Phillips, P.C.B. and S.N. Durlauf (1986): "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies* 53, 473-96.

Phillips, P.C.B. and B.E. Hansen (1990): "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies* (1990) 57, 99-125.

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

Phillips, P.C.B., and J.Y. Park (1988): "On the Formulation of Wald Tests of Nonlinear Restrictions," *Econometrica* 56, 1065-1083.

Phillips, P. C. B. and P. Perron (1988): "Testing for a Unit Root in Time Series Regression," *Biometrica*, 75, 335-46.

Podivinsky, J.M. (1990): "Testing Misspecified Cointegrating Relationships," manuscript, University of Southampton.

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

Saikkonen, P. (1991): "Estimation and Testing of Cointegrated Systems by an Autoregressive Approximation," *Econometric Theory* 8, 1-27.

Schwert, G.W. (1989): "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business & Economic Statistics* 7, 147-159.

Stock, J.H. (1987): "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica* 55, 1035-1056.

Stock, J.H. and M.W. Watson (1988a): "Testing for Common Trends," *Journal of the American Statistical Association* 83: 1097-1107.

—— (1988b): "Variable Trends in Economic Time Series," *Journal of Economic Perspectives* 2: 147-174.

—— (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica* 61: 783-820.

Watson, M.W. (1992): "Vector Autoregressions and Cointegration," forthcoming, *Handbook of Econometrics, Volume IV*. Amsterdam: Elsevier.

West, K.D. (1988): "Asymptotic Normality, When Regressors Have a Unit Root," *Econometrica* 56, 1397-1417.

TABLE 1

Critical Values of Park's J(p,q) Tests for the Null of Different Stationarity

Size	0.010	0.025	0.050	0.100
J(0,3)	0.1118	0.2072	0.3385	0.5773
J(1,5)	0.1228	0.1977	0.2950	0.4520

Note: These critical values are from Park and Choi (1988).

TABLE 2

Critical Values of Park's I(p,q) Tests for Null of No Cointegration

Number of Regressors	Size	I(0,3)	I(1,5)
1	0.01	0.06864	0.10269
	0.05	0.23286	0.25064
	0.10	0.39897	0.49845
2	0.01	0.05520	0.00819
	0.05	0.17539	0.21040
	0.10	0.29622	0.32251

Note: These critical values are from Park, Ourliaris, and Choi (1988).