Chapter 14

COINTEGRATING AND SPURIOUS REGRESSIONS

This chapter reviews properties of regression estimators and test statistics based on the estimators when the regressors and regressant are difference stationary. 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). Let $\mathbf{z}_t$ be a $n \times 1$ vector of difference stationary random variables with $\Delta\mathbf{z}_t$ being stationary. If there exists a nonzero vector of real numbers $\boldsymbol{\beta}$ such that $\boldsymbol{\beta}'\mathbf{z}_t$ is stationary, then $\mathbf{z}_t$ is said to be *cointegrated* with a *cointegrating vector* $\boldsymbol{\beta}$. If $\boldsymbol{\beta}$ is a cointegrating vector, $b\boldsymbol{\beta}$ is also a cointegrating vector for any real number b. There may exist more than one linearly independent cointegrating vector. This chapter covers the case in which there is only one linearly independent cointegrating vector, and the case in which there exists no cointegrating vector. Chapter 16 concerns the case when there are more than one linearly independent cointegrating vectors.

When there is one cointegrating vector, a regression of one variable in $\mathbf{z}_t$ on the others is called a *cointegrating regression*. What is striking about cointegration is that a cointegrating vector that eliminates the stochastic trends can be estimated consistently by a cointegrating regression without using instrumental variables, even when no variables are exogenous.

When there is no cointegrating vector, a regression of one variable in $\mathbf{z}_t$ on the others is called a *spurious regression*. One reason why macroeconomists need to be careful about unit root nonstationary variables is that the standard asymptotic theory for regressions in Chapter 5 can be very misleading when variables in a regression are difference stationary.

In the first section, 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 and the null of cointegration as well as tests for the number of cointegrating vectors are presented. Section 14.6 discusses how cointegration may be combined with standard econometric methods that assume stationarity.

14.1 Definitions

If $\boldsymbol{\beta}$ is a cointegrating vector, $b\boldsymbol{\beta}$ is also a cointegrating vector for any real number b. It is often convenient to normalize one of the elements of $\boldsymbol{\beta}$ by one. Suppose that the first element of $\boldsymbol{\beta}$ is nonzero, then partition $\mathbf{z}_t$ by $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$ and normalize $\boldsymbol{\beta}$ by $\boldsymbol{\beta} = (1, -\mathbf{c}')'$. Here y_t is a difference stationary process, $\mathbf{x}_t$ is a vector difference stationary process, and $\mathbf{c}$ is a normalized cointegrating vector.

For most macroeconomic time series such as aggregate income, consumption, and investment, we observe secular upward trends. A secular upward trend of a time series implies that the expected value of the first difference of the series is positive, which implies that the drift term of the series is positive if the series is difference stationary.

Nonzero drift terms in a system of difference stationary series introduce the deterministic trends in addition to the stochastic trends. Hence the cointegrating vector, which eliminates the stochastic trends, may or may not eliminate the deterministic trends from the system. In order to distinguish these cases, we now introduce the notions of stochastic cointegration and the deterministic cointegration restriction, as defined by Ogaki and Park (1997).[1] Consider a vector difference stationary process $\mathbf{x}_t$ with drift:

$$\mathbf{x}_t - \mathbf{x}_{t-1} = \boldsymbol{\mu}_x + \mathbf{v}_t \tag{14.1}$$

for $t \geq 1$ where $\boldsymbol{\mu}_x$ is an $(n-1)$-dimensional vector of real numbers and $\mathbf{v}_t$ is stationary with mean zero. Recursive substitution in (14.1) yields

$$\mathbf{x}_t = \boldsymbol{\mu}_x t + \mathbf{x}_t^0 \tag{14.2}$$

where $\mathbf{x}_t^0$ is difference stationary without drift. Relation (14.2) decomposes the difference stationary process $\mathbf{x}_t$ into deterministic trends arising from drift $\boldsymbol{\mu}_x$ and the difference stationary process without drift, $\mathbf{x}_t^0$. 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 (14.2):

$$y_t = \mu_y t + y_t^0. \tag{14.3}$$

Difference stationary processes y_t and $\mathbf{x}_t$ are said to be *stochastically cointegrated*

[1] Ogaki (1988) introduces these notions and calls them the stochastic and deterministic parts of cointegration. West (1988) considers estimation under the deterministic cointegration restriction for the special case of one stochastic trend in the system. Hansen (1992a) and Park (1992) consider the deterministic cointegration restriction under more general cases.

with a *normalized cointegrating vector* $\mathbf{c}$ when there exists an (n-1)-dimensional vector $\mathbf{c}$ such that $y_t - \mathbf{c}'\mathbf{x}_t$ is trend stationary.[2] This property means that stochastic cointegration only requires that stochastic trend components of the series are cointegrated. We may then write $y_t^0 - \mathbf{c}'\mathbf{x}_t^0 = \theta_c + \epsilon_t$, where ϵ_t is stationary with mean zero. Then by (14.2) and (14.3),

$$y_t = \theta_c + m_c t + \mathbf{c}'\mathbf{x}_t + \epsilon_t \tag{14.4}$$

where

$$m_c = \mu_y - \mathbf{c}'\boldsymbol{\mu}_x. \tag{14.5}$$

Suppose that

$$\mu_y = \mathbf{c}'\boldsymbol{\mu}_x \tag{14.6}$$

holds. Then the deterministic cointegration restriction is said to hold. This means that the cointegrating vector that eliminates the stochastic trends also eliminates the deterministic trends. If this restriction is satisfied, then

$$y_t = \theta_c + \mathbf{c}'\mathbf{x}_t + \epsilon_t. \tag{14.7}$$

and $(y_t, \mathbf{x}_t)'$ is cointegrated.

Another way to explain the deterministic cointegration is to use an idea of cotrending. Suppose that a vector $\mathbf{c}^*$ satisfies

$$\mu_y = \mathbf{c}^{*\prime}\boldsymbol{\mu}_x. \tag{14.8}$$

Then $y_t - \mathbf{c}^{*\prime}\mathbf{x}_t$ does not possess any deterministic trend, and y_t and $\mathbf{x}_t$ are *cotrended* with a *normalized cotrending* vector $\mathbf{c}^*$. If $n > 2$ and if one of the components

[2] If $y_t^0 - \mathbf{c}'\mathbf{x}_t^0$ is stationary rather than trend stationary, y_t and $\mathbf{x}_t$ are said to be cointegrated.

of $\boldsymbol{\mu}_x$ is nonzero, there are infinitely many cotrending vectors. Consider an extra restriction that the normalized cointegrating vector $\mathbf{c}$ 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, Equation (14.7) holds and $(y_t, \mathbf{x}_t')'$ is cointegrated.

14.2 Exact Finite Sample Properties of Regression Estimators

This section studies exact finite sample properties of cointegrating and spurious regression estimators. In the literature on unit root econometrics, asymptotic theory and the method of Monte Carlo studies have been typically used. However, the conditional Gauss-Markov theorem in Chapter 5 can be applied to study exact finite sample properties as in Ogaki and Choi (2001).

Consider a regression of the form

$$y_t = \mathbf{h}'\mathbf{d}_t + \mathbf{c}'\mathbf{x}_t + \epsilon_t. \tag{14.9}$$

where $\mathbf{d}_t$ is a function of time, t. For example, $\mathbf{d}_t = (1, t)'$ as in (14.4) or $\mathbf{d}_t = 1$ as in (14.7). If ϵ_t is stationary for some $\mathbf{c}$, (14.9) is a cointegrating regression. If ϵ_t is difference stationary for any $\mathbf{c}$, then (14.9) is a spurious regression.

14.2.1 Spurious Regressions

Suppose that y_t is a random walk and x_t is a random walk that is independent of y_t. Granger and Newbold (1974) find that the standard Wald test statistic for the hypothesis that the coefficient on x_t is zero tends 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) show that the Wald test statistic 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 tends to spuriously conclude that two independent difference stationary variables are related.

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

The Gauss Markov theorem provides us with a tool to understand exact small sample properties of estimators and test statistics of spurious regressions. The asymptotic theories of Phillips (1987, 1998) have been used to understand the spurious regression problem, but have not been used to provide a solution to the problem. The Gauss Markov theorem indicates a simple solution to the problem.

Let y_t be a random walk that is generated from

$$\Delta y_t = \epsilon_t \tag{14.10}$$

with an initial random variable y_0 and a white noise ϵ_t that is conditionally homoskedastic. Let x_t be another random walk that is generated from

$$\Delta x_t = v_t \tag{14.11}$$

with an initial random variable x_0 and a white noise v_t that is conditionally homoskedastic. We assume that $\{\epsilon_t\}_{t=1}^T$ and y_0 are independent, and that they are independent from $\{v_t\}_{t=1}^T$ and x_0, so that x_t and y_t are independent random walks. Let $\mathbf{y} = \{y_t\}_{t=1}^T$, $\mathbf{X} = \{x_t\}_{t=1}^T$ and $\mathbf{e} = \{e_t\}_{t=1}^T$ where $\Delta e_t = \epsilon_t$, and consider the OLS estimator for $\mathbf{y} = \mathbf{X}b_0 + \mathbf{e}$. Then the true value of the regression coefficient is zero: $b_0 = 0$.

Let I_x be the information set generated from y_0 and $\mathbf{X}$. Assumptions 5.1, 5.2, and 5.4 of the strict version of the theorem in Chapter 5 hold for the spurious regression. However, Assumption 5.3 is violated because

$$E(\mathbf{ee}'|\mathrm{I}_x) = \sigma^2\mathbf{\Phi} \tag{14.12}$$

where $\sigma^2 = E(\epsilon_t^2)$, and

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 2 & \cdots & 2 & 2 \\ 1 & 2 & 3 & \cdots & 3 & 3 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 1 & 2 & 3 & \cdots & T-1 & T-1 \\ 1 & 2 & 3 & \cdots & T-1 & T \end{bmatrix}. \tag{14.13}$$

Thus the spurious regression violates Assumption 5.3, but not the other assumptions. The OLS estimator is still unbiased. One can apply a GLS correction and obtain a more efficient estimator.

When Assumption 5.5 is made, by applying GLS to the spurious regression, we can solve the spurious regression problem: we can obtain the exact (unconditional) t distribution for the usual t statistic.

We now consider spurious regressions of the form (14.9) which do not satisfy the strict exogeneity assumption. For this purpose, we consider a particular data generating process that leads to a spurious regression.

Let e_t and $\mathbf{z}_t$ be two time series of dimensions 1 and k, respectively, that are generated from

$$\Delta e_t = \epsilon_t, \qquad t = 1, 2, 3, \cdots . \tag{14.14}$$

$$\Delta \mathbf{z}_t = \boldsymbol{\mu} + \mathbf{v}_t, \qquad t = -q, \cdots, -1, 0, 1 \cdots \tag{14.15}$$

where $(\epsilon_t, \mathbf{v}_t')'$ is a covariance stationary series, $\boldsymbol{\mu}$ is a k-dimensional vector of real numbers, $e_0 = 0$, and $\mathbf{z}_{-q}$ is a given random vector. We assume that the long-run covariance matrix of $\mathbf{v}_t$,

$$\boldsymbol{\Omega} = \lim_{j\to\infty} \sum_{-j}^{j} E(\mathbf{v}_t \mathbf{v}_{t-j}') \tag{14.16}$$

is nonsingular. We assume that $\mathbf{z}_t$ is strictly exogenous with respect to ϵ_t.

An implication of the strict exogeneity assumption is that e_t and $\mathbf{z}_t$ are not cointegrated: that is, there is no nonzero vector $\boldsymbol{\beta}$ such that $e_t - \boldsymbol{\beta}'\mathbf{z}_t$ is stationary. This property results because the assumption implies that

$$\lim_{j\to\infty} \sum_{-j}^{j} E(\epsilon_t \mathbf{v}_{t-j}') = 0. \tag{14.17}$$

Consider a series y_t that is generated from

$$y_t = \mathbf{h}'\mathbf{d}_t + \mathbf{c}'\mathbf{z}_t + \boldsymbol{\gamma}(L^{-1})\Delta\mathbf{z}_t + \boldsymbol{\eta}(L)\Delta\mathbf{z}_t + e_t, \quad t = 1, 2, 3, \cdots , \tag{14.18}$$

where $\boldsymbol{\gamma}(L^{-1}) = \boldsymbol{\gamma}_1 L^{-1} + \cdots + \boldsymbol{\gamma}_p L^{-p}$, $\boldsymbol{\eta}(L) = \boldsymbol{\eta}_0 + \boldsymbol{\eta}_1 L + \cdots + \boldsymbol{\eta}_q L^q$ and $\mathbf{d}_t$ is a vector of deterministic variables that is $(1, t)'$ or 1 for example. Here $\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_p$, and $\boldsymbol{\eta}_0, \cdots, \boldsymbol{\eta}_q$ are $1 \times k$ vectors, and we assume that at least one of them is nonzero.

Under these assumptions, consider a regression of y_t onto $\mathbf{d}_t$ and $\mathbf{z}_t$:

$$y_t = \mathbf{h}^{*\prime}\mathbf{d}_t + \mathbf{c}^{*\prime}\mathbf{z}_t + e_t^* \tag{14.19}$$

This regression is a spurious regression: that is, for any vector $\mathbf{c}^*$, e_t^* is unit root nonstationary. To see this property, assume that e_t^* is stationary for a vector $\mathbf{c}^*$. Then (14.18) implies that $e_t - (\mathbf{c}^* - \mathbf{c})'\mathbf{z}_t$ is stationary. It follows that a cointegrating relationship exists between e_t and $\mathbf{z}_t$ contradicting the strict exogeneity assumption.

Given that e_t in (14.18) satisfies the exogeneity condition, $\mathbf{c}$ can be considered the true value of the spurious regression coefficient $\mathbf{c}^*$ in (14.19). With this interpretation, one problem with (14.19) is that the strict exogeneity assumption is violated.

Let $\mathbf{X}$ be a matrix whose t-th row is given by $(\mathbf{d}_t', \mathbf{z}_t', \Delta\mathbf{z}_{t+p}', \Delta\mathbf{z}_{t+p-1}', \cdots, \Delta\mathbf{z}_t', \Delta\mathbf{z}_{t-1}', \cdots, \Delta\mathbf{z}_{t-q}')$, $\mathbf{y} = \{y_t\}_{t=1}^T$, and $\mathbf{e}^* = \{e_t^*\}_{t=1}^T$. When

$$E(\mathbf{e}^*\mathbf{e}^{*\prime}|\mathbf{X}) = \sigma^2\mathbf{\Psi} \tag{14.20}$$

with a known matrix $\mathbf{\Psi}$ and a possibly unknown number σ, then the GLS can be applied to (14.18). If e_t is a random walk, then with $\mathbf{\Phi}$ given by (14.13) and $\sigma^2 = E(\epsilon_t^2)$. Just as in the strict exogenous case, the finite sample properties of the GLS estimators and test statistics based on GLS can be analyzed.

The GLS correction is basically the same as taking first differences for the case of strictly exogenous regressors. The GLS correction, however, can be useful in applications for which the strict exogeneity assumption is violated.

14.2.2 Cointegrating Regressions

Let e_t and $\mathbf{z}_t$ be two time series of dimensions 1 and k, respectively. We assume that $\mathbf{z}_t$ is generated from (14.15), where $(e_t, \mathbf{v}_t')'$ is a covariance stationary series, $\boldsymbol{\mu}$ is a k-dimensional vector of real numbers, $e_0 = 0$, and $\mathbf{z}_{-q}$ is a given random vector. We assume that the long-run covariance matrix of $\mathbf{v}_t$, $\mathbf{\Omega} = \lim_{j\to\infty}\sum_{-j}^{j} E(\mathbf{v}_t\mathbf{v}_{t-j}')$ is nonsingular. We assume that $\mathbf{z}_t$ is strictly exogenous with respect to e_t.

Consider a series y_t that is generated from

$$y_t = \mathbf{h}'\mathbf{d}_t + \mathbf{c}'\mathbf{z}_t + \boldsymbol{\gamma}(L^{-1})\Delta\mathbf{z}_t + \boldsymbol{\eta}(L)\Delta\mathbf{z}_t + e_t, \quad t = 1, 2, 3, \cdots, \tag{14.21}$$

where $\boldsymbol{\gamma}(L^{-1})$, $\boldsymbol{\eta}(L)$ and $\mathbf{d}_t$ are defined in (14.18).

Under these assumptions, consider a regression of y_t onto $\mathbf{d}_t$ and $\mathbf{z}_t$:

$$y_t = \mathbf{h}^{*\prime}\mathbf{d}_t + \mathbf{c}^{*\prime}\mathbf{z}_t + e_t^* \tag{14.22}$$

This regression is a cointegrating regression. With an appropriate choice of $\mathbf{h}^*$ and $\mathbf{c}^* = \mathbf{c}$, e_t^* is stationary. However, since the strict exogeneity assumption is not satisfied, the OLS estimator for (14.22) is biased.

In contrast, the OLS estimator for (14.21) is unbiased. It is the BLUE if e_t is serially uncorrelated. This is because Assumptions 5.1, 5.2, 5.3 and 5.4 are satisfied, and the conditional Gauss-Markov theorem applies. The OLS estimator for (14.21) is called the dynamic OLS estimator. The GLS estimator for (14.21) is called the dynamic GLS estimator.

14.3 Large Sample Properties

An important feature of the cointegration regression is that the OLS estimator is consistent without any exogeneity assumption (see Phillips and Durlauf, 1986; Stock, 1987). Along with the spurious regression results discussed in the last section, it is another example of the fact that the standard asymptotic theory in Chapter 5 does not apply to regressions in the presence of unit root nonstationary variables. This fact is well known in the literature. On the other hand, the fact that the conditional probability version of the Gauss Markov theorem applies to cointegrating regressions under the assumptions of the theorem has not been emphasized in the literature. In the context of cointegration, an assumption of the theorem requires that $\mathbf{x}_t$ is strictly exogenous.

In most applications, the strict exogeneity assumption is too restrictive. This section discusses econometric methods for when the assumption is violated. The OLS estimator is consistent (see Phillips and Durlauf, 1986; Stock, 1987), but is asymptotically biased. It also has a nonstandard distribution, which makes 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 some of these problems have been developed. Dynamic OLS and GLS estimators introduced in the last section were proposed by Stock and Watson (1993). Phillips and Loretan (1991) and Saikkonen (1991) have proposed similar estimators.

Dynamic OLS and GLS estimators correct the endogeneity problem parametrically. Estimators proposed by Phillips and Hansen (1990) and Park's (1992) Canonical Cointegrating Regressions correct the endogeneity problem nonparametrically. In Chapter 16, we will explain Johansen (1988, 1991) Maximum Likelihood Estimation method.

14.3.1 Canonical Cointegrating Regression

Johansen's maximum likelihood estimation makes a parametric correction for long-run correlation of $\Delta\mathbf{x}_t$ and $\mathbf{e}_t$. Another way to obtain an efficient estimator is to utilize a nonparametric estimate of the long-run covariance parameters. Both Phillips and Hansen (1990) and Park (1992) employ such covariance estimates. Here, attention is confined to Park's Canonical Cointegration Regressions (CCR).

Consider a cointegrated system

$$y_t = \mathbf{h}'\mathbf{d}_t + \mathbf{c}'\mathbf{x}_t + \epsilon_t \tag{14.23}$$

$$\Delta\mathbf{x}_t = \mathbf{v}_t, \tag{14.24}$$

where $\mathbf{d}_t$ is a deterministic term that are usually constants, time trends, or both, y_t and $\mathbf{x}_t$ are difference stationary, and ϵ_t and $\mathbf{v}_t$ are stationary with zero mean. Here y_t is a scalar and $\mathbf{x}_t$ is a $(n-1)\times 1$ random vector. Let

$$\mathbf{w}_t = (\epsilon_t, \mathbf{v}_t')'. \tag{14.25}$$

Define $\mathbf{\Phi}(i) = E(\mathbf{w}_t\mathbf{w}_{t-i}')$, $\mathbf{\Sigma} = \mathbf{\Phi}(0)$, $\mathbf{\Gamma} = \sum_{i=0}^{\infty}\mathbf{\Phi}(i)$, and $\mathbf{\Omega} = \sum_{i=-\infty}^{\infty}\mathbf{\Phi}(i)$. Here $\mathbf{\Omega}$ is the matrix version of (14.16) and is the long run variance (or covariance) matrix of $\mathbf{w}_t$. Partition $\mathbf{\Omega}$ as

$$\mathbf{\Omega} = \begin{bmatrix} \Omega_{11} & \mathbf{\Omega}_{12} \\ \mathbf{\Omega}_{21} & \mathbf{\Omega}_{22} \end{bmatrix} \tag{14.26}$$

where Ω_{11} is a scalar, and $\mathbf{\Omega}_{22}$ is a $(n-1)\times(n-1)$ matrix, and partition $\mathbf{\Gamma}$ conformably. Define

$$\Omega_{11.2} = \Omega_{11} - \mathbf{\Omega}_{12}\mathbf{\Omega}_{22}^{-1}\mathbf{\Omega}_{21} \tag{14.27}$$

and $\mathbf{\Gamma}_2 = (\mathbf{\Gamma}_{12}', \mathbf{\Gamma}_{22}')'$. The CCR procedure assumes that $\mathbf{\Omega}_{22}$ is positive definite, implying that $\mathbf{x}_t$ is not itself cointegrated (see, e.g., Phillips, 1986; Engle and Granger, 1987). This assumption assures that $(1, -\mathbf{c})$ is the unique cointegrating vector (up to a scale factor).[3]

[3]For many applications, it is natural to assume that $\Delta^{-1}\epsilon_t$ is not cointegrated with $\mathbf{x}_t$. This assumption implies that $\Omega_{11.2}$ is positive. Park (1992) calls cointegration between y_t and $\mathbf{x}_t$ 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).

The OLS estimator in (14.23) is super-consistent in that the estimator converges to $\mathbf{c}$ at the rate of T (sample size) even when $\Delta\mathbf{x}_t$ and ϵ_t are correlated. The OLS estimator, however, is not asymptotically efficient. Consider transformations

$$y_t^* = y_t + \boldsymbol{\pi}_y'\mathbf{w}_t \tag{14.28}$$

$$\mathbf{x}_t^* = \mathbf{x}_t + \boldsymbol{\pi}_x'\mathbf{w}_t. \tag{14.29}$$

Since $\mathbf{w}_t$ is stationary, y_t^* and $\mathbf{x}_t^*$ are cointegrated with the same cointegrating vector $(1, -\mathbf{c})$ as y_t and $\mathbf{x}_t$ for any $\boldsymbol{\pi}_y$ and $\boldsymbol{\pi}_x$. The idea of the CCR is to choose $\boldsymbol{\pi}_y$ and $\boldsymbol{\pi}_x$, so that the OLS estimator is asymptotically efficient when y_t^* is regressed on $\mathbf{x}_t^*$.[4] This requires

Masao needs to check this!

$$\boldsymbol{\pi}_y = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Gamma}_2\mathbf{c} + (0, \boldsymbol{\Omega}_{12}\boldsymbol{\Omega}_{22}^{-1})' \tag{14.30}$$

$$\boldsymbol{\pi}_x = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Gamma}_2. \tag{14.31}$$

In practice, long-run covariance parameters in these formulas are estimated, and estimated $\boldsymbol{\pi}_y$ and $\boldsymbol{\pi}_x$ are used to transform y_t and $\mathbf{x}_t$. As long as these parameters are estimated consistently, the resultant CCR estimator 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 econometric system of interest, then asymptotically it is more efficient to apply seemingly unrelated regressions. Park and Ogaki (1991a) develop a method of Seemingly Unrelated Canonical Cointegrating Regressions (SUCCR) for this case. In the SUCCR, transformations of y_t and $\mathbf{x}_t$ that are slightly different from (14.28) and (14.29) are applied

[4] Under general conditions, a sequence of functions $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{w}_t$ converges in distribution to a vector Brownian motion B with covariance matrix $\boldsymbol{\Omega}$. The OLS estimator converges in distribution to ????

in each regression. After transforming the variables, the standard seemingly unrelated regression method is applied to the transformed variables.

14.3.2 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 $\mathbf{\Omega}$ and $\mathbf{\Gamma}$.

In many applications of cointegration, the order of serial correlation is unknown. Let $\mathbf{\Phi}(\tau) = E(\mathbf{w}_t\mathbf{w}'_{t-\tau})$,

$$\mathbf{\Phi}_T(\tau) = \frac{1}{T}\sum_{t=\tau+1}^{T} \hat{\mathbf{w}}_t\hat{\mathbf{w}}'_{t-\tau} \quad \text{for} \quad \tau \geq 0, \tag{14.32}$$

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

$$\mathbf{\Omega}_T = \frac{T}{T-p}\sum_{\tau=-T+1}^{T-1} k(\frac{\tau}{S_T})\mathbf{\Phi}_T(\tau), \tag{14.33}$$

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

$$\mathbf{\Gamma}_T = \frac{T}{T-p}\sum_{\tau=0}^{T-1} k(\frac{\tau}{S_T})\mathbf{\Phi}_T(\tau), \tag{14.34}$$

Park and Ogaki (1991b) extend Andrews and Monahan's (1992) VAR prewhitening method to the estimation of $\mathbf{\Gamma}$ so that it can be applied to cointegrating regressions. The first step in the VAR prewhitening method is to run a VAR:

$$\mathbf{w}_t = \mathbf{A}_1\mathbf{w}_{t-1} + \mathbf{A}_2\mathbf{w}_{t-2} + \cdots + \mathbf{A}_k\mathbf{w}_{t-k} + \mathbf{f}_t. \tag{14.35}$$

Note that the model (14.35) need not be a true model in any sense. Then the estimated VAR is used to form an estimate $\mathbf{f}_t$ and estimators of the form (14.33) and (14.34) are applied to the estimated $\mathbf{f}_t$ to estimate the long-run variance of $\mathbf{f}_t$, $\mathbf{\Omega}^*$ and the parameter $\mathbf{\Gamma}$ for $\mathbf{f}_t$, $\mathbf{\Gamma}^*$. The estimator based on the QS kernel with the automatic bandwidth parameter can be used for $\mathbf{f}_t$ for example. Then the sample counterpart of the formulas

$$\text{(14.36)} \qquad \mathbf{\Omega} = [\mathbf{I} - \sum_{i=1}^{k} \mathbf{A}_i]^{-1} \mathbf{\Omega}^* [\mathbf{I} - \sum_{i=1}^{k} \mathbf{A}'_i]^{-1}$$

$$\text{(14.37)} \qquad \mathbf{\Gamma} = \mathbf{\Phi}(0) + [\mathbf{I} - \sum_{i=1}^{k} \mathbf{A}_i]^{-1} (\mathbf{\Gamma}^* - E(\mathbf{f}_t \mathbf{f}'_t)) [\mathbf{I} - \sum_{i=1}^{k} \mathbf{A}'_i]^{-1} + [\mathbf{I} - \sum_{i=1}^{k} \mathbf{A}_i]^{-1} \sum_{j=0}^{k-1} \sum_{i=j+1}^{k} \mathbf{A}_i \mathbf{\Phi}(-i)$$

are used to form estimates of $\mathbf{\Omega}$ and $\mathbf{\Gamma}$.[5]

Monte Carlo experiments in Park and Ogaki (1991b) show that the VAR prewhitening improves small sample properties of CCR estimators substantially.

14.4 Tests for the Null Hypothesis 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 nonstationarity are applied to residuals, the null of no cointegration is 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. The asymptotic prop-

[5]See Park and Ogaki (1991a) for a derivation of (14.36) and (14.37).

erties of the ADF test are 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 14.3 above should 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 AR.

Phillips and Ouliaris (1990) 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 (1990) 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:

$$y_t = \sum_{\tau=0}^{p} \mu_\tau t^\tau + \sum_{\tau=p+1}^{q} \mu_\tau t^\tau + \mathbf{c}'\mathbf{x}_t + \epsilon_t. \tag{14.38}$$

Here, time polynomials up to the order of p represent maintained trends, while higher order time polynomials are spurious trends. Part of Park, Ouliaris, and Choi's (1988) table of critical values for $I(p,q)$ tests are reproduced here in Table 14.1. This test has an advantage over ADF and Phillips-Ouliaris tests in that neither the order of AR nor the bandwidth parameter needs to be chosen.

Table 14.1: 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, Ouliaris, and Choi (1988).

14.5 Tests for the Null Hypothesis 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 (1994), Hansen (1992b), and Kwiatkowski, Phillips, Schmidt, and Shin (1992), among others, have developed tests for the null of cointegration.

Park's (1990) $H(p,q)$ test is computed by applying the CCR to (14.38). Thus, this test essentially applies Park's $G(p,q)$ test to CCR residuals. A similar test was originally developed by Park, Ouliaris, 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-square distributions with $q-p$ degrees of freedom. Under the alternative of no cointegration, the $H(p,q)$ statistic diverges to infinity because spurious trends try to mimic 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 stationary with drift. Each variable possess a linear deterministic trend as well as a stochastic trend in Section 14.1 because of drift. 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 linear deterministic trends.

14.6 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 regression problems. Hence econometricians detrend data by taking growth rates of variables, for example. However, by detrending data, the econometrician 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 16.4 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 case is particularly useful because many estimators can be considered special cases of GMM.

The asymptotic theory of GMM does not make strong distributional assumptions, such as that the variables are normally distributed. However, Hansen assumes that $\mathbf{x}_t$ 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 $f(\mathbf{x}_t, \mathbf{b}_0)$ while retaining moment conditions. In such cases, it may be possible to use cointegrating relationships to induce stationarity 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 cointegration 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 (1996) and Ogaki and Reinhart (1998a,b) explained in the next chapter, 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, $f(\mathbf{x}_t, \mathbf{b}_0)$. This two step procedure is similar to Engle and Granger's two step procedure for the error correction model discussed in Section 16.4. 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 $\sqrt{T}$.

Appendix

14.A Procedures for Cointegration Tests

14.A.1 Park's CCR and H Test (CCR.EXP)

Park's canonical cointegrating regression (CCR) and $H(p, q)$ test proceed as follows:

(i) Define a regressand (the corresponding variable to be specified is `Y`) and regressors. The latter includes both a vector of deterministic regressors[6] (the corresponding variable to be specified is `X1`) and difference stationary regressors (the corresponding variable to be specified is `X2`).

(ii) Choose the order of the maintained trend (the corresponding variable to be specified is `P`) in the residuals to test for the null hypothesis of cointegration ($H(p, q)$ test). If either each variable exhibits no secular trend or some variables show a secular trend with the deterministic cointegration restriction[7], set `P=0`. On the other hand, when some variables exhibit a time trend without the deterministic cointegration, set `P=1`.

(iii) Select the largest order of additional time polynomials (the corresponding variable to be specified is `Q`). If either each variable exhibits no secular trend or some variables show a secular trend with the deterministic cointegration restriction, set `Q=1`. But when some variables exhibit a time trend without the deterministic cointegration, set `Q=2`. Choose an appropriate `DQ` depending on how many test results you want. We recommend either `DQ=2` or `DQ=3`.

(iv) Determine an appropriate method to estimate the long-run covariance matrix, Ω_T. See chapter 6 for details (the corresponding variables to be specified are `MAXD, ST, BST,` and `MSERHO`. The default of the program is the prewhitened QS kernel with automatic bandwidth selection).

(v) Impose restrictions on the cointegrating vector (the corresponding variable in

[6]It is typically either a constant or a constant and a linear time trend

[7]Typically, the economic model for the application tells us whether or not this restriction is satisfied

the program is `B`), if any (the corresponding variables to be specified are `R` and `RV` matrices).

(vi) Check the statistical evidence about estimates and tests. For CCR estimates, report the third stage result. For the $H(p, q)$ test, report the fourth stage result. For a linear restriction `RB=RV`, report the third stage result when the alternative hypothesis is cointegration with `RB` $\neq$ `RV`, and report the fourth stage result when the alternative hypothesis is no cointegration.

14.A.2 Park's I Test (IPQ.EXP)

Park's I(p,q) test proceeds as follows:

(i) Define a regressand (the corresponding variable to be specified is `Y`) and regressors. The latter includes both a vector of deterministic regressors[8] (the corresponding variable to be specified is `X1`) and difference stationary regressors (the corresponding variable to be specified is `X2`).

(ii) Choose the order of the maintained trend in the regression (the corresponding variable in the program is `P`). If the variable of interest does not exhibit a secular time trend, the maintained hypothesis is that it includes only a constant (set `P=0`). However, if it shows a secular time trend, the maintained hypothesis is that it possesses a linear time trend (set `P=1`).

(iii) Select the largest order of additional time polynomials (the corresponding variable in the program is `Q`) and its range (the corresponding variable in the program is `DQ`) in the regression. If the variable of interest does not exhibit a

[8] It is typically either a constant or a constant and a linear time trend

secular time trend, the maintained hypothesis is that it includes only a constant (set `Q=1`). However, if it shows a secular time trend, the maintained hypothesis is that it possesses a linear time trend (set `Q=2`). Choose an appropriate DQ depending on how many test results you want. We recommend either `DQ=2` or `DQ=3`.

(iv) Impose restrictions on the cointegrating vector (the corresponding variable in the program is `B`), if any (the corresponding variables to be specified are `R` and `RV` matrices).

(v) If $I(p, q)$ is smaller than the appropriate critical value, then reject the null of no cointegration.

14.B Weak Convergence to Stochastic Integral

Before we provide formal theorems of weak convergence to stochastic integral, by using a a cointegrating regression, we show why the FCLT alone (even with the CMT) is not enough to establish the asymptotic properties of the OLS estimator.

Consider the following cointegrating regression:

$$y_t = \beta x_t + u_t,$$

where x_t is an I(1) process and u_t is an I(0) process. The OLS estimator is given by

$$\widehat{\beta} = \frac{\sum_{t=1}^{T} y_t u_t}{\sum_{t=1}^{T} x_t^2},$$

and its sample error can be written as

$$T(\widehat{\beta} - \beta) = \frac{\dfrac{1}{T}\sum_{t=1}^{T} x_t u_t}{\dfrac{1}{T^2}\sum_{t=1}^{T} x_t^2}.$$

For the denominator, by the FCLT (along with the CMT) it can be shown that:

$$\frac{1}{T^2}\sum_{t=1}^{T} x_t^2 \xrightarrow{d} \int_0^1 W^2(r)dr.$$

However, the numerator cannot be analyzed by the FCLT alone. It is evident that the asymptotic distribution of the numerator cannot be established by the FCLT alone, because the numerator is a mixture of I(1) and I(0) random variables. Therefore we need a different too, so-called "weak convergence to the stochastic integral." In below, we present the most general version of the theorem.

Theorem 14.1 *Let $\{U_{nt}, W_{nt}\}$ be a (2×1) stochastic array, let $X_n(r) = \sum_{t=1}^{[nr]} U_{nt}$ and $Y_n(r) = \sum_{t=1}^{[nr]} W_{nt}$, and suppose that $(X_n(r), Y_n(r)) \xrightarrow{d} (B_X(r), B_Y(r))$. Assume $\{U_{nt}\}$ is L_r-bounded and L_2-NED of size -1 on $\{\mathbf{V}_{nt}\}$ with respect to constants $\{c_{nt}^U\}$. If the one of the following assumptions hold:*

1. *$\{W_{nt}, \mathcal{H}_{nt}\}$ is a martingale difference array, where $\mathcal{H}_{nt} = \sigma((W_{nk}, U_{n,k-1}, k \leq t)$, and $E(W_{n,t+1}^2 | \mathcal{H}_{n,t}) \ll (c_{n,t}^W)^2 < \infty$, a.s.*

2. *$W_{nt} = \sum_{k=0}^{\infty} \theta_k V_{1n,t-k}$ where $V_{1nt} \in \mathbf{V}_{nt}$ is a L_r-bounded zero-mean random variable, independent of $V_{n,t'}$ for all $t \neq t'$, and $\sum_{t=0}^{\infty} \sum_{k=t}^{\infty} |\theta_k| < \infty$*

Then,

$$G_n = \sum_{j=1}^{n-1}\left(\sum_{i=1}^{j} U_{n,i}\right) W_{n,j+1} \xrightarrow{d} \int_0^1 B_X(r)\mathrm{d}B_Y(r) + \Lambda_{XY}$$

where $\Lambda_{XY} = \lim_{n\to\infty} \sum_{i=1}^{n-1} \sum_{m=0}^{i-1} E(U_{n,i-m} W_{n,i+1})$

References

Andrews, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59(3), 817–858.

ANDREWS, D. W. K., AND J. C. MONAHAN (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60(4), 953–966.

COOLEY, T. F., AND M. OGAKI (1996): "A Time Series Analysis of Real Wages, Consumption, and Asset Returns," *Journal of Applied Econometrics*, 11(2), 119–134.

DURLAUF, S. N., AND P. C. B. PHILLIPS (1988): "Trends versus Random Walks in Time Series Analysis," *Econometrica*, 56, 1333–1354.

EICHENBAUM, M., AND L. P. HANSEN (1990): "Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data," *Journal of Business and Economic Statistics*, 8, 53–69.

EICHENBAUM, M., L. P. HANSEN, AND K. J. SINGLETON (1988): "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty," *Quarterly Journal of Economics*, 103, 51–78.

ENGLE, R. F., AND C. GRANGER (1987): "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251–276.

FUKUSHIGE, M., M. HATANAKA, AND Y. KOTO (1994): "Testing for the Stationarity and the Stability of Equilibrium," in *Advances in Econometrics, Sixth World Congress*, ed. by C. A. Sims, vol. I, pp. 3–45. Cambridge University Press.

GRANGER, C. W. J., AND P. NEWBOLD (1974): "Spurious Regressions in Econometrics," *Journal of Econometrics*, 2, 111–120.

HANSEN, B. E. (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 Regression with I(1) Processes," *Journal of Business and Economic Statistics*, 10, 321–335.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–1054.

JOHANSEN, S. (1988): "Statistical Analysis of Cointegration 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(6), 1551–1580.

KWIATKOWSKI, D., P. C. B. PHILLIPS, P. SCHMIDT, AND Y. C. 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(1–3), 159–178.

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

OGAKI, M. (1988): "Learning about Preferences from Time Trends," Ph.D. thesis, University of Chicago.

OGAKI, M., AND C.-Y. CHOI (2001): "The Gauss-Markov Theorem and Spurious Regressions," Working Paper No. 01-13, Department of Economics, Ohio State University.

Ogaki, M., and J. Y. Park (1997): "A Cointegration Approach to Estimating Preference Parameters," *Journal of Econometrics*, 82(1), 107–134.

Ogaki, M., and C. M. Reinhart (1998a): "Intertemporal Substitution and Durable Goods: Long-Run Data," *Economics Letters*, 61, 85–90.

——— (1998b): "Measuring Intertemporal Substitution: The Role of Durable Goods," *Journal of Political Economy*, 106, 1078–1098.

Park, J. Y. (1990): "Testing for Unit Roots and Cointegration by Variable Addition," in *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, ed. by T. Fomby, and G. Rhodes, vol. 8, pp. 107–133. JAI Press, Greenwich, CT.

——— (1992): "Canonical Cointegrating Regressions," *Econometrica*, 60(1), 119–143.

Park, J. Y., and M. Ogaki (1991a): "Seemingly Unrelated Canonical Cointegrating Regressions," RCER Working Paper No. 280.

——— (1991b): "VAR Prewhitening to Estimate Short-Run Dynamics: On Improved Method of Inference in Cointegrated Models," RCER Working Paper No. 281.

Park, J. Y., S. Ouliaris, and B. Choi (1988): "Spurious Regressions and Tests for Cointegration," CAE Working Paper No. 88-07, Cornell 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(2), 277–301.

——— (1998): "New Tools for Understanding Spurious Regressions," *Econometrica*, 66, 1299–1325.

Phillips, P. C. B., and S. N. Durlauf (1986): "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, 53, 473–495.

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

Phillips, P. C. B., and M. Loretan (1991): "Estimating Long-Run Economic Equilibria," *Review of Economic Studies*, 58, 407–436.

Phillips, P. C. B., and S. Ouliaris (1990): "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 58(1), 165–193.

Saikkonen, P. (1991): "Asymptotically Efficient Estimation of Cointegrating Regressions," *Econometric Theory*, 7, 1–21.

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

Stock, J. H., and M. W. Watson (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61(4), 783–820.

West, K. D. (1988): "Asymptotic Normality, When Regressors Have a Unit Root," *Econometrica*, 56(6), 1397–1417.