David A. Dickey, Dennis W. Jansen
and Daniel L. Thornton

David A. Dickey is a professor of statistics at North Carolina State University, Dennis W. Jansen is an associate professor of economics at Texas A&M University and Daniel T. Thornton is an assistant vice president at the Federal Bank of St. Louis. The authors would like to thank Mark Watson, Tom Fomby and David Small for helpful comments on an earlier draft of this article. Lora Holman and Kevin L. Kliesen provided research assistance.

A Primer On Cointegration with an Application to Money and Income

FOR SOME TIME NOW, macroeconomists have been aware that many macroeconomic time series are not stationary in their levels and that many time series are most adequately represented by first differences.[1] In the parlance of time-series analysis, such variables are said to be integrated of order one and are denoted I(1). The level of such variables can become arbitrarily large or small so there is no tendency for them to revert to their mean level. Indeed, neither the mean nor the variance is a meaningful concept for such variables.

Nonstationarity gives rise to several econometric problems.[2] One of the most troublesome stems from a common prediction of macroeconomic theory that there should be a stable long-run relationship among the levels of certain economic variables. That is, theory often suggests that some set of variables cannot wander too far away from each other. If individual time series are integrated of order one, however, they may be "cointegrated." Cointegration means that one or more linear combinations of these variables is stationary even though individually they are not. If these variables are cointegrated, they cannot move "too far" away from each other. In contrast, a lack of cointegration suggests that such variables have no long-run link;

---

[1]That is, formal statistical tests often cannot reject the null hypothesis of a unit root. The results of these tests, however, are sensitive to how the tests are performed— that is, whether the tests assume a non-zero mean or a time trend, whether an MA or AR data generating processes is assumed [Schwert (1987)] and whether the test is performed using classical or Bayesian statistical inference [Sims (1988), and Sims and Uhlig (1988)]. These sensitivities are partly due to the lack of power these tests have against an alternative hypothesis of a stationary but large root.

[2]It can give rise to the possibility of a spurious relationship among the levels of the economic variables. Also, the parameter estimates from a regression of one such variable on others are inconsistent unless the variables are cointegrated.

in principle, they can wander arbitrarily far away from each other.[3]

This article illustrates the salient features of cointegration and tests for cointegration. The discussion, initially motivated by the simple example of Irving Fisher's "equation of exchange," draws an analogy between cointegration and unit roots on the one hand and tests for cointegration among multiple time series and the usual tests for unit roots in univariate time-series analysis on the other. The article then addresses the broader question of the economic interpretation of cointegration by contrasting it with the usual linear, dynamic, simultaneous equation model which is frequently used in macroeconomics.

The article goes on to compare three recently proposed tests for cointegration and outlines the procedures for applying these tests. An application of these tests to U.S. time-series data using alternative monetary aggregates, income and interest rates suggests that there is a stable long-run relationship among real output, interest rates and several monetary aggregates, including the monetary base.

TESTING FOR COINTEGRATION: A GENERAL FRAMEWORK

Because of the close correspondence between tests for cointegration and standard tests for unit roots, it is useful to begin the discussion by considering the univariate time-series model

(1) $y_t - \mu = \varrho(y_{t-1} - \mu) + e_t,$

where y_t denotes some univariate time series, μ is the series' mean and e_t is a random error with an expected value of zero and a constant, finite variance. The coefficient ϱ measures the degree of persistence of deviations of y_t from μ.

When $\varrho = 1$, these deviations are permanent. In this case, y_t is said to follow a random walk— it can wander arbitrarily far from any given constant if enough time passes.[4] In fact, when $\varrho = 1$ the variance of y_t approaches infinity as t increases and the mean of y_t, μ, is not defined. Alternatively, when $|\varrho| < 1$, the series is said to be mean reverting and the variance of y_t is finite.

Although there is a similarity between tests for cointegration and tests for unit roots, as we shall see below, they are not identical. Tests for unit roots are performed on univariate time series. In contrast, cointegration deals with the relationship among a group of variables, where (unconditionally) each has a unit root.

To be specific, consider Irving Fisher's important equation of exchange, $MV = Pq$, where M is a measure of nominal money, V is the velocity of money, P is the overall level of prices and q is real output.[5] This equation can be rewritten in natural logarithms as:

(2) $\ln M + \ln V - \ln P - \ln q = 0.$

In this form, the equation of exchange is an identity. The theory of the demand for money, however, converts this identity into an equation by making velocity a function of a number of economic variables; both the form of the function and its arguments change from one theoretical specification to another. In the theory of money demand, V is unobservable and in applied work it is proxied with some function of economic variables, V^*, $\ln V^* = \ln V + E$, where E denotes a random error associated with the use of the proxy for V.[6] The proxy is a function of one or more observed variables, other than income and prices, that are hypothesized to determine the demand for money. Hence, equation 2 is replaced with

---

[3]At the present time, tests for cointegration deal only with looking for stable *linear* relationships among economic variables. Consequently, a failure to find cointegration does not necessarily mean that there is no stable, long-run relationship among the variables, it only suggests that there is no stable, long-run, linear relationship among them.

[4]That is, for any numbers C > 0 and 0 < p < 1 and for any starting value Y_0 there is a time, T, such that, for all t > T, $Pr(|Y_t| > C) > p$. When $|\varrho| < 1$ the process generating Y_t is stationary in that it does not wander too far from its mean, i.e., for any given probability p we can find a constant C > 0 such that $Pr(|Y_t - \mu| > C) < p$.

[5]The cointegrating vector could be different from the hypothesized one for other reasons as well. For exam-

ple, money holders might have a money illusion or money demand might not be homogenous of degree one in real income.

[6]For the classic discussion of velocity and a long list of its potential determinants, see Friedman (1956). Empirical proxies for velocity often contain one or more of these determinants.

(2') $\ln M + \ln V^* - \ln P - \ln q = E$.

If the proxy is good, the expected value of E should be zero. Furthermore, E should be stationary, so that, V^* might deviate from its true value in the short-run, but should converge to it in the long-run. Failure to find a stationary relationship among these variables—that is, to find that they are not cointegrated—implies either that V^* is a poor proxy for V or that the long-run demand for money does not exist in any meaningful sense.

In essence, the Fisher relationship embodies a long-run relationship among money, prices, output and velocity. In particular, it hypothesizes that the cointegrating vector $(1, 1, -1, -1)$ exists. This vector combines the four series into a univariate series, E. Given this assumption, a test for cointegration can be performed by applying any conventional unit root test to E.

Using conventional unit root tests to test for cointegration [such as tests prepared by Dickey and Fuller (1979) and Phillips (1987)], however, requires prior knowledge of the cointegrating vector. And most often, this vector is unknown. Therefore, some linear combination of these variables, for example,

(3) $b_1 \ln M + b_2 \ln V^* + b_3 \ln P + b_4 \ln q$,

is hypothesized to be stationary, where the cointegrating vector (b_1, b_2, b_3, b_4) is unknown and must be estimated.

Locating Stationary Linear Combination of Variables

Engle and Granger (1987), Stock and Watson (1988), and Johansen (1988) have suggested alternative tests for cointegration and methods for estimating the cointegrating vectors. While differing in a number of respects, all of these procedures involve locating the "most stationary" linear combinations (among all of the possible ones) of the vector time series in question. If the linear combinations being compared are not chosen a priori, but are determined by choosing among all possible vectors, tests for cointegration encounter the type of distributional problems associated with order statistics and multiple comparisons. Hence, it is useful to discuss some of these problems in more detail.

In multiple comparison tests, an experimenter is usually concerned with, say, comparing the highest and lowest sample means among several. He wants to find the pair of sample means with the largest difference to see whether the difference is statistically significant. When the means to be compared are chosen ahead of time, tests for a significant difference between the means can be done using the usual t-statistics. If, however, the means to be compared are chosen simply because they are the largest and smallest from a sample of, say, five means, the rejection rate under the null hypothesis will be much higher than that implied by the percentile of the t-distribution. In order to control for the experimentwise error rate, as it is called, tables of distributions of highest mean minus lowest mean (standardized) have been computed for the case of no true differences in the population means. These "tables of the studentized range" are then used to test for significant differences between the highest and lowest means.[7]

The price paid for controlling the experimentwise error rate is a loss in power. That is, the difference between the means must be much larger than in the case of the standard t-test before it can be declared significant at some predetermined significance level. Thus, the power of the test to detect significant differences is reduced.

In an analogous way, it is difficult to reject the null hypothesis that there are no stationary linear combinations when the observed data are used to estimate the most stationary-looking linear combination before testing for cointegration. This loss of power is evident in the test tables given by Stock and Watson or Johansen, where the percentiles are shifted far away from those of the standard unit root distribution of Dickey and Fuller (1979).[8] Consequently, detecting cointegrating relationships among variables is relatively hard. Some power can be gained, however, if economic theory is used to assign values to some coefficients, a priori. Indeed, if theory fully specifies the cointegrating vector, as in our example of the Fisher equation, using conventional unit root tests to test for cointegration would be appropriate.

---

[7] See Steel and Torrie (1980), p. 588, for these tables.

[8] See Stock and Watson (1988) and Johansen (1988) for further discussion of the relationship to order statistics.

Multiple Cointegrating Vectors

Until now, the possibility that only one linear combination of variables is stationary has been considered; however, this need not be the case. In cases where more than two time series are being considered, more than one stable linear combination can exist. Also, until now, cointegration was discussed without any explicit reference to a dynamic specification of the levels of the economic variables. Nevertheless, the fact that cointegration is related to their dynamic specification was implicit in the fact that all of the univariate series are I(1). While a number of alternative multivariate representations could be used, it is convenient to use the following multivariate AR(1) representation:

$$(4)\ Y_t\ =\ AY_{t-1}\ +\ \varepsilon_t,$$

where Y_t, an n by 1 vector, is Z_t minus μ, where Z_t is a vector of economic time series (in our example, M, V*, P and q) and μ is the vector of the means of Z.[9] A is an n by n matrix and ε_t is a vector of independent random disturbances that are stationary around zero—that is, $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \Omega$ for all t.

The possibility of k cointegrating vectors means there exists a k by n matrix β', of rank k, such that $\beta'Y_t$ is stationary in the sense that it is mean reverting. It is assumed that all of the elements of Y_t are integrated of the same order, I(1); however, for the sake of illustration (here and elsewhere) we first consider the possibility that the elements of Y_t are I(0). In this case, the long-run stationary solution to equation 4 is $Y_t = (I-A)^{-1}\varepsilon_t$.[10] But there is no need to ask whether these variables are cointegrated, because clearly any linear combination of Y_t is mean reverting.

Now consider the case where each element of Y_t is I(1). Assume further that the elements of Y_t are mutually independent so that $A = I$. In this case, no long-run equilibrium exists because the matrix $(I-A)$ is of rank zero. Since any linear combination of these independent I(1) series must itself be I(1), these variables are not cointegrated. No stationary linear combination of these variables exists!

Finally, consider the case where all of the univariate series are I(1), but A is not an identity matrix. In this case, not every linear combination of Y_t is stationary because $(I-A)$ is not of full rank, that is, $(I-A)^{-1}$ does not exist. However, as we will show, some linear combinations of Y_t may be stationary. The number of such cointegrating vectors is determined by the rank of $(I-A)$. From a purely statistical point of view, cointegration places some restrictions on the matrix A. From an economic perspective, economic theory, which determines the matrix A, places some restrictions on the long-run behavior of Y_t.

From a somewhat broader perspective then, the objective of cointegration analysis is to find an n by n matrix B', of rank n, such that $B'Y_t$ decomposes Y_t into its stationary and nonstationary components. This is accomplished by obtaining a k by n sub-matrix of B', β', of rank k such that the transformed series $\beta'Y_t$ is stationary. The k rows of B' associated with these stationary series are called "cointegrating vectors." The remaining $n-k$ unit root combinations are termed "common trends."[11]

Tests for Cointegration and Their Relation to Unit Root Tests

In illustrating tests for cointegration, we draw the analogy between tests for cointegration and tests for unit roots. Any autoregressive time series of order p can be written in terms of its first difference, one lag level and $p-1$ lag differences. Consider first the univariate case,

$$(5)\ y_t^*\ =\ a_1 y_{t-1}^* + a_2 y_{t-2}^* + \ldots + a_p y_{t-p}^* + e_t,$$

where $y_t^* = y_t - \mu$. The equation can be reparameterized as

$$(6)\ \Delta y_t^*\ =\ b_1 \Delta y_{t-1}^* + b_2 \Delta y_{t-2}^* + \ldots + b_{p-1}\Delta y_{t-p+1}^* - c y_{t-p} + e_t,$$

where $b_i\ =\ -1 + a_1 + a_2 + \ldots + a_i$ and $c\ =\ 1 - a_1 - a_2 - \ldots - a_p$.

---

[9]Note that the system given by Y could be thought of as a multivariate p^{th}-order AR system that has been rewritten as an AR(1) system.

[10]$(I-A)^{-1}$ exists if all of the eigenvalues of A are less than one in absolute value.

[11]This terminology stems from the fact that the original series are retrieved from the transformed ones by multi-

plying the transformed vector by the inverse of B. In general, this spreads the unit root "trend" processes through all the original series depending on where the zeros occur in B^{-1}. See Stock and Watson (1988).

Alternatively, equation 5 can be reparameterized as

$$(7) \quad \Delta y_t^* = d_1 \Delta y_{t-1}^* + d_2 \Delta y_{t-2}^* + \ldots + d_{p-1} \Delta y_{t-p+1}^* - c y_{t-1} + e_t,$$

where $d_{p-1} = -a_p$, $d_{p-2} = -a_p - a_{p-1}, \ldots,$
$d_1 = -a_p - a_{p-1} - \ldots - a_2$.

The Dickey and Fuller test uses the t-statistic from the ordinary least squares regression to test the null hypothesis that the coefficient c in equation 7 is equal to zero. The t-statistic is the likelihood ratio test of the null hypothesis of a unit root, where the likelihood is conditional upon an initial value, y_o.[12]

Consider now the multivariate analogue to equation 5:

$$(8) \quad Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + A_3 Y_{t-3} + \ldots + A_p Y_{t-p} + \varepsilon_t,$$

where $A_1, A_2, \ldots A_p$ are n by n matrices. Equation 8, too, can be reparameterized as either

$$(9) \quad \Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \Gamma_2 \Delta Y_{t-2} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} - \psi Y_{t-p} + \varepsilon_t,$$

or

$$(10) \quad \Delta Y_t = \theta_1 \Delta Y_{t-1} + \theta_2 \Delta Y_{t-2} + \ldots + \theta_{p-1} \Delta Y_{t-p+1} - \psi Y_{t-1} + \varepsilon_t,$$

If the matrix $\psi = (I - A_1 - A_2 - \ldots - A_p)$ is full rank, then any linear combination of Y_t will be stationary. If ψ is a matrix of zeros, then any linear combination of Y_t will be a unit root process and, hence, nonstationary.[13]

This leaves an intermediate case where ψ is not a matrix of zeros, but is less than full rank.

The rank of ψ, k, is the number of linearly independent and stationary linear combinations of Y_t that can be found. In other words, it is the number of linearly independent cointegrating relations among the variables in Y_t. Of course, the estimate of ψ, $\hat{\psi}$, will almost always be of full rank in a numerical sense. The objective of tests for cointegration is to test for the rank of ψ by testing whether the eigenvalues of $\hat{\psi}$ are significantly different from zero.[14] From this perspective, the tests for cointegration purposed by Stock and Watson (1988) and Johansen (1988) are multivariate analogues of the Dickey-Fuller test.[15]

IS THERE AN ECONOMIC INTERPRETATION OF COINTEGRATION VECTORS?

It should be clear now that, if several I(1) variables are cointegrated, then one or more linear combinations of them will have a finite variance. But the broader question of the economic interpretation of such cointegrating vectors remains. In general, cointegrating vectors are obtained from the reduced form of a system where all of the variables are assumed to be jointly endogenous. Consequently, they cannot be interpreted as representing structural equations because, in general, there is no way to go from the reduced form back to the structure.[16] Nevertheless, they might be thought of as arising from a constraint that an economic structure imposes on the long-run relationship among the jointly endogenous variables. For example, economic theory suggests that arbitrage

---

[12]An initialization assumption is necessary for unit root processes. The critical values of the test statistics change depending on whether μ is equal to zero or whether a time trend is included. Dickey and Fuller (1979) present versions of both tests where μ is assumed to be 0 and where μ is replaced with a linear time trend.

[13]For example, let p=2, if $A_2 = 0$ and $A_1 = I$, then $(I - A_1 - A_2) = 0$. In this case, all the elements of X_t are unit root processes and $(I - A_1 - A_2)$ is rank zero. But, if $A_2 = 0$ and A_1 is $(\varrho_1, \varrho_2, \ldots \varrho_n)I$, where $|\varrho_i| < 1$ for all i, i = 1,2, . . . ,n, then $(I - A_1 - A_2)$ is full rank and all the elements of X_t are stationary AR(1) processes.

[14]To our knowledge, Theil and Boot (1962) were the first to prepare a test for the statistical significance of the eigenvalues of $\hat{\psi}$. In so doing, they were the first to develop a test for cointegration.

[15]There are three versions of the univariate test for unit roots (zero mean, constant, non-zero mean and a constant linear time trend). Johansen (1988) presents the multivariate analogue of the likelihood ratio test when μ is zero.

Johansen and Juselius (1990) extend this to the case where μ is non-zero. In practice, the difference between Y_t and Y_1, the first sample observation, should be used if the $\mu=0$ assumption is maintained; hence, it seems reasonable to use the unconstrained μ versions of the tests [See Dickey, Bell and Miller (1986) for a discussion of which test to use]. Note that the formulas for computing the test statistics are straightforward least-squares formulas. It is the distribution of the test statistics that are nonstandard.

[16]Of course, because equation 17 is essentially a multivariate VAR, in principle, it should be possible to give these a "structural" interpretation by imposing identifying restrictions on the reduced-form parameters, as has been done recently for VAR models. For example, see Bernanke (1986) and Blanchard and Quah (1989).

will keep nominal interest rates—especially those on assets with the same or similar maturity—from getting too far away from each other. Thus, it is not surprising that such interest rates are cointegrated.[17]

Cointegration with Exogenous Variables

The importance of cointegration in economics can be highlighted by noting the close correspondence between cointegration and the typical linear, dynamic, simultaneous equation model used in macroeconomics and econometrics,

$$(11) \quad AY_t = BY_{t-1} + CX_t + u_t.$$

Y_t is a n by 1 vector of endogenous variables, X_t is a g by 1 vector of exogenous variables and u_t is a n by 1 vector of random disturbances, and A, B and C are matrices of unknown parameters. It is assumed that A^{-1} exists, so that the dynamic reduced form can be written as

$$(12) \quad Y_t = \pi Y_{t-1} + \Gamma X_t + \varepsilon_t,$$

where π $(= A^{-1}B)$ and Γ $(= A^{-1}C)$ are matrices of unknown reduced-form parameters. Equation 12 contains only predetermined variables on the right-hand side, so that the dynamic response of Y_t to X_t can be studied by recursive substitution using equation 12. Letting L denote the lag operator $(LY_t = Y_{t-1})$, the result is written as

$$(13) \quad Y_t = (I + \pi L + \pi^2 L^2 + \ldots)\Gamma X_t$$
$$+ (I + \pi L + \pi^2 L^2 + \ldots)\varepsilon_t.$$

The infinite series $(I + \pi L + \pi^2 L^2 \ldots)$, evaluated at $L = 1$, converges to $(I-\pi)^{-1}$ if all of the eigenvalues of π are less than 1 in absolute value.

The expected value of equation 13, conditional on X_t, is

$$(14) \quad E(Y_t) = (I + \pi L + \pi^2 L^2 \ldots)\Gamma X_t$$
$$= (I - \pi L)^{-1} \Gamma X_t.$$

Equation 14 is used to investigate the long-run response of Y_t to a change in one or more of the elements of X_t. Assuming that X_t is a vector

of zeros for all $t \leq 0$ and $X_t = \delta$ for all $t > 0$, and that the system converges, then, in the limit

$$(15) \quad E(Y) = (I - \pi)^{-1}\Gamma\delta.$$

Thus, $(I - \pi)^{-1}\Gamma\delta$ gives the long-run response of Y_t to a permanent change, δ, in X. Theil and Boot (1962) have termed $(I - \pi)^{-1}\Gamma$ the final-form multipliers.

The cointegration techniques of Johansen and Stock and Watson start with an equation similar to equation 12. There are two key differences, however. First, all of the variables are explicitly endogenous. Second, because π has unit eigenvalues corresponding to the common trends in the system, $(I - \pi)$ is not invertible, and, therefore, the final-form multipliers in 15 are undefined.[18]

A related technical point is that the initial conditions are not transient. To see this in equation 12, let $X_t = 0$ for all t and substitute recursively to obtain

$$(16) \quad Y_t = \varepsilon_t + \pi\varepsilon_{t-1} + \pi^2\varepsilon_{t-2} + \ldots + \pi^{t-1}\varepsilon_1$$
$$+ \pi^t Y_0.$$

Because the matrix π^i does not converge to a matrix of zeros as t approaches infinity, the initial condition is not transient and must be specified. Often, the initial vector, Y_0, is assumed to be a vector of zeros.

Can cointegration be used to determine the long-run response to changes in exogenous variables? In general, the answer is yes; however, how this is done, the interpretation of the estimated cointegrating vectors and the method of estimation, all depend on the assumptions made about X_t. Consider, for example, the case where the elements of X_t are non-stochastic and fixed in repeated samples. Rewrite equation 12 as

$$(17) \quad \Delta Y_t = -(I-\pi)Y_{t-1} + \Gamma X_t + \varepsilon_t.$$

If Y_t is cointegrated, the n by n matrix $(I-\pi)$ is rank $k < n$, and π can be represented as $\alpha\beta'$, where α and β are n by k matrices of rank k. Then

$$(18) \quad \beta'Y_t = (I-\beta'\alpha)\beta'Y_{t-1} + \beta'\Gamma X_t + \beta'\varepsilon_t$$

[17]Stock and Watson (1988) find that the nominal federal funds, the three-month Treasury bill and one-year Treasury bill rates are cointegrated.

[18]Finding unit eigenvalues in π is equivalent to finding zero eigenvalues in ψ. This is so because $\psi = 1-\pi$ so that $|\psi - \lambda I| = |\lambda(1-\pi)-I| = |\pi-(1-\lambda)I| = |\pi-\delta I| = 0$, so $\lambda = 0$ is equivalent to $\delta = 1$.

is stationary.[19] Equation 18 can be expressed as the k-dimensional system, with $\beta' Y_t = Y^*_t$,

(19) $Y^*_t = \pi^* Y^*_{t-1} + \Gamma^* X_t + \varepsilon^*_t.$

Because the k by k matrix $I - \pi^*$ is full rank, $(I - \pi^*)^{-1}$ exists. Thus, this lower-dimensional system has the steady-state representation,

(20) $Y^*_t = (I - \pi^*)^{-1} \Gamma^* X_t + (I - \pi^*)^{-1} \varepsilon^*_t.$

The expected value of 20 gives the final-form multipliers for the cointegrated subsystem of Y_t. Consequently, there is a representation for the long-run average response of $\beta' Y_t$ to a change in X.

It must be remembered, however, that because the estimated cointegrating vector is conditional on the information set in the model, X_t needs to be included along with lagged differences of the endogenous variables when estimating the cointegrating vectors. Moreover, the distribution of the test statistic will not be invariant to X_t. Consequently, the Monte Carlo experiments used to derive the distributions of the test statistics used by, say, Johansen (1988), would have to be redone including X_t in the model.

In the case where X_t is stochastic, the method used to obtain estimated cointegrating vectors and their interpretation changes with the assumed structure of the model. A general form of such a model is

(21) $\begin{bmatrix} A & C \\ O & I \end{bmatrix} \begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} B & O \\ O & D \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix},$

where it is assumed that

$E(\varepsilon_t \eta_t)' (\varepsilon_t \eta_t) = \begin{bmatrix} \sigma^2_\varepsilon \Sigma_\varepsilon & 0 \\ 0 & \sigma^2_\eta \Sigma_\eta \end{bmatrix},$

for all t. Again, assume that A^{-1} exists, so 21 can be rewritten as

(22) $\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} A^{-1}B & -A^{-1}CD \\ O & D \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta_t \\ \eta_t \end{bmatrix}$

or

(23) $\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} \pi & \Gamma D \\ O & D \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta_t \\ \eta_t \end{bmatrix}.$

Cointegrating vectors and their interpretation depend on the assumptions made about π and D. For example, if all of the eigenvalues of π are less than one, so that conditional on X_t, Y_t is stationary, the common trends in the system are the result of X_t being non-stationary. The number of common trends will be equal to the number of unit roots in D. If $D = I$, for example, then there will be g common trends and n cointegrating vectors. In this case, there is a steady-state representation for the endogenous variables conditional on the exogenous variables.[20] That is, the final-form multipliers of Theil and Boot are conditional on the exogenous variables. They do not exist, however, because the variance of the exogenous variables is unbounded and, hence, so is the unconditional variance of the endogenous variables.

If π also has some unit eigenvalues, however, then part of the nonstationality of the system is due to instability in the dynamics of the endogenous variables in the system.[21] In this case, the conditional final-form multipliers, analogous to those obtained by Theil and Boot, exist along the stationary directions given by the cointegrating vectors as before. An important aspect of this formulation is that, given the proposed structure of the system, the researcher can identify whether the common trends stem from the structural dynamics or are simply a manifestation of the stochastic properties of the exogenous variables. Under the assumption on $(\varepsilon_t \eta_t)'$, estimating cointegrating vectors for the system given by the equations in 23 is the same as estimating them for the system given by the equations in 8.[22] Hence, any of the multivariate methods discussed in the next section can be employed. The reader is cautioned, however, that these procedures would have to be modified to impose the upper triangular structure of the system given by the equations in 23.[23]

---

[19]More technically, the eigenvalues of $I - \beta' \alpha$ are less than one in absolute value.

[20]A form of this model has been proposed recently by Hoffman and Rasche (1990). This is the Stock-Watson formulation, except that X_t is a set of latent variables. Conditional on the unobserved X_t, the endogenous variables are assumed to be stationary.

[21]This so because

$\begin{vmatrix} \pi - \lambda I & \Gamma D \\ 0 & D - \lambda I \end{vmatrix} = |\pi - \lambda I| \, |D - \lambda I| = 0.$

[22]If these error terms are not independently distributed obtaining consistent estimates of the cointegrating vector will be more difficult. For an example of this case, see Stock and Watson (1989).

[23]Indeed, both Stock and Watson (1988) and Johansen (1988) allow for the possibility of exogenous variables in the sense of equation 22; however, they do not impose the exogeneity restriction *ex ante*.

Should There Be Many or Few Cointegrating Vectors?

Since it is possible to have n-1 cointegrating vectors in the system given by 17, the question naturally arises, is it better to have many or few cointegrating vectors? Providing a general answer to this question is difficult. Cointegrating vectors can be thought of as representing constraints that an economic system imposes on the movement of the variables in the system in the long-run. Consequently, the more cointegrating vectors there are, the "more stable" the system. Other things the same, it is desirable for an economic system to be stationary in as many directions as possible.

To see why, consider a model with no common trends, so the system is stationary: Y_t never wanders "too far" from its steady-state equilibrium value (in the model discussed in the text, the vector of means). If there is one common trend and $n-1$ cointegrating vectors, however, $n-1$ of the variables must be solved for in terms of the nth, and the structure of these variables follows a single common trend. Hence, there are only $n-1$ directions (as opposed to n in the previous example) where the variance is finite and one direction in which it is infinite. On the other hand, if there is only one cointegrating vector, the nth variable must be solved for in terms of the other $n-1$ variables. The system can wander off in $n-1$ independent directions; it is stable in only one direction.[24]

To see this point from a geometric perspective, consider the case were there are three endogenous variables that span R^3. If these variables are stationary, the system converges to a steady-state equilibrium, a point in R^3, and variation around that point is finite. If the variables are non-stationary, and there is one common

trend, however, the system converges to a long-run equilibrium represented by a line, determined by the intersection of the planes defined by the two cointegrating vectors, in R^3.[25] This is a stationary equilibrium in the sense that the variance about this line is finite. If there are two common trends and one cointegrating vector, the long-run equilibrium is represented by a plane defined by the single cointegrating vector. The variables are unbounded in the plane, but cannot move too far from it. That is to say, that the variance in the plane is infinite, but the variance about the plane is finite. If there are no cointegrating vectors, the variables are free to wander anywhere in R^3—they are unbounded!

Consequently, when non-stationary variables are cointegrated, there exists a direction where a meaningful long-run relationship among them exists. The fewer the number of cointegrating vectors, the less constrained is the long-run relationship. Hence, all other things the same, it seems desirable to have many cointegrating vectors.[26] Alternatively stated, we prefer economic models that have unique steady-state equilibria. Accordingly, researchers are interested not only in testing to see whether variables are cointegrated, but in obtaining precise estimates of the cointegrating vectors.

ALTERNATIVE TESTS FOR COINTEGRATION

With a number of tests for cointegration being available, it is important to understand their similarities and their differences. The purpose of this section is to discuss the salient features of alternative tests for cointegration. Step-by-step instructions on how to perform two of the more difficult of these (the Stock-Watson and Johansen tests) are presented in a shaded insert on the following pages.

---

[24]In terms of the model given by the equations in 22, when $(I-\pi)$ is full rank, the more cointegrating vectors in D, the more directions in which the "final-form" multipliers will exist.

[25]Of course, the subsystem defined by equation 20 converges to a point in R^2. This point is given by the intersection of this plane and the line given by the intersection of the two cointegrating vectors.

[26]Having said that, we should hasten to add that it is very unlikely that these tests will indicate that there are a large number of cointegrating vectors. These tests lack the power to reject the null hypothesis of zero cointegrating

vectors. At a more practical level, it is well known that macroeconomic time-series data are highly correlated so that, typically, the generalized variance of the matrix of such variables is concentrated on relatively few principal components. See footnote 27.

Step-By-Step Application Of The Johansen And Stock-Watson Approaches To Cointegration

Because the procedures developed by Johansen, Stock and Watson are more difficult to employ, this insert provides step-by-step procedures for applying these approaches. Both of these procedures can be illustrated with the multivariate model (equation 8 of the text)

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + \varepsilon_t.$$

Step-by-Step Application of Johansen's Approach to Integration:

1. Pick an autoregressive order p for the model.
2. Run a regression of ΔY_t on ΔY_{t-1}, ΔY_{t-2}, ..., ΔY_{t-p+1} and output the residuals, D_t. For each t, D_t has n elements.
3. Regress Y_{t-p} on ΔY_{t-1}, ΔY_{t-2}, ..., ΔY_{t-p+1} and output the residuals, L_t. For each t, L_t has n elements.
4. Compute squares of the canonical correlations between D_t and L_t, calling these $\varrho_1^2 > \varrho_2^2 > \ldots > \varrho_n^2$.
5a. Letting N denote the number of time periods available in the data, compute the trace test as

$$\text{TRACE TEST} = -N \sum_{i=k+1}^{n} \ln(1 - \varrho_i^2). [1]$$

The null hypothesis is "there are k or less cointegrating vectors."

5b. You may choose to use the maximal eigenvalue test (which really uses the $k+1^{th}$ largest squared canonical correlation or eigenvalue) as follows:

MAX EIGENVALUE TEST =
$-N \ln(1 - \varrho_{k+1}^2)$.

6. Compare the test statistic to the appropriate table in Johansen and Juselius (1990).[2]
Note: The squared canonical correlations are the solution to the determinantal equation

$$|\varrho_i^2 S_{kk} - S_{k0} S_{00}^{-1} S_{k0}'| = 0 \text{ where}$$

$$S_{kk} = N^{-1} \sum_{t=1}^{N} L_t L_t' \quad S_{00} = N^{-1} \sum_{t=1}^{N} D_t D_t' \text{ and}$$

$$S_{k0} = N^{-1} \sum_{t=1}^{N} L_t D_t' \text{ and } D_t \text{ and } L_t \text{ are col-}$$

umn vectors of residuals from steps 2 and 3. The maximum likelihood estimates of the k cointegrating vectors (k columns of β_i for which $\varrho_i^2 S_{kk} \beta_i = S_{k0} S_{00}^{-1} S_{k0}' \beta_i$).

Step-by-Step Application of the Stock-Watson Approach to Cointegration:

In the Stock-Watson approach the null hypothesis is that there are m common trends $(n - k = m)$ against the alternative that there are less than m, say $m - q$, common trends. There are six steps:

1. Pick the autoregressive order p for the model.
2. Compute the eigenvectors of $\Sigma Y_t Y_t'$, that is, do a principal components analysis of Y_t.

---

[1] Theoretically, a cointegrating vector is associated with $\varrho_1^2, \varrho_2^2, \varrho_3^2, 0, 0$. The theoretical counterpart of the trace test for H_0: k = 3 or less cointegrating vectors is $-N \Sigma \ln(1-0) = 0$ so the test statistic would be within sampling error of 0. For H_0: k = 2 or less cointegrating vectors, the theoretical value of the test is $-N \Sigma_{i=3} \ln(1-\varrho_i^2) = -N \ln(1-\varrho_3^2) > 0$ and as N gets large, this diverges to $+\infty$. Note that ϱ_4^2 and ϱ_5^2 (which are both zero) do not contribute to the test statistic and this is the motivation for the so-called "maximal eigenvalue" test of Johansen.

[2] The appropriate table as well as the handling of the intercept term in estimation depends on the role of the intercept in the model. For a discussion of this, refer to Dickey and Rossana (1990).

3. Using the m principal components with highest variance, that is, largest eigenvalues, fit a vector autoregression to the differences. If P_t is the vector of m principal components (selected as described in the text) then the autoregressive model is denoted $\Delta P_t = A_1 \Delta P_{t-1} + \ldots + A_{p-1} \Delta P_{t-p-1} + \varepsilon_t$, where, as before, p stands for the number of lags in the "original" autoregression. This provides a filter to use in step 4.

4. Compute a filtered version, $\hat{F}_t$, of P_t by $\hat{F}_t = P_t - \hat{A}_1 P_{t-1} - \ldots - \hat{A}_p P_{t-p}$. This

reduces the multi-lag model to a one lag model.

5. Regress $\Delta \hat{F}_t$ on $\hat{F}_{t-1}$ getting coefficient matrix B.

6. Compute the eigenvalues of B, normalize, and compare to the distributional tables of Stock and Watson (1988). Rejecting the null hypothesis of m common trends in favor of the alternative of m-q common trends means a reduction in the number of common trends by q and thus an increase of q in the number of cointegrating vectors.

The relationship among tests for cointegration developed by Johansen (1988), Engle and Granger (1987), Stock and Watson (1988) and Fountis and Dickey (1989) can be illustrated by considering the multivariate model

$$(24)\ Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + \varepsilon_t.$$

Johansen reparameterizes equation 24 as

$$(25)\ \Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \Gamma_2 \Delta Y_{t-2} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} - \psi Y_{t-p} + \varepsilon_t,$$

where , as before, $\psi = (I - A_1 - A_2 \ldots - A_p)$. He then makes use of the fact that any n by n matrix, ψ, of rank k < n can be written as the product of two n by k matrices of rank k—that is, $\psi = \alpha\beta'$, where α and β are n by k matrices of rank k. He maximizes the likelihood function for Y_t conditional on any given β using standard least squares formulas for regression of ΔY_t on $\Delta Y_{t-1}, \Delta Y_{t-2}, \ldots, \Delta Y_{t-p+1}$ and $\beta' Y_{t-p}$. The solution to this maximization problem gives estimates of $\Gamma_1, \Gamma_2, \ldots, \Gamma_{p-1}$ and α conditional on β. Once this is done, β, or more specifically, the row space of β, is estimated.

In cointegration, it is only possible to determine the rank of $\alpha\beta'$.[27] Specifically, obtaining unique elements of α and β without imposing arbitrary constraints is impossible. The rank of ψ can be obtained by computing canonical correlations between ΔY_t and Y_{t-p}, adjusting for all intervening lags. Johansen chooses to put the lag level at the largest lag but, as has been shown earlier, this is not crucial.

In the Johansen approach, $-\alpha\beta'$ is the coefficient matrix on the lagged level. Upon pre-multiplying equation 25 by β', the last term in equation 25 is $\beta'\alpha\beta'Y_{t-p}$. Note that $\beta'\alpha$ has no zero eigenvalues so that $\beta'Y_t$ is a stationary vector time series of dimension k. Thus, rows of β' are the cointegrating vectors.

Now consider adding to the k rows of β', n-k more rows orthogonal to the columns of α. Denote the resulting matrix B'. Note that B'α is simply n-k rows of zeros appended to the bottom of $\beta'\alpha$. In equation 25, after this transormation, the last n-k rows involve only differences. Hence, $C_t = B'Y_t$ is a column vector of k stationary processes followed by n-k unit root processes (common trends). In Johansen's approach, then, the matrix B' has the k cointegrating vectors as its first k rows and coefficients yielding the n-k common trends as its last n-k rows.

By standard results in unit root estimation, the vector $C_t = B'Y_t$ is such that $N^{-2} \Sigma C_t C_t'$ (where N is the series' length) converges to an n by n matrix with zeros everywhere except in the lower right n-k by n-k submatrix. This result underlies the other approaches to cointegration.

For example, Stock and Watson point out that if $W_t = (B')^{-1}C_t$, then the sum of squares and cross products matrix $N^{-2}\Sigma W_t W_t'$ converges to the same limit as $(B')^{-1}N^{-2}\Sigma C_t C_t' B^{-1}$. This limit matrix has rank n-k and, thus, has n-k nonzero eigenvalues. Because the generalized variance of

---

[27]The space spanned by the columns of the α and β' matrices is important because $\alpha\beta'$ can be obtained by many choices of matrices. To see this, note that for any nonsingular k by k matrix, H, $(\alpha H^{-1})(H\beta') = \alpha\beta'$.

a sum-of-squares and cross-products matrix is equal to the sum of the eigenvalues of this matrix, they suggest a test based on computing the largest eigenvalues of the sum-of-squares and cross-products matrix.[28]

In multivariate statistics, a vector random variable Y can be transformed into a canonical vector random variable C = TY by choosing T (a matrix of constants) so that the elements of C are uncorrelated. The elements of C are referred to as "principal components" of Y. Stock and Watson reason that if a vector process has $n - k$ common trends (that is, the vector $C_t = B'Y_t$ has $n - k$ unit root processes and k stationary ones), then the $n - k$ principal components with largest variance should correspond to the unit root processes or "common trends." This reasoning is based on the previously stated notion that the normalized sum-of-squares and cross-product matrix $N^{-2}\Sigma Y_t Y_t'$ converges to a singular limit so the variances of the $n - k$ common trends correspond to principal components giving the nonzero eigenvalues of the limit matrix. The k principal components with smaller variances correspond to the zero eigenvalues of the limit matrix and each of the k rows is a cointegrating vector.

A Note about Distributions

Having more than one lag in the model introduces matrices of nuisance parameters (that is, parameters which must be estimated but are not of primary interest) that affect the form of the test statistics. How the existence of these parameters complicates the test procedures is illustrated for the univariate case.[29] Consider a simple version of the univariate model given by equation 5:

$$(26)\ (y_t - \mu) = (\varrho + \lambda)(y_{t-1} - \mu) - \lambda\varrho(y_{t-2} - \mu) + e_t.$$

If $\varrho = 1$, μ drops out. The series y_t is a unit root process with no tendency to move toward any

level. In this case, however, the test of the null hypothesis that $\varrho = 1$ is complicated by the presence of a nuisance parameter, λ. Regardless of ϱ, as noted before, the model can be reparameterized as

$$(27)\ \Delta y_t = -(\varrho - 1)(\lambda - 1)(y_{t-1} - \mu) + \lambda\varrho\Delta y_{t-1} + e_t.$$

Notice that the coefficient on $y_{t-1} - \mu$ is now a *multiple* of $\varrho - 1$. In ordinary regression, multiplying a regressor by a constant changes the distribution of the regression coefficient by a multiplicative constant, but does *not* change the distribution of its t-statistic. The same holds true asymptotically in cointegration.[30]

If λ is known, the estimated coefficient on $(y_{t-1} - \mu)$ could be divided by $(1 - \lambda)$ and the limit distribution $n(\hat{\varrho}_\mu - 1)$ listed in Fuller (1976, section 8.5) could be obtained. If λ is unknown, however, an approximation must be obtained by dividing by $(1 - \hat{\lambda})$, where $\hat{\lambda}$ is estimated by regressing Δy_t on Δy_{t-1}, suggested by imposing the null hypothesis $\varrho = 1$.

Alternatively, the univariate model could be written as

$$(28)\ \Delta(y_t - \lambda y_{t-1}) = (\varrho - 1)(y_{t-1} - \lambda y_{t-2} - \mu(1 - \lambda)) + e_t.$$

Since $\hat{\lambda}$ is a consistent estimate of λ, it can be used to filter y_t. That is, $\hat{F}_t = y_t - \hat{\lambda}y_{t-1}$. $\Delta\hat{F}_t$ can then be regressed on $\hat{F}_{t-1}$ (with an intercept) or on $F_{t-1} - \bar{F}$, where $\bar{F}$ is the mean of the series, $\hat{F}_t$, to get a test statistic asymptotically equivalent to $n(\hat{\varrho}_\mu - 1)$.

Other Approaches to Cointegration

The approach of Fountis and Dickey (1989) is similar to that of Stock and Watson but only allows for the possibility that there is one unit root. As such, it is much less general than either of the above approaches. It does, however, provide estimates of the cointegrating vectors and common trend, based on the coeffi-

---

[28]For example, let X be an n by k matrix and let J = (X'X) be a k by k matrix of rank k. Then there exists a k by k matrix T such that T' T = Λ, where Λ is a diagonal matrix with the eigenvalues of J on the diagonal. The columns of T are the eigenvectors of J. These eigenvectors are called "principal components" because the generalized variance of J, |J|, is equal to the trace of Λ. Because often many of the eigenvalues of J are close to zero, its generalized variance can be closely approximated by a relatively few "principal components." A good discussion of this can be found in Dhrymes (1970), pp. 53-59.

[29]These ideas can also be found in Fuller's text (1976) or in Dickey, Bell and Miller (1986).

[30]Johansen, using a multivariate analogue of the Dickey-Fuller τ statistic, only needs to have a correctly specified likelihood function. Stock and Watson, using a multivariate analogue of the normalized coefficient tests $n(\hat{\varrho} - 1)$, $n(\hat{\varrho}_\mu - 1)$, or $n(\hat{\varrho}_\tau - 1)$ of Dickey and Fuller need to adjust for this fact multiplier (corresponding to $1 - \lambda$ in our univariate example).

cient matrix of the lagged levels. Although there is an asymptotic link between variances and lag coefficients as established above, the actual definition of a unit root process is in terms of lag relationships. All approaches to cointegration use lags in *testing*. However, in *transforming* Y_t, Stock and Watson use only the variance-covariance matrix while Fountis and Dickey (and Johansen) use the lag information.

Finally, we mention an approach by Engle and Granger which is especially easy to use and to explain in the bivariate case. Consider two univariate series y_t and z_t. The first step is to check that each is a unit root process. Next regress y_t on z_t (or z_t on y_t) getting the linear combination $y_t - \hat{b}z_t$ with smallest variance. Notice that this does not use lag information to obtain the transformation $y_t - \hat{b}z_t$. The next step is to test if the series $y_t - \hat{b}z_t$ is stationary (if so, the vector $(1, -\hat{b})$ is a cointegrating vector). Had the parameter b been pre-specified instead of estimated, an ordinary Dickey-Fuller test would be appropriate. Since the data are used to estimate b, however, the previously mentioned problems analogous to the "multiple comparisons" problem exist. Engle and Granger have provided the appropriately adjusted percentiles. Other than using these special tables, this approach uses only the standard univariate unit root testing strategy.

One problem with this approach is that it requires the researcher to choose one of the jointly endogenous variables to put on the left-hand side. While the test is asymptotically invariant to this so-called direction normalization rule, the test results may be very sensitive to it in finite samples. Indeed, practical experience indicates that the result of the test depends qualitatively on which variable is chosen to be on the left-hand side. The alternative multivariate approaches have the advantage that all of the variables are explicitly endogenous, so the researcher does not have to make such arbitrary normalization choices.

AN APPLICATION OF COINTEGRATION: THE DEMAND FOR MONEY

One important macroeconomic relationship that has received considerable attention is the link between money and income. This relationship, embedded in the demand for money, is commonly represented by the income velocity of money. Since the income velocity of M1 has drifted upward over time, it does not appear to be stationary. Furthermore, formal tests indicate that the income velocity of M1 is not stationary, indicating that M1 and income, in the form of income velocity, are not cointegrated.[31] In contrast, the income velocity of M2 appears to move around an unchanged mean, and formal tests suggest that M2 and income are cointegrated.[32] These results have been interpreted by some as evidence against the existence of a stable long-run demand for M1 and for the existence of a stable long-run demand for M2.[33]

But if M1 is the relevant measure of "money," the Fisher equation suggests that there should be at least one cointegrating relationship between M1, its velocity, real income and the price level. The specification of tests for cointegration depend on the specification of the demand for money. Consequently, it is important to review the theory of money demand before performing tests for cointegration.

A general specification for the long-run demand for money is

$$(29) \quad M^d = f(P, Q, Z),$$

where M and Q denote the nominal money stock and the nominal income level, respectively, P denotes the level of prices and Z denotes all other variables that affect money demand—for example, current and expected future real interest rates, the expected rate of inflation, etc. Assuming that economic agents do not suffer

---

[31]Strictly speaking, the nonstationarity indicates only that the cointegrating vector for M1 and nominal GNP is not (1, -1). For example, see Nelson and Plosser (1982). Also, see Engle and Granger (1987) who test for cointegration between M1 and income.

[32]For example, see Engle and Granger (1987). However, the results for M2 appear to be sensitive to the sample period and how the test is performed. See Hallman, Porter and Small (1989, 1990) and Hafer and Jansen (1991).

[33]For example, Hallman, Porter and Small (1989) have predicated their P-star model on such a long-run stable demand for M2.

from a money illusion, equation 29 can be written as

(30) $M^d/P = m^d = f(Q/P, Z) = f(q, Z)$.

That is, the demand for real money balances, m^d, is a function of real income, q, and some other variables. Furthermore, it is commonly assumed that the demand for money is homogenous of degree one in real income, so that equation 30 can be written as

(31) $m^d/q = h(Z)$,

where $h(Z)$ is the famous k in the Cambridge cash balance equation—that is, the reciprocal of the income velocity of money. In equilibrium, the demand for real money equals the supply of real money, m^s, so that $h(Z)$ is observed simply as the ratio of real money stock to real income.

The Velocity of M1 and M2

Consider now two alternative monetary aggregates, M1 and M2. In this framework, the reciprocals of their velocities can be written as

(32) $m1/q = h1(Z)$, and

(33) $m2/q = h2(W)$,

respectively. The specification for M2 allows for the possibility that there are factors that affect its demand that do not affect the demand for M1—that is, Z is a subset of W. Since M2 is simply M1 plus some other financial assets, equation 33 can be written as

(34) $m1/q + nm1m2/q = h1(Z) + v(W)$,

where nm1m2 denotes the real non-M1 components of M2 and $v(W) = h2(W) - h1(Z)$, hereafter, called the reciprocal of NM1M2 velocity.[34]

The above analysis has two important implications. First, because velocity is not directly and independently observable, its proxy must be specified to perform tests for cointegration. Second, if, in fact, M1 and income are not cointegrated but M2 and income are, a long-run, stable inverse relationship must exist between h1(Z) and v(W). On average, movements in h1(Z) must be offset by movements in v(W), that is, h1(Z) and v(W) must be cointegrated.

A simple analysis of M1, M2 and GNP data is consistent with this conjecture. Figures 1 and 2 show the observed income velocities of M1 and M2—that is, (q/m1) and (q/m2)—respectively, for the period from 1953.1 to 1988.4. M1 velocity trends upward through the early 1980s and then declines. In contrast, M2 velocity appears to cycle with no apparent trend.[35] Also, the sharp break in the pattern of M1 velocity in the early 1980s is not as apparent in M2 velocity.

Figure 3 shows the reciprocals of M1, M2 and NM1M2 velocities. As expected, the downward trend in the reciprocal of M1 velocity through 1980 appears to be matched by an upward trend in the reciprocal of the velocity of NM1M2.[36] Also, a comparison of the series in figure 3 reveals that much of the variability in the reciprocal of M2 velocity is associated with variability in the reciprocal of NM1M2 velocity, rather than variability of the reciprocal of M1 velocity.[37]

---

[34]These relationships hold for seasonally adjusted M2 as well, because M1 and the non-M1 components of M2 are seasonally adjusted separately and added together.

[35]Dickey-Fuller tests cannot reject the null hypothesis of a unit root over some sample periods. Figure 2 suggests that these results are likely driven by the low power of the test against the alternative of a large, but stationary, root. Hence, one might wish to rely on his eyes rather than the formal test results.

[36]At a more formal level, the proposition that the decline in reciprocal of M1 velocity is just offset by the rise in the reciprocal of the velocity of NM1M2 was tested by a simple linear regression the reciprocal of NM1M2 on the reciprocal of M1 velocity and testing the null hypothesis that the coefficient is equal to -1. The estimated slope coefficient was -1.055 with a t-statistic of -41.47. While the estimated coefficient was very close to -1, the null hypothesis was rejected at the 5 percent significance level. The t-statistic was 2.17. Nevertheless, a formal test for cointegration using an augmented Dickey-Fuller test on the residuals from this equation suggests that these variables are cointegrated. This proposition also was tested by

estimating a simple linear time trend for the reciprocals of the velocities of each M1 and NM1M2 and testing the hypothesis that trends are equal and opposite in sign. The estimated trend coefficients for M1 and NM1M2 were $-.00143$ and $.00154$, respectively, and both were statistically significant at well below the 5 percent level. But, again, despite the closeness of the estimates, the null hypothesis was rejected at the 5 percent significance level. The F-statistic was 8.67.

[37]This can be illustrated by a simple linear regression of the reciprocal of M2 velocity on the reciprocals of the velocities of each M1 and NM1M2. These regressions indicate that the reciprocal of the velocity of NM1M2 explains 19 percent of the variation in the reciprocal of the velocity of M2, while the reciprocal of M1 velocity alone explains only .03 percent of the variation.
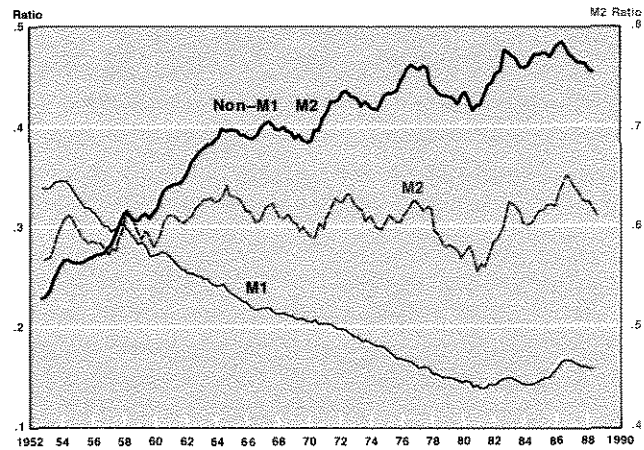
Figure 1
The Income Velocity of M1
I/1953 to IV/1988
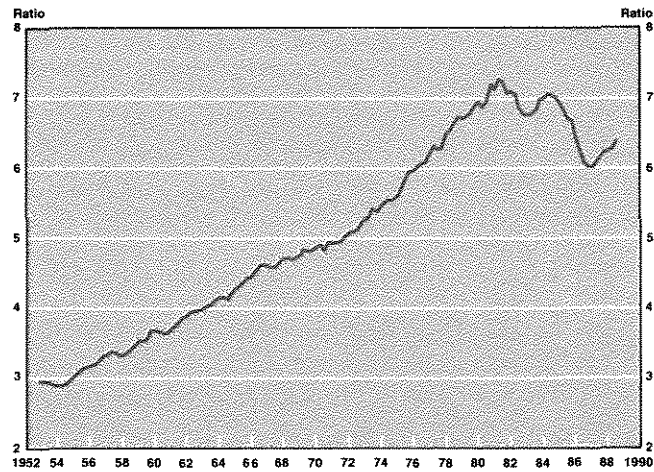
Figure 2
The Income Velocity of M2
I/1953 to IV/1988
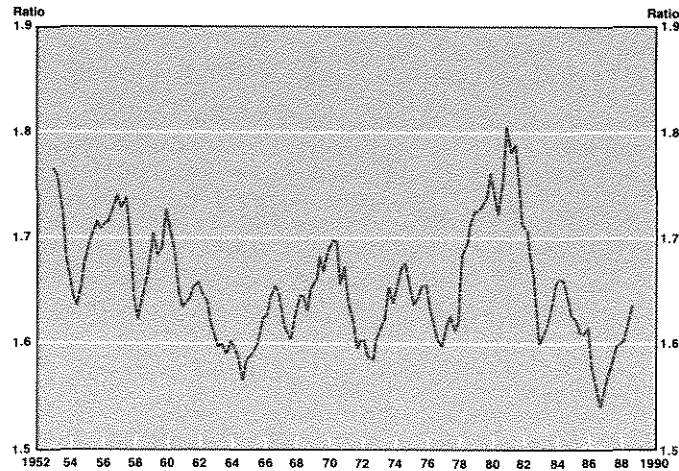
Figure 3
**The Reciprocals of the Income Velocities of M1,
M2 and the Non–M1 Components of M2**
I/1953 to IV/1988

The Velocity of the Monetary Base

The concept of income velocity can be extended directly to a broader or narrower range of monetary aggregates. One such aggregate, the monetary base, is particularly important because the monetary authority can control it fairly well. While the observed monetary base velocity is simply the ratio of income to the monetary base, the measure is only meaningful in the context of the demand for money.[38] Monetary base velocity can be incorporated into the money demand framework by noting that the nominal money supply, M^s, can be expressed as

$$(35) \quad M^s = mm(H)MB,$$

where MB denotes the adjusted monetary base and mm denotes the money multiplier which is a function of a set of variables H, for example, interest rates, portfolio preferences of the public, etc. Dividing both sides of equation 35 by the price level and substituting the result for m^d in equation 31, yields

$$(36) \quad mm(H)mb = qh(Z),$$

where mb denotes the real monetary base. Because the money multiplier is not observed independently, tests of cointegration involving the monetary base must include not only q and Z, but H.[39]

Empirical Results

The empirical work is presented in two parts. The first presents tests for cointegration using methodologies suggested by Johansen, Stock-Watson and Engle-Granger. The results in this part are presented only for M1. In the second part, the analysis is extended to a broader set of monetary aggregates using the Johansen approach. The Johansen methodology was chosen because it is based on the well-accepted likeli-

hood ratio principle. Moreover, recent Monte Carlo evidence by Gonzalo (1989) suggests that Johansen's maximum likelihood technique for estimating and testing cointegrating relationships performs better than both single equation methods and alternative multivariate methods. Nevertheless, because the other two approaches are widely known and the Engle-Granger approach is especially widely used in the empirical literature to date, we report test results for M1 using all three techniques.

In the second part, the analysis is extended to other monetary aggregates. These aggregates are the adjusted monetary base (calculated by the Federal Reserve Bank of St. Louis), M2 and the non-M1 component of M2, denoted NM1M2. When the monetary base is used, the ratio of currency to total checkable deposits, denoted K, is also included because it is the most important determinant of the money multiplier.[40] In both parts, the income and price level measures are real GNP, q, and the GNP deflator, P, respectively. Two measures of nominal interest rates, R, are used: the three-month Treasury bill rate, R3M, and the yield on 10-year government securities, R10Y. The data consist of quarterly observations from 1953.2 to 1988.4 and all data are transformed to natural logarithms.

Tests for the Order of Integration: Before testing for cointegration, the order of integration of the individual time series must be determined. Tests for unit roots are performed on all of the data using the augmented Dickey-Fuller test with three lagged differences. The null hypothesis is that the variable under investigation has a unit root, against the alternative that it does not. The substantially negative values of the reported test statistic lead to rejection of the null hypothesis.

The tests are performed sequentially. The first column in the top half of table 1 reports tests

---

[38]For example, it is not reasonable to obtain this result by simply assuming that the monetary base is the appropriate measure for money because it is composed of currency and bank reserves.

[39]It is not necessary that H include variables that are not included in Z. If it does not, however, there is an identification problem. That is, one cannot tell the difference between the above model and simply treating the monetary base as the appropriate monetary aggregate. Since this possibility is difficult to conceive of, it is useful if H includes variables that are not in Z, as in our empirical work which follows.

[40]A simple linear regression of the multiplier (or its growth rate) on K (or its growth rate) using quarterly data, indi-

cates that K alone explains 95 percent of the variation in the level of the multiplier. Moreover, the growth rate of K alone explains 84 percent of the variation in the growth rate of the multiplier.

Table 1

Augmented Dickey-Fuller Test For A Unit Root

Variable	t_μ 3 lags	t_τ 3 lags
M2/P	−0.801	−2.456
M1/P	−0.817	−1.444
MB/P	0.308	−2.285
q	−0.723	−2.167
R3M	−2.689	−3.884*
R10Y	−1.873	−2.444
K	−0.561	−2.274
NM1M2/P	−2.150	−1.580
Δ(M2/P)	−4.006*	
Δ(M1/P)	−3.618*	
Δ(MB/P)	−3.067*	
Δq	−5.923*	
ΔR3M	−6.634*	
ΔR10Y	−5.586*	
ΔK	−5.127*	
Δ(NM1M2/P)	−3.855*	

*Indicate statistical significance at the 5 percent level. Critical values t_μ (T = 100) = −2.89, t_τ (T = 100) = −3.45.

of stationarity of the levels of the time series about a non-zero mean. The critical values of the test statistic [t_μ] are tabulated in Fuller (1976) and discussed in Dickey and Fuller (1979). The reported test statistics indicate that the null hypothesis cannot be rejected for any variable. We then test for stationarity about a deterministic time trend, using the Dickey-Fuller statistic [t_τ]. The results of this test are given in the second column in the top half of table 1. Critical values for this test statistic are tabulated in Fuller (1976). With the exception of R3M, the null hypothesis that the time series has a unit root cannot be rejected.

The bottom half of table 1 reports results for the augmented Dickey-Fuller test on first differences of the variables. The null hypothesis of a unit root is rejected for all of the time series using differenced data. These results are broad-

ly consistent with the hypothesis that the individual time-series are individually I(1). Because these data appear to be stationary in first differences, no further tests are performed.

Tests for Cointegration Using Three Methodologies: Tests for cointegration for real M1, real income and either R3M or R10Y using methodologies proposed by Johansen, Stock-Watson and Engle-Granger are presented in table 2. For the Johansen and Engle-Granger tests, three lagged differences were used. Both the test statistics and the estimated cointegrating vector (setting the coefficient on M1/P equal to one) are reported. The estimated cointegrating vector is reported even when the test does not indicate cointegration.

For the Johansen method, there are two test statistics for the number of cointegrating vectors: the trace and maximum eigenvalue statistics. In the trace test, the null hypothesis is that the number of cointegrating vectors is less than or equal to k, where k is 0, 1 or 2. In each case the null hypothesis is tested against the general alternative. The maximum eigenvalue test is similar, except that the alternative hypothesis is explicit. The null hypothesis k = 0 is tested against the alternative that k = 1, k = 1 against the alternative k = 2, etc. The critical values for these tests are tabulated by Johansen and Juselius (1990). For the trace test, the hypotheses k ≤ 1 and k ≤ 2 cannot be rejected for either of the two interest rates, while the hypothesis k = 0 can be rejected.[41] Consequently, we conclude that there is one cointegrating vector.

Turning to the maximum eigenvalue test, the hypothesis k = 0 is uniformly rejected in favor of the alternative k = 1. Consequently, this test indicates that real M1 is cointegrated with real income and either of the two nominal interest rates. Moreover, there appears to be a single cointegrating vector. The maximum eigenvalue test of k = 1 vs. k = 2 fails to reject the null hypothesis of k = 1. Thus, there are two common trends and one cointegrating vector.

The Johansen test produced results that were markedly different from those obtained using

[41]Johansen and Juselius (1990) note, however, "One would, however, expect the power of this procedure [the trace test] to be low, since it does not use the information that the last three eigenvalues have been found not to differ significantly from zero. Thus one would expect the maximum eigenvalue test to produce more clear cut results," (p.19).

Table 2
Tests for Cointegration for M1

Test Statistics					Cointegrating Vector		
					q	R3M	R10Y
Johansen test[1]							
Trace			Max. Eigenvalue				
k = 0	k ≤ 1	k ≤ 2	k = 0	k = 1			
32.3*	5.2	1.2	27.1*	4.0	.680	− .369	—
28.4	6.0	1.8	22.4*	4.2	.845	—	− .570
Stock-Watson test[2]							
q(3,2)							
− 24.91					.398	− .152	—
− 15.61					.635	—	− .353
Engle-Granger test[3]							
Dependent Variable							
(M1/P) − 2.26					.353	− .131	—
− 2.08					.558	—	− .303
q − 3.92					.671	− .271	—
− 3.18					.794	—	− .449
R − 4.86*					.730	− .359	—
− 3.36					.826	—	− .496

*Indicates statistical significance at the 5 percent level.

[1]Critical values:

	Trace			Max. Eigenvalue	
	k = 0	k ≤ 1	k ≤ 2	k = 0	k = 1
	31.3	17.8	8.1	21.3	14.6

[2]Critical value for q(3,2) is − 31.5.

[3]Critical values for ADF taken from Engle and Yoo (1987)

100 observations	200 observations
− 3.93	− 3.78

either the Engle-Granger or Stock-Watson methodologies. Because the results of the Engle-Granger can change with the variable chosen as the dependent variable, the test was performed with each variable on the right-hand side. Not surprisingly, the test was sensitive to this choice. The test indicated cointegration only when R3M was the dependent variable (although the test indicated cointegration at slightly higher significance level if real GNP is the dependent variable and R3M is used). It is interesting to note, however, that the estimated cointegrating vectors obtained from the Johansen and Engle-Granger approaches are nearly identical when both indicate cointegration. This is not the case, however, when the Engle-Granger or Stock-Watson tests do not indicate cointegration.[42]

---

[42]The estimated cointegrating vector from the Engle-Granger approach when the equation is normalized on real output is (M1/P) = − .271R + .671Q. While the estimated coefficient on the three-month Treasury bill rate is markedly different from that when cointegration is indicated, the estimated coefficient on output is nearly identical.

Table 3
Tests for Cointegration for the Broader Monetary Aggregates

Aggregate	Trace Statistic			Max. Eigenvalue	
	k = 0	k ≤ 1	k ≤ 2	k = 0	k = 1
M2	33.7*	13.9	1.6	19.8	12.3
	38.7*	17.3	4.4	21.4*	12.9
NM1M2	53.4*	20.6*	2.7	32.8*	17.9*
	39.7*	16.3	3.7	23.4*	12.6

Critical Values	Trace			Max. Eigenvalue	
	k = 0	k ≤ 1	k ≤ 2	k = 0	k = 1
	31.3	17.8	8.1	21.3	14.6

Although evidence concerning cointegration between real M1, interest rates and output is sensitive to the method used, the results using Johansen methodology are similar to those of Hoffman and Rasche (1989) using monthly data and the same methodology. Moreover, for the Johansen results, the hypothesis that the normalized coefficient on output is unity is not rejected using either interest rate, while the hypothesis of a zero coefficient for the interest rate is rejected for both interest rates (see table 4).[43] Thus, the data appear to support the notion that there is a stable long-run relationship between real M1, real income and interest rates.

Cointegration Using Alternative Monetary Aggregates: Tests for cointegration using M2 and NM1M2 are presented in table 3. Both the trace and maximum eigenvalue tests indicate one cointegrating vector for M2, and both indicate cointegrating vectors for NM1M2.[44] The previous discussion of the relationship between M1, M2 and NM1M2 suggests that certain long-run relationships should exist between the estimated cointegrating vectors using the various aggregates. In order to examine this suggestion, the estimated cointegrating vectors normalized on the real value of the respective aggregate are presented in table 4, along with

tests of the hypotheses that the coefficient on output is unity and the coefficient on the interest rate is zero. These test statistics are asymptotically distributed $\chi^2(1)$.

The apparent nonstationarity of M1 velocity and the stationarity of M2 velocity implies that M1 and NM1M2 must have compensating nonstationary behavior. This suggests that the sum of the income elasticities or interest elasticities for M1 and NM1M2 should equal that of M2. While there are no formal tests of these cross-equation restrictions, the point estimates in table 4 indicate that, with the exception of the income elasticity when R3M is used, these restrictions do not do too much violence to the data. For the three-month rate, the sum of the interest elasticities for M1 and NM1M2 is −.15, compared with the estimated elasticity for M2 of .02; the sum of the income elasticities is 1.72, compared with an estimated income elasticity for M2 of .97. For the 10-year rate, the sum of the elasticities is −.08, compared with the estimated elasticity of −.03; the sum of the income elasticities is .92, compared with an estimated income elasticity of 1.04. Nevertheless, the fact that the hypothesis that the income elasticity is unity cannot be rejected for either M1 or M2 is troubling.

[43]The $\chi^2(1)$ statistics for the test of the coefficient on output are 2.23 and .61, respectively, for the three-month and 10-year rates, and the test statistics for the interest rates are 21.37 and 14.66, for the two rates, respectively.

[44]If a 10 percent significance level is used, the test indicates there are two cointegrating vectors for M2 when the 10-year bond rate is used.

Table 4
Normalized Cointegrating Vectors and Hypothesis Tests

Aggregate	Cointegrating Vector			Hypothesis Test		
	q	R3M	R10Y	q^1	$R3M^2$	$R10Y^2$
M1	0.680	−0.369	—	2.23	21.37*	—
	0.845	—	−0.570	0.61	—	14.66*
M2	0.971	0.016	—	0.20	0.16	—
	1.044	—	−0.031	0.21	—	0.24
NM1M2	1.043	0.217	—	0.09	8.58*	—
	0.078	—	0.646	2.72	—	4.30*

*Indicates statistical significance at the 5 percent level.
[1]Null hypothesis is the income elasticity is one.
[2]Null hypothesis is the interest elasticity is zero.

Cointegration and the Monetary Base

As was noted earlier, the monetary base must be regarded as a supply-side variable, and cointegration of the monetary base with income and interest rates arises due to the relationship between the monetary base and the relevant money stock measure. Consequently, it is necessary to include a proxy for the money multiplier in an investigation of cointegration for the monetary base. Because the primary determinant of the multiplier is the currency-deposit ratio, K, it is included along with the real monetary base, real income and an interest rate in tests for cointegration.

The results, presented in table 5, indicate that there are two cointegrating vectors linking the real monetary base, real income, the nominal interest rate and K when R3M is used, but only one cointegrating vector when R10Y is used. Because a cointegrating vector merely represents a long-run, stable relationship among jointly endogenous variables, in general, they cannot be interpreted as structural equations. Consequently, neither of the estimated cointegrating vectors necessarily represents either the long-run demand for or long-run supply of money. All that can be said is that there are two linear combinations for which the variance is bounded. Nevertheless, it is interesting to note that the second reported cointegrating vector is broadly consistent with equation 36.[45]

Nevertheless, because a stable long-run demand for money implies that there is a stable long-run relationship between real money, real income and either a short- or long-term interest rate. Consequently, these results are consistent with the proposition that the long-run demand for money is stable, even though they may not be estimates of the long-run money demand function itself.[46]

They also suggest that the reason M2 velocity is stable is because it includes transactions and

---

[45]Taking the log of equation 36 and letting the money multiplier be a function of both K and the interest rate and h(Z) be solely a function of the interest rate, results in ln mb = ln q + ln h(R) − ln mm (R,K), where h′< 0 and ∂ mm/∂R > 0 and ∂ mm/∂K < 0.

This equation implies that the long-run elasticity of ln mb with respect to ln q is unity, the elasticity of ln mb with respect to R is negative, but, smaller than the estimate for the long-run demand for money, and that the elasticity with respect to K is negative.

[46]These results are both quantitatively and qualitatively similar to those obtained by Hoffman and Rasche (1989).

Table 5
Tests for Cointegration Using the Monetary Base

Trace test				Max. Eigenvalue			Cointegrating vector			
k=0	k≤1	k≤2	k≤3	k=0	k=1	k=2	q	K	R3M	R10Y
65.5*	30.6	7.4	2.3	34.9*	23.2*	5.1	.465	.333	-.271	—
							.883	-.278	-.204	—
52.6*	22.5	5.3	1.4	30.1*	17.2	3.9	1.239	-.540	—	-.269

*Indicates statistical significance at the 5 percent level.

non-transactions components that are close substitutes for each other in the long run. In particular, the upward drift in M1 velocity appears to be largely due to a relatively steady shift from M1 to the non-transactions deposits in M2. The magnitude of the trend movements in these variables is approximately equal so that M2 velocity is essentially trendless over the estimation period.[47]

SUMMARY AND CONCLUSIONS

This paper reviews the concept of cointegration, notes the relationship between tests for it and common tests for unit roots and considers its implications for the relationship among real money balances, real income and nominal interest rates. We argue that if M2 and nominal income are cointegrated, while M1 and nominal income are not, there necessarily exists a stationary long-run relationship between M1 and the non-M1 components of M2. We also argue that, if M1, real income and the nominal interest rate are cointegrated, the same could be true for real income, the nominal interest rate, the monetary base and a proxy for the monetary base/money multiplier.

Tests for cointegration among real M1, real income and one of two interest rates using three alternative procedures show that the results are sensitive to the method used. Nevertheless, the technique proposed by Johansen indicates that there is a single cointegrating relationship among these variables. While the cointegrating vector cannot be interpreted as the long-run demand for money, the estimated long-run income and interest elasticities are consistent with those often hypothesized and estimated for the long-run demand for money.

We also show that the hypothesized long-run relationship for the cointegrating vectors for M1, M2 and the non-M1 components of M2, namely that the sum of the income and interest elasticities for M1 and the non-M1 components of M2 equal the income and interest elasticities of M2, is supported by the data. Finally, we show that if the currency-deposit ratio is used to proxy the monetary base multiplier, the real monetary base, real income, the interest rate and the currency-deposit ratio are cointegrated.

The last two results are consistent with the notion of a stable long-run relationship between monetary aggregates and prices when both real income and nominal interest rates are taken into account. Moreover, that there appears to be a stable long-run relationship between real money, real income and nominal interest rates establishes the potential for achieving price level stability by controlling the growth rates of either M1 or the monetary base.

---

[47]The stable long-run relationship between real income, the real monetary base, nominal interest rates and the currency-deposit ratio, is also consistent with the idea recently put forth by McCallum (1987) that nominal GNP can be controlled in the long-run by monetary base targeting.

REFERENCES

Bernanke, Ben S. "Alternative Explanations of the Money-Income Correlation," in *Carnegie-Rochester Conference on Public Policy*, K. Brunner and A. Meltzer, eds., (Autumn 1986), pp. 49-100.

Blanchard, Oliver Jean, and Danny Quah. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review* (September 1989), pp. 655-73.

Dhrymes, Phoebus J. *Econometrics: Statistical Foundations and Applications* (Harper and Row, 1970).

Dickey, David A., William R. Bell and Robert B. Miller. "Unit Roots in Time Series Models: Tests and Implications," *American Statistician* (February 1986), pp. 12-26.

Dickey, David A., and Wayne A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* (June 1979), pp. 423-31.

Dickey, David A., and Robert J. Rossana. "Cointegrated Time Series: A Guide to Estimation and Hypothesis Testing," Manuscript, N.C. State University, Department of Statistics and Economics (October 1990).

Engle, Robert F., and C. W. J. Granger. "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica* (March 1987), pp. 251-76.

Engle, Robert F., and Byung Sam Yoo. "Forecasting and Testing in Co-integrated Systems," *Journal of Econometrics* (May 1987), pp. 143-59.

Fountis, Nicolaos G., and David A. Dickey. "Testing For a Unit Root Nonstationarity in Multivariate Autoregressive Time Series," *Annals of Statistics* (1989), pp. 419-28.

Friedman, Milton. "The Quantity Theory of Money—A Restatement," in *Studies in The Quantity Theory of Money* (University of Chicago Press, 1956), pp. 3-21.

Fuller, Wayne A. *Introduction to Statistical Time Series* (John Wiley & Sons, Inc., 1976).

Gonzalo, Jesus. "Comparison of Five Alternative Methods of Estimating Long-Run, Equilibrium Relationships," Discussion Paper 89-55, University of California at San Diego, 1989.

Hafer, R.W., and Dennis W. Jansen. "The Demand for Money in the United States: Evidence from Cointegration Tests," *Journal of Money, Credit and Banking* (May 1991), pp. 155-68.

Hallman, Jeffrey J., Richard D. Porter and David H. Small. "M2 per Unit of Potential GNP as an Anchor for the Price Level," Board of Governors of the Federal Reserve System Staff Study No. 157, April 1989.

_____ . "Is the Price Level Tied to the Stock of M2 in the Long Run?" Unpublished manuscript (February 1990).

Hoffman, Dennis, and Robert H. Rasche. "Long-Run Income and Interest Elasticities of Money Demand in the United States," NBER Working Paper No. 2949 (April 1989).

_____ . "Identification and Inference in Cointegrated Systems: A Synthesis of Recent Developments," Working Paper, February 1990.

Johansen, Soren. "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* (1988), pp. 231-54.

Johansen, Soren, and K. Juselius. "Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics* (May 1990), pp. 169-210.

McCallum, Bennett T. "The Case for Rules in the Conduct of Monetary Policy: A Concrete Example," Federal Reserve Bank of Richmond *Economic Review* (September/October 1987), pp. 10-18.

Nelson, Charles R., and Charles I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* (September 1982), pp. 139-62.

Phillips, P. C. B. "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika* (September 1987), pp. 535-47.

Schwert, G. William. "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data," *Journal of Monetary Economics* (July 1987), pp. 73-104.

Sims, Christopher A. "Bayesian Skepticism on Unit Root Econometrics," Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, Discussion Paper No. 3 (May 1988).

Sims, Christopher A., and Harold Uhlig. "Understanding the Unit Rooters: A Helicopter Tour," Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, Discussion Paper No. 4 (June 1988).

Steel, Robert George Douglas, and James H. Torrie. *Principles and Procedures of Statistic: A Biometrical Approach*, 2nd Ed. (McGraw-Hill, 1980).

Stock, James H., and Mark W. Watson. "Testing for Common Trends," *Journal of the American Statistical Association* (December 1988), pp. 1097-1107.

_____ . "A Simple MLE of Cointegrating Vectors in Higher Order Integrated Systems," Unpublished manuscript (November 1989).

Theil, H., and J. C. G. Boot. "The Final Form of Econometric Equation Systems," *Review of the International Statistical Institute* (1962), pp. 136-52.