Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 415

November 1991

COINTEGRATION, EXOGENEITY, AND POLICY ANALYSIS: AN OVERVIEW

Neil R. Ericsson

ABSTRACT

This overview describes the concepts of cointegration and exogeneity, focusing on analytical structure, statistical inference, and implications for policy analysis. Examples help clarify the concepts. The remainder of the overview summarizes the articles in a special issue of the *Journal of Policy Modeling* entitled *Cointegration, Exogeneity, and Policy Analysis.*

Key words and phrases: cointegration, conditional models, error-correction models, exogeneity, parameter constancy, policy analysis, predictive accuracy.

Cointegration, Exogeneity, and Policy Analysis: An Overview

Neil R. Ericsson[1]

1. Introduction

The concept of "cointegration" is a central development in the econometric literature over the last decade. Introduced by Granger (1981) and Engle and Granger (1987), cointegration is a statistical property which may describe the long-run behavior of economic time series. Importantly, cointegration ties together several apparently disparate fields. First, cointegration links the economic notion of a long-run relationship between economic variables to a statistical model of those variables. If a long-run relationship exists, the variables involved are "cointegrated". Second, the technical and previously somewhat obscure statistical theory on unit-root processes provides the basis for statistical inference about the empirical existence of cointegration. Third, cointegration implies and is implied by the existence of an error-correction representation of the relevant variables. Thus, cointegration establishes a firmer statistical and economic basis for the empirically successful error-correction models. Fourth, via error-correction models, cointegration brings together short- and long-run information in modeling the data. That unification resolves the "debate" on whether to use levels or differences, with Box-Jenkins-type time-series models and classical "structural" models both being special cases of error-correction models. Fifth, via the distributional theory of integrated processes, cointegration clarifies the "spurious regressions" or "nonsense-correlations" problem associated with trending time-series data.

Cointegration by itself is proving a useful conceptual and empirical tool. However, to derive implications of cointegration for policy analysis, it is fruitful to introduce a second concept, exogeneity.

Virtually concurrent to developments in cointegration, Richard (1980), Engle, Hendry, and Richard (1983), and Florens and Mouchart (1985a, 1985b) clarified and refined the

[1] This paper is the introduction to a special issue of the *Journal of Policy Modeling* entitled *Cointegration, Exogeneity, and Policy Analysis*, forthcoming in 1992. I wish to thank George Tavlas for inviting me to edit this special issue, and for his support and guidance throughout its preparation. All papers for the special issue were refereed anonymously and subjected to the usual standards of acceptance for the *Journal*. Each referee was provided with the author's data, and this led to substantive, fruitful exchanges on empirical aspects of the papers. I thank the authors and referees for their promptness under a tight editorial and publishing schedule.

The author of this paper is a staff economist in the Division of International Finance, Federal Reserve Board. The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting those of the Board of Governors of the Federal Reserve System or other members of its staff. Helpful discussions with and comments from Lisa Barrow, Jim Boughton, Julia Campos, Jonathan Eaton, Jon Faust, David Hendry, Søren Johansen, Katarina Juselius, Ed Leamer, Jaime Marquez, Will Melick, and Hong-Anh Tran are gratefully acknowledged. I am indebted to Lisa Barrow for invaluable research assistance.

concept of exogeneity, building on Koopmans (1950) and Barndorff-Nielsen (1978). Critically, the exogeneity of a variable depends on the parameters of interest to the investigator, and on the purpose of the model, whether for statistical inference, forecasting, or policy (scenario) analysis. These three purposes define three types of exogeneity, which Engle, Hendry, and Richard call weak, strong, and super.

A clear understanding of exogeneity is critical for analyzing the implications of cointegration for policy analysis, so this overview discusses exogeneity first (Section 2), then cointegration (Section 3). Each section considers the analytical structure of the associated concept, statistical inference, and policy implications, in that order. Examples en route help demonstrate the roles of exogeneity and cointegration in policy analysis. The issues in Sections 2 and 3 are discussed at greater length in Ericsson, Campos, and Tran (1991), itself summarizing David Hendry's empirical econometric methodology; cf. Hendry and Richard (1982, 1983), Hendry (1983), Spanos (1986), and Hendry (1987, 1989, 1990). Section 4 summarizes the articles in this special issue. The Appendix lists the titles of those articles.

2. Exogeneity

Whether or not a variable is exogenous depends upon whether or not that variable can be taken as "given" without losing information for the purpose at hand. The distinct purposes of statistical inference, forecasting, and policy analysis define the three concepts of weak, strong, and super exogeneity. Valid exogeneity assumptions may permit simpler modeling strategies, reduce computational expense, and help isolate invariants of the economic mechanism. Invalid exogeneity assumptions may lead to inefficient or inconsistent inferences and result in misleading forecasts and policy simulations. Section 2.A defines the exogeneity concepts and illustrates them with a static bivariate normal process (Example 1), the well-known cobweb model (Example 2), and a first-order vector autoregression (Example 3). Section 2.B discusses tests of exogeneity and, in particular, tests of super exogeneity, since that concept is relevant for policy analysis. Section 2.C comments on the policy implications of exogeneity. The examples here and in Section 3 generalize straightforwardly to linear multivariate dynamic processes.

A. Concepts and Structure

Weak exogeneity. The essential concept is weak exogeneity, which is required for efficient inference (i.e., estimation and hypothesis testing) in a conditional model. Weak exogeneity can be explained via one of the simplest processes, the bivariate normal. In this first example, the bivariate normal density is factorized into its conditional and marginal densities. Analyzing the conditional density leads to the concepts "parameters of interest" and "variation free", and so to weak exogeneity. Example 2 discusses these concepts in greater detail for the cobweb model.

Example 1: joint, conditional, and marginal densities. Consider two variables, y_t and z_t, which are jointly normally distributed and serially independent:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} \sim IN(\boldsymbol{\mu}, \boldsymbol{\Omega}) \qquad t = 1, \ldots, T. \tag{1}$$

The subscript t denotes time, T is the total number of observations on $(y_t, z_t)'$, the notation "$\sim IN(\boldsymbol{\mu}, \boldsymbol{\Omega})$" denotes "is distributed independently and normally, with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Omega}$", and bold face denotes a vector or matrix, rather than a scalar. In an economic context, y_t and z_t might be money and an interest rate, or consumers' expenditure and income, or wages and prices. Let $\boldsymbol{x}_t$ be $(y_t, z_t)'$, and define $\boldsymbol{\varepsilon}_t$ as the "error" $\boldsymbol{x}_t - \mathcal{E}(\boldsymbol{x}_t)$, which is $\boldsymbol{x}_t - \boldsymbol{\mu}$, where $\mathcal{E}(\cdot)$ is the expectation operator. Then (1) becomes:

$$\boldsymbol{x}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t \sim IN(\mathbf{0}, \boldsymbol{\Omega}). \tag{2}$$

Equation (2) is in "model form", rather than being written directly as a distribution, as in (1). Below, it will be helpful to express $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ explicitly in terms of their scalar elements:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \tag{3a}$$

and

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}. \tag{3b}$$

Without loss of generality, (1) may be factorized into the "conditional" density of y_t given z_t and the "marginal" density of z_t:[2]

$$y_t \mid z_t \sim IN(a + bz_t, \sigma^2) \tag{4a}$$

$$z_t \sim IN(\mu_2, \omega_{22}), \tag{4b}$$

where $b = \omega_{12}/\omega_{22}$, $a = \mu_1 - b\mu_2$, $\sigma^2 = \omega_{11} - \omega_{12}^2/\omega_{22}$, and the vertical bar $|$ is the conditioning operator. In model form, (4) is:

$$y_t = a + bz_t + \nu_{1t} \qquad \nu_{1t} \sim IN(0, \sigma^2) \tag{5a}$$

$$z_t = \mu_2 + \varepsilon_{2t} \qquad \varepsilon_{2t} \sim IN(0, \omega_{22}), \tag{5b}$$

where ν_{1t} $[= \varepsilon_{1t} - (\omega_{12}/\omega_{22})\varepsilon_{2t}]$ is the error in the conditional model for y_t given z_t, and the error $\boldsymbol{\varepsilon}_t$ is $(\varepsilon_{1t}, \varepsilon_{2t})'$. In the standard regression framework, (5a) would be obtained by conditioning y_t on z_t. Since (5a) is a conditional model, ν_{1t} is $y_t - \mathcal{E}(y_t \mid z_t)$. Thus, ν_{1t} contains only that part of y_t which is uncorrelated with z_t, and so which is uncorrelated with ε_{2t}, in light of (5b). It then follows that $\mathcal{E}(z_t \cdot \nu_{1t}) = 0$ and $\mathcal{E}(\varepsilon_{2t} \cdot \nu_{1t}) = 0$.

Symbolically, the relationship between (1) and (4) [or between (2) and (5)] is:

$$F_{\boldsymbol{x}}(\boldsymbol{x}_t;\, \boldsymbol{\theta}) = F_{y|z}(y_t \mid z_t;\, \boldsymbol{\lambda}_1) \cdot F_z(z_t;\, \boldsymbol{\lambda}_2), \tag{6}$$

where $F_v(\cdot)$ denotes the density function for variable v. Thus, $F_{\boldsymbol{x}}(\boldsymbol{x}_t;\, \boldsymbol{\theta})$ is the joint density of $\boldsymbol{x}_t$, $F_{y|z}(y_t \mid z_t;\, \boldsymbol{\lambda}_1)$ is the conditional density of y_t given z_t, and $F_z(z_t;\, \boldsymbol{\lambda}_2)$ is the marginal density of z_t. The parameter vector $\boldsymbol{\theta}$ is the full set of parameters in the joint process; $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are the parameters of the conditional and marginal models; and the respective parameter spaces are $\boldsymbol{\Theta}$, $\boldsymbol{\Lambda}_1$, and $\boldsymbol{\Lambda}_2$. Defining $\boldsymbol{\lambda}$ as $(\boldsymbol{\lambda}_1', \boldsymbol{\lambda}_2')'$ and denoting its parameter space as $\boldsymbol{\Lambda}$, then there is a one-to-one function $\boldsymbol{g}(\cdot)$ such that $\boldsymbol{\lambda} = \boldsymbol{g}(\boldsymbol{\theta})$.

[2] The word "marginal" is used here and elsewhere in its statistical sense. Its usage in statistics arises from summing a tabulated joint distribution function across its rows or down its columns and entering those sums in the margin, to obtain what is known as the marginal distribution; see Kendall and Stuart (1977, p. 22). "Marginal" as used here is not to be confused with its economic sense in (e.g.) "marginal versus average cost".

Above, $\boldsymbol{\theta} = [\boldsymbol{\mu}', \mathrm{vec}(\boldsymbol{\Omega})']'$, $\boldsymbol{\lambda}_1 = (a, b, \sigma^2)'$, and $\boldsymbol{\lambda}_2 = (\mu_2, \omega_{22})'$. The representation in (6) is important throughout this overview.

In (6), the joint density of $\boldsymbol{x}_t$ is factorized into the conditional density of y_t given z_t and the marginal density of z_t. This factorization is ***without*** loss of generality. Even so, analyzing the conditional density $F_{y|z}(y_t \mid z_t;\ \boldsymbol{\lambda}_1)$ while ignoring the corresponding marginal density $F_z(z_t;\ \boldsymbol{\lambda}_2)$ is ***with*** loss of generality, and in general implies a loss of information about the conditional process being modeled. Analyzing the conditional model alone is the statistical formalization of taking z_t as "given", so the remainder of this section considers the corresponding implications.

Modeling the conditional density (4a) by itself ignores some information about the conditional model's parameters (a, b, σ^2) when any of a, b, and σ^2 are linked to the marginal model's parameters (μ_2, ω_{22}), e.g., via cross-equation restrictions. However, in spite of the definitions of a, b, and σ^2, such dependence may be absent. For instance, if ω_{12} took values in proportion to ω_{22} as (e.g) ω_{22} varied across different regimes for z_t, then the value of ω_{22} (in $\boldsymbol{\lambda}_2$) would be uninformative about b, which is ω_{12}/ω_{22}, and is in $\boldsymbol{\lambda}_1$.

"Lack of dependence" between $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ is an overly strong condition for inference. Instead, a related concept, the ***sequential cut*** of a density function, is used. Specifically, the factorization (6) operates a sequential cut if and only if $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are ***variation free***, i.e., $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ belong to $\boldsymbol{\Lambda}_1 \times \boldsymbol{\Lambda}_2$, the product of their individual parameter spaces. Thus, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are variation-free if the parameter space $\boldsymbol{\Lambda}_1$ is not a function of the parameter $\boldsymbol{\lambda}_2$, and the parameter space $\boldsymbol{\Lambda}_2$ is not a function of the parameter $\boldsymbol{\lambda}_1$. Expressed slightly differently, knowledge about the value of one parameter provides no information on the other parameter's range of potential values. Thus, under weak exogeneity, permissible $\boldsymbol{\theta}$ are always reconstructed correctly from separately selected values of $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$. Example 2 below examines the concept of "variation free" in greater detail.

The parameters $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ being variation free is not enough to ensure valid inference about the parameters of interest, using the conditional model (4a) alone. For instance, if an investigator were interested in $\boldsymbol{\mu}$, both (4a) and (4b) would need to be estimated: $\boldsymbol{\mu}$ can not be retrieved from only (a, b, σ^2). Thus, the formal notion of ***parameters of interest*** (denoted $\boldsymbol{\psi}$) is introduced. This leads to the definition of weak exogeneity.

> *Definition.* The variable z_t is weakly exogenous over the sample period for the parameters of interest $\boldsymbol{\psi}$ if and only if there exists a reparameterization of $\boldsymbol{\theta}$ as $\boldsymbol{\lambda}$, with $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1', \boldsymbol{\lambda}_2')'$ such that
> (i) $\boldsymbol{\psi}$ is a function of $\boldsymbol{\lambda}_1$ alone, and
> (ii) the factorization in (6) operates a sequential cut, i.e.,
> $$F_{\boldsymbol{x}}(\boldsymbol{x}_t;\ \boldsymbol{\theta}) = F_{y|z}(y_t \mid z_t;\ \boldsymbol{\lambda}_1) \cdot F_z(z_t;\ \boldsymbol{\lambda}_2)$$
> where $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_1 \times \boldsymbol{\Lambda}_2$.
>
> [Engle, Hendry, and Richard (1983, p. 282)]

If weak exogeneity holds, then efficient estimation and testing may be conducted by analyzing only the conditional model (4a), ignoring the information of the marginal process (4b). In economic analysis, there may be many variables in z_t but relatively few in y_t, so weak exogeneity can greatly reduce the modeling effort required.

Example 2: the cobweb model. The concepts "parameters of interest", "variation free", and "weak exogeneity" are illustrated clearly with the standard economic cobweb model.

This model characterizes a market with lags in the production process, as might occur with agricultural commodities. See Tinbergen (1931) and Suits (1955) for pivotal contributions, and Henderson and Quandt (1971, pp. 142-145) for an exposition.

The cobweb model is obtained by a simple generalization of the static bivariate normal model in Example 1. Specifically, let the mean of z_t depend linearly upon y_{t-1}, the lagged value of y_t: Example 3 gives details. From (5a) and (5b), the resulting model is:

$$y_t = a + bz_t + \nu_{1t} \qquad \nu_{1t} \sim IN(0, \sigma^2) \tag{7a}$$

$$z_t = ky_{t-1} + \varepsilon_{2t} \qquad \varepsilon_{2t} \sim IN(0, \omega_{22}), \tag{7b}$$

where k is a parameter capturing the linear dependence of z_t on y_{t-1}. In the cobweb model, y_t and z_t are interpreted as the logs of price and quantity respectively. Denoting those logs as p_t and q_t, and ignoring the constant a (for ease of exposition), (7a)-(7b) becomes:

$$p_t = bq_t + \nu_{1t} \qquad \nu_{1t} \sim IN(0, \sigma^2) \tag{8a}$$

$$q_t = kp_{t-1} + \varepsilon_{2t} \qquad \varepsilon_{2t} \sim IN(0, \omega_{22}). \tag{8b}$$

As before, $\mathcal{E}(q_t \cdot \nu_{1t}) = 0$ and $\mathcal{E}(\varepsilon_{2t} \cdot \nu_{1t}) = 0$.

The cobweb model (8) has the following interpretation and properties. The first equation, (8a), is derived from a demand equation: the price (p_t) clears the market for a given quantity (q_t) supplied. The value $1/b$ is the price elasticity of demand. The second equation, (8b), is a supply equation, capturing (e.g.) how much farmers decide to produce this year (q_t), depending upon the price they were able to obtain in the previous year (p_{t-1}). The value k is the price elasticity of supply. The stability of (8a)-(8b) as a system is sometimes of interest, and can be determined from the reduced form for p_t [e.g., by substituting (8b) into (8a)]:

$$p_t = \rho p_{t-1} + \varepsilon_{1t} \qquad \varepsilon_{1t} \sim IN(0, \omega_{11}), \tag{9}$$

where ρ is the root of (9), and is equal to $b \cdot k$. If $|\rho| < 1$, the market is dynamically stable. If $|\rho| = 1$, the market generates prices which oscillate without dampening; and if $|\rho| > 1$, the market is dynamically unstable.

Now, consider how parameters of interest and parameter spaces determine whether or not the quantity q_t in (8a) is weakly exogenous. Specifically, consider conditions (i) and (ii) for weak exogeneity individually, recognizing that both conditions must be satisfied for weak exogeneity to hold. In the notation from Example 1, the parameters of the conditional model (8a) are $\boldsymbol{\lambda}_1 = (b, \sigma^2)'$, and the parameters of the marginal model (8b) are $\boldsymbol{\lambda}_2 = (k, \omega_{22})'$.

Condition (i) for weak exogenei`-- requires that the parameters of interest $\boldsymbol{\psi}$ be a function of the conditional model's parameters $\boldsymbol{\lambda}_1$ only. If the parameter of interest is the demand elasticity ($1/b$), this condition is satisfied: b enters $\boldsymbol{\lambda}_1$, and $\boldsymbol{\lambda}_1$ alone. However, if the stability of the system is at issue and so the parameter of interest is the root ρ, condition (i) is violated. The parameter ρ requires knowledge of both b (in $\boldsymbol{\lambda}_1$) and k (in $\boldsymbol{\lambda}_2$), and so necessitates analysis of the full system. Thus, q_t is not weakly exogenous for the root ρ; but q_t may be weakly exogenous for the elasticity $1/b$, depending upon whether or not condition (ii) is satisfied. Choosing the parameters of interest is not an innocuous decision.

Condition (ii) for weak exogeneity requires that the parameters of the conditional and marginal models ($\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$) are variation free. The following three situations show how these parameters might or might not be so. For ease of exposition, ignore the presence of σ^2 and ω_{22} in $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, thereby allowing analysis of the parameter space $\boldsymbol{\Lambda}$ [now the parameter space of (b, k)] on a plane.

First, suppose that b and k are completely unrestricted real values. Their parameter space $\boldsymbol{\Lambda}$ is $\Re^2$, the complete real plane. For every value of k, the parameter b can take any value in the interval $(-\infty, +\infty)$, which is $\boldsymbol{\Lambda}_1$. The value of the marginal model's parameter k does not affect the range of the conditional model's parameter b, and conversely, so b and k (i.e., $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$) are variation free. Equivalently, the parameter space $\boldsymbol{\Lambda}$ is the product space of $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$, i.e., $(-\infty, +\infty) \times (-\infty, +\infty)$, which is $\Re^2$. Thus, q_t is weakly exogenous for (e.g.) the elasticity $1/b$.

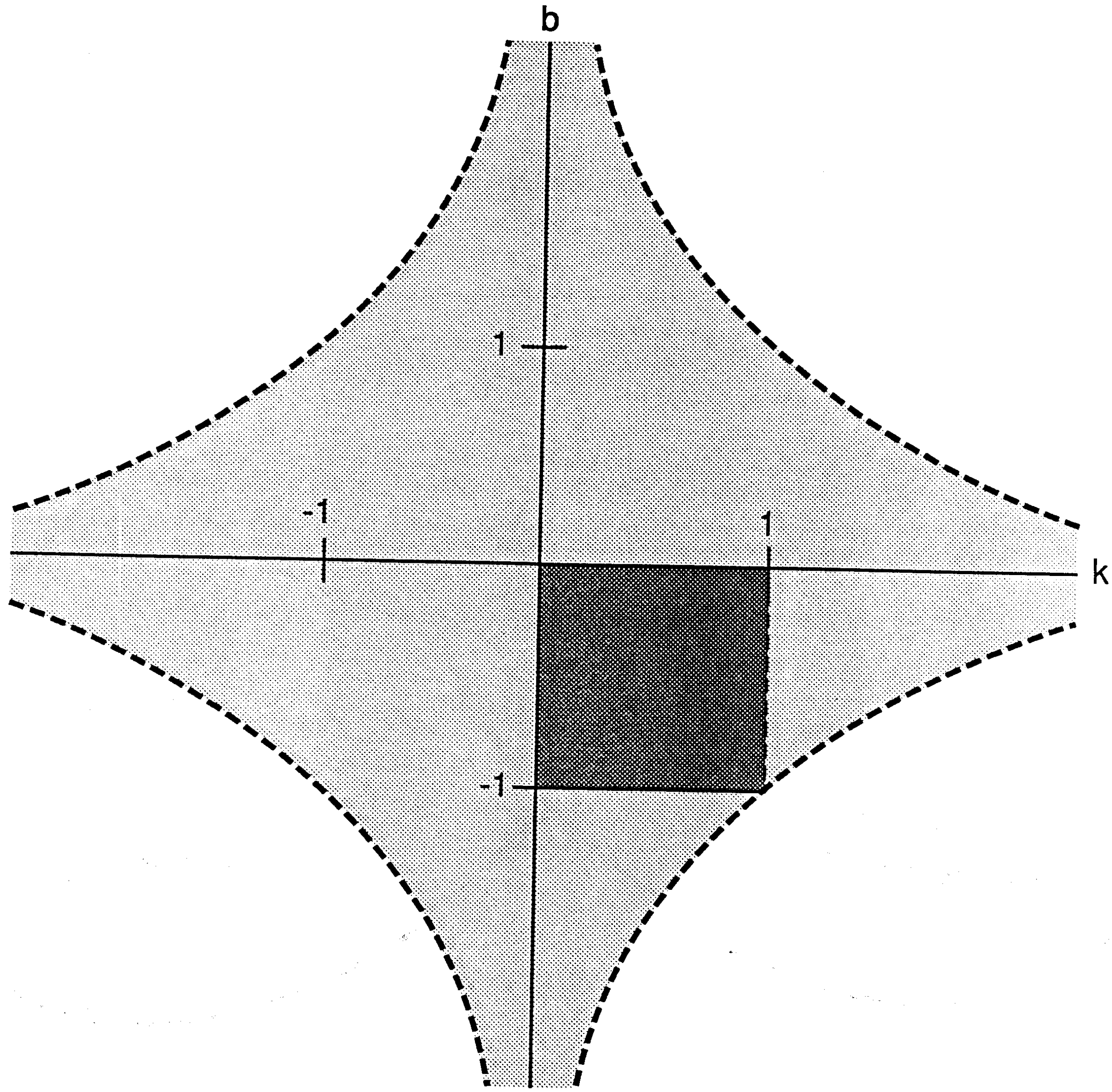
Second, suppose that b and k are restricted such that the system (8a)-(8b) is stable, i.e., such that $|b \cdot k| < 1$. Their corresponding parameter space $\boldsymbol{\Lambda}$ appears in Figure 1, labelled as "stability". Unlike the previous case, the value of k does affect the range of b. For instance, if $k = 0.5$, then b must lie in the interval $(-2, +2)$, whereas if $k = 0.2$, then b lies in the interval $(-5, +5)$. Formally, $\boldsymbol{\Lambda}_1$, which is the parameter space of b (or $\boldsymbol{\lambda}_1$), depends upon the values of the marginal process's parameter k (or $\boldsymbol{\lambda}_2$). Thus, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are not variation free.

Equally, the parameter space $\boldsymbol{\Lambda}$ is not $\boldsymbol{\Lambda}_1 \times \boldsymbol{\Lambda}_2$, the product of the spaces of b and k. For example, for $k = 0.2$, $\boldsymbol{\Lambda}_1$ is $(-5, +5)$; and for $b = 1$, $\boldsymbol{\Lambda}_2$ is $(-1, +1)$. However, the product space $\boldsymbol{\Lambda}_1 \times \boldsymbol{\Lambda}_2$, which is $(-5, +5) \times (-1, +1)$, is not the parameter space $\boldsymbol{\Lambda}$ for (b, k), which is defined by $|b \cdot k| < 1$. Put somewhat differently, the value of k is informative about the value of b, even though k does not determine the specific value of b. Inference using the conditional density alone loses information about b from k in the marginal density, so q_t is not weakly exogenous for the elasticity $1/b$.

Third, suppose that (e.g.) economic theory or intuition suggests the following restrictions: that the supply elasticity k lies in the unit interval $[0, 1)$, and that the demand elasticity $1/b$ is negative and is greater than or equal to unity in absolute value (implying that $-1 \leq b \leq 0$). The corresponding parameter space $\boldsymbol{\Lambda}$ appears in Figure 1, labelled as "elasticity restrictions". The parameter b lies in the interval $[-1, 0]$, regardless of the value of k; and k lies in the interval $[0, 1)$, regardless of the value of b. The parameters *are* variation free: the product space $[-1, 0] \times [0, 1)$, which is $\boldsymbol{\Lambda}_1 \times \boldsymbol{\Lambda}_2$, is also the space in which (b, k) lies, i.e., $\boldsymbol{\Lambda}$. Thus, under the "elasticity restrictions", q_t is weakly exogenous for the elasticity $1/b$. As with parameters of interest, the choice of parameter space is important in the determination of a variable's status as exogenous or endogenous.

To summarize, the parameter space and the parameters of interest are important concepts, both statistically and economically. Their choice is critical to the exogeneity status of a given variable. The introduction of dynamics in (7) above leads naturally to another concept of exogeneity, strong exogeneity.

Strong exogeneity. Strong exogeneity is the conjunction of weak exogeneity and Granger non-causality, and insures valid conditional forecasting. Figure 2 shows the relationship between weak exogeneity, Granger non-causality, and strong exogeneity via a Venn diagram. Figure 2 also includes a set for the property of invariance, which helps



Stability: {(b,k) such that |b • k| <1}

Elasticity restrictions: {(b,k) such that -1 ≤ b ≤ 0 and 0 ≤ k < 1}

Figure 1. Possible parameter spaces of the cobweb model.

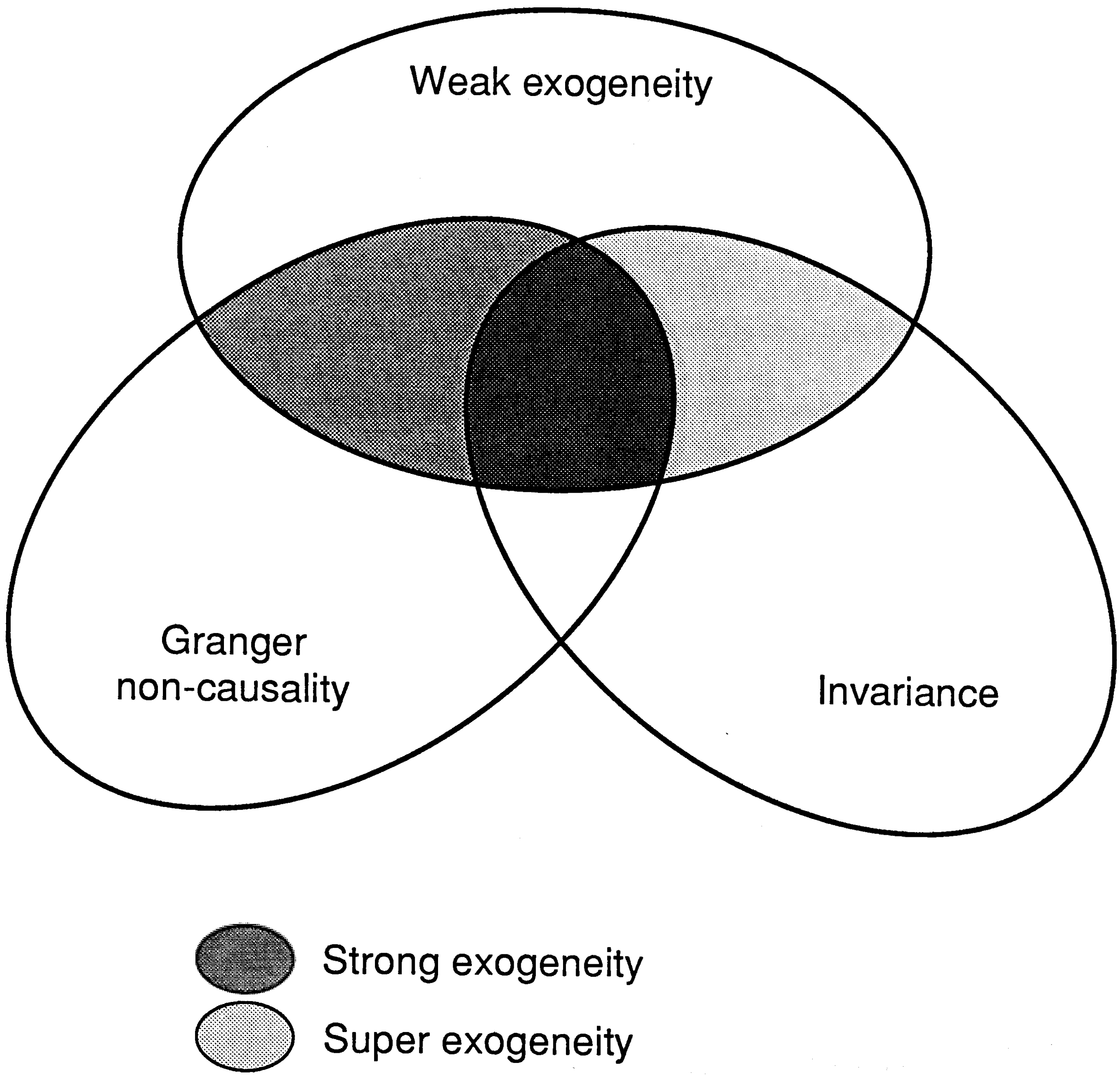


Figure 2. The relationship between Granger non-causality, invariance, and three forms of exogeneity.

define super exogeneity (discussed after strong exogeneity).

To discuss the concept of strong exogeneity, Example 1 is modified to include dynamics by reinterpreting (1) as the joint density of $\boldsymbol{x}_t$, conditional on the past of $\boldsymbol{x}_t$ (denoted $\boldsymbol{X}_{t-1}$). In general, $\boldsymbol{X}_{t-1}$ may affect the distribution of $\boldsymbol{x}_t$ through both $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$, and in a rather arbitrary fashion. For simplicity, only *linear* dependence of $\boldsymbol{\mu}$ on the *first* lag of $\boldsymbol{x}_t$ is considered. Thus, the mean $\boldsymbol{\mu}$ in (1) is interpreted as the conditional mean of $\boldsymbol{x}_t$ given $\boldsymbol{X}_{t-1}$. That is, $\boldsymbol{\mu} = \boldsymbol{\pi}_1 \boldsymbol{x}_{t-1}$, where $\boldsymbol{\pi}_1$ is a matrix of coefficients, and the constant term is ignored for simplicity.

Example 3: joint, conditional, and marginal densities with lags. Under this simplifying assumption about lags and linearity, (1) becomes a first-order vector autoregression (VAR) in model form:

$$\boldsymbol{x}_t = \boldsymbol{\pi}_1 \boldsymbol{x}_{t-1} + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t \sim IN(\boldsymbol{0}, \boldsymbol{\Omega}). \tag{10}$$

Equations (5a) and (5b) become conditional and marginal autoregressive distributed lag (AD) models:

$$y_t = b_0 z_t + b_1 z_{t-1} + b_2 y_{t-1} + \nu_{1t} \qquad \nu_{1t} \sim IN(0, \sigma^2) \tag{11a}$$

$$z_t = \pi_{22} z_{t-1} + \pi_{21} y_{t-1} + \varepsilon_{2t} \qquad \varepsilon_{2t} \sim IN(0, \omega_{22}), \tag{11b}$$

where π_{ij} denotes the *(i,j)*th element of $\boldsymbol{\pi}_1$, and (b_0, b_1, b_2) are derived from $\boldsymbol{\pi}_1$ and $\boldsymbol{\Omega}$, paralleling the relation between b and $(\boldsymbol{\mu}, \boldsymbol{\Omega})$ in (4a). Specifically, $b_0 = \omega_{12}/\omega_{22}$, $b_1 = \pi_{12} - (\omega_{12}/\omega_{22})\pi_{22}$, and $b_2 = \pi_{11} - (\omega_{12}/\omega_{22})\pi_{21}$; see Engle, Hendry, and Richard (1983, p. 297).

Valid prediction of y from its conditional model (11a) requires more than weak exogeneity. With weak exogeneity alone, y_{t-1} influences z_t if $\pi_{21} \neq 0$ in the marginal model (11b), in which case z_t in the conditional model (11a) can not be treated as "fixed" for prediction of y_t. The requisite additional restriction is that $\pi_{21} = 0$, or (in general) that y does not Granger-cause z.[3] Weak exogeneity plus Granger non-causality generates strong exogeneity. Strong exogeneity permits valid multi-step ahead prediction of y from (11a), conditional on predictions of z generated from (11b) (with $\pi_{21} = 0$), where the predictions of z depend upon only their own lags. With $\pi_{21} \neq 0$, valid prediction of y must account for the feedback of y onto z in (11b), period by period: valid multi-step prediction therefore *requires* joint analysis of (11a) and (11b), violating the exogeneity of z_t in (11a).

"Error-correction" models (and so cointegrated processes) are closely related to the autoregressive distributed lag model in (11a). By appropriately adding and subtracting y_{t-1} and z_{t-1} from (11a), that equation may be written as an error-correction model (ECM):

$$\Delta y_t = \gamma_1 \Delta z_t + \gamma_2 (y_{t-1} - \delta z_{t-1}) + \nu_{1t}, \tag{12}$$

where $\gamma_1 = b_0$, $\gamma_2 = b_2 - 1$, and $\delta = -(b_0 + b_1)/(b_2 - 1)$.[4] Equation (12) involves no loss

---

[3] Specifically, Granger non-causality means that only lagged values of z_t enter (11b) or its generalization. Equally, Granger non-causality means that lags of other variables (other than of z_t) do not enter the marginal model for z_t.

[4] With the lag operator L defined as $Lx_t = x_{t-1}$, the difference operator Δ is $(1 - L)$; hence $\Delta x_t = x_t - x_{t-1}$. More generally, $\Delta_j^i x_t = (1 - L^j)^i x_t$. If i (or j) is undefined, it is taken to be unity.

of generality relative to (11a), provided $b_2 \neq 1$. The term $\gamma_1 \Delta z_t$ is the immediate impact that a change in z_t has on y_t. The term $\gamma_2(y_{t-1} - \delta z_{t-1})$ (with $\gamma_2 < 0$ required for dynamic stability) is the impact on Δy_t of having y_{t-1} out of line with δz_{t-1}: $y = \delta z$ is the long-run static solution to (12). Discrepancies between y_{t-1} and δz_{t-1} could arise from errors in agents' past decisions, with the presence of $\gamma_2(y_{t-1} - \delta z_{t-1})$ reflecting their attempts to correct such errors: hence the name ***error-correction*** model. Several papers in this special issue develop conditional ECMs in the course of modeling.

Hendry, Pagan, and Sargan (1984, pp. 1040-1049) discuss the properties of ECMs in greater detail, and show that many classes of models common in empirical research are subsumed by the ECM. The ECM representation is also central to the discussion of cointegration, as shown in Example 7 below.

Super exogeneity. Super exogeneity is the conjunction of weak exogeneity and "invariance" (see Figure 2), and insures valid policy simulations. The concept of invariance can be motivated as follows.

Frequently, the reduced form (2) is empirically nonconstant, due to (e.g.) OPEC shocks, changes in policy rules, and financial innovations. The factorization (6) may aim to isolate those nonconstancies into the sub-vector $\boldsymbol{\lambda}_2$, leaving the parameters of the conditional model $\boldsymbol{\lambda}_1$ invariant to the changes which have occurred. Thus, the concept of *invariance* is introduced. The parameter $\boldsymbol{\lambda}_1$ is invariant to a class of interventions to the marginal process for z_t (i.e., to a set of changes in $\boldsymbol{\lambda}_2$) if $\boldsymbol{\lambda}_1$ is not a function of $\boldsymbol{\lambda}_2$ for that class of interventions.[5] For invariance, lack of dependence between the parameters themselves matters, and not just lack of dependence between parameters and parameter spaces.

Policy analysis (or counter-factual analysis) often involves changing the marginal process of z_t. Valid analysis of the conditional model under such changes requires that the parameters $\boldsymbol{\lambda}_1$ be *invariant* to those changes (or "interventions"). The relevant concept is super exogeneity, whereby z_t is weakly exogenous for the parameters of interest $\boldsymbol{\psi}$, and $\boldsymbol{\lambda}_1$ is invariant to the class of interventions to $\boldsymbol{\lambda}_2$ under consideration (see Figure 2). Importantly, a variable is super exogenous with respect to a *specified* class of interventions (e.g., those that occurred within sample): the variable need not be super exogenous with respect to interventions outside that class, although it may be.

To illustrate the concept of super exogeneity, interpret y_t and z_t as money and an interest rate, (11a) as a money demand function, and (11b) as the Fed's interest rate reaction function (assuming that there is one). Then, policy analysis of (11a) for money demanded at different levels of the interest rate is valid only if the parameters in the money-demand function ($\boldsymbol{\lambda}_1$) are invariant to the specified changes in the parameters of the interest-rate reaction function ($\boldsymbol{\lambda}_2$) which would generate those levels. That condition is equivalently that the interest rate is super exogenous for the parameters in the money demand function with respect to the changes in the reaction function under consideration.

Predeterminedness and strict exogeneity. Before turning to tests of exogeneity (Section 2.B), the concepts of predeterminedness and strict exogeneity are considered briefly, as they often appear in discussions of exogeneity. Neither concept specifies parameters of interest;

[5] Engle, Hendry, and Richard (1983, p. 284) call this structural invariance, rather than invariance.

and that is a major drawback of each concept, given the preceding exposition. For instance, in the cobweb model, q_t is predetermined in the conditional model (8a), and OLS of (8a) is consistent for b. However, that conditional model is not sufficient to obtain the parameter of interest if the parameter of interest is the root ρ. Likewise, z_t is strictly exogenous in (5a), but that conditional equation by itself is not sufficient to obtain μ_1. Neither predeterminedness nor strict exogeneity is sufficient for efficient statistical inference; lack of necessity can be shown as well. See Engle, Hendry, and Richard (1983) for details and examples, including of the dynamic simultaneous equations model.

B. Inference

Whether or not z_t is exogenous depends *inter alia* upon the process generating y_t and z_t, so exogeneity may be testable.

Even so, testing for weak exogeneity *per se* is often difficult because doing so involves modeling z_t, whereas a motivation for assuming weak exogeneity is to avoid modeling z_t. Also, weak exogeneity depends upon the parameters of interest, which are chosen by the investigator. Still, a few tests of weak exogeneity have been proposed, notably in the context of cointegration; see the discussion of Johansen (1990) in Section 3.B below.

While strong exogeneity requires weak exogeneity, the former is refuted by finding Granger causality of y onto z. Thus, Granger non-causality, which is a necessary condition for strong exogeneity, generates an easily calculated (albeit incomplete) test of strong exogeneity.

Super exogeneity requires weak exogeneity of z_t for the parameters of interest $\boldsymbol{\psi}$, and invariance of the conditional model's parameters $\boldsymbol{\lambda}_1$ (on which $\boldsymbol{\psi}$ depends) to changes in the parameters of the marginal process ($\boldsymbol{\lambda}_2$). Thus, two common tests for super exogeneity are as follows.

(i) Establish the constancy of $\boldsymbol{\lambda}_1$ and the nonconstancy of $\boldsymbol{\lambda}_2$. With $\boldsymbol{\lambda}_1$ constant and $\boldsymbol{\lambda}_2$ not, then $\boldsymbol{\lambda}_1$ must be invariant to $\boldsymbol{\lambda}_2$, and so super exogeneity holds; cf. Hendry (1988).

(ii) Having established (i), further develop the marginal model for z_t until it is empirically constant. For instance, by adding dummies and/or other variables, model the way in which $\boldsymbol{\lambda}_2$ varies over time. Then test for the significance of those dummies and/or other variables when they are added to the conditional model. Their *in*significance in the conditional model demonstrates invariance of the conditional model's parameters $\boldsymbol{\lambda}_1$ to the changes in the marginal process; cf. Engle and Hendry (1989).

Tests of parameter constancy thus are central to tests of super exogeneity.

Parameter constancy is of more general interest as well. Economic theory focuses on the invariants of the economic process, as reflected by the continuing debates on autonomy, "deep" or "structural" parameters, and the Lucas critique. Given the intimate link between parameter constancy and predictive accuracy, valid forecasting also relies on constant parameters; cf. Hendry (1979). Most estimation techniques require parameter constancy for valid inference, so parameter constancy is a central concept from a statistical perspective. Even models with time-varying parameters posit "meta-parameters", which are assumed constant over time and whose empirical constancy could be tested. Recursive estimation is an incisive tool for investigating parameter constancy, both through the sequence of estimated coefficient values and via the associated Chow (1960) statistics for constancy.

Thus, the constancy of a model bears on its economic interpretability, forecasting, and policy analysis, the latter specifically via tests of super exogeneity. Because of the critical role of parameter constancy in policy analysis, many of the papers in this special issue employ recursive estimation and tests of parameter constancy and predictive accuracy. Further, several papers in this issue analyze and develop tests of parameter constancy and predictive accuracy themselves. See Section 4 below.

C. Policy Implications

Super exogeneity has several implications for policy analysis.

First, the empirical presence of super exogeneity immunizes the conditional model from the Lucas (1976) critique; see Hendry (1988), Engle and Hendry (1989), Ericsson and Hendry (1989), and Favero and Hendry (1989). For example, suppose that the conditional and marginal models represent agents' and policy makers' decision rules respectively. Then, under super exogeneity, the agents' parameter vector $\boldsymbol{\lambda}_1$ is invariant to changes in policy makers' rules (via $\boldsymbol{\lambda}_2$), which is opposite to the implication of the Lucas critique.

Second, "inverting" the conditional model is invalid. For example, inverting a money-demand equation to obtain a price equation is invalid. The bivariate normal distribution in (1) demonstrates how and why that is so.

In (4a) and (4b), the joint density (1) was factorized into the conditional density of y_t given z_t and the marginal density of z_t. Equation (1) also may be factorized into the conditional density of z_t given y_t and the marginal density of y_t:

(13*a*) $$z_t \mid y_t \sim IN(c + dy_t, \tau^2)$$

(13*b*) $$y_t \sim IN(\mu_1, \omega_{11}),$$

where $d = \omega_{21}/\omega_{11}$, $c = \mu_2 - d\mu_1$, and $\tau^2 = \omega_{22} - \omega_{21}^2/\omega_{11}$. The factorization in (13) is "opposite" to that in (4). In model form, (13) is:

(14*a*) $$z_t = c + dy_t + \nu_{2t} \qquad \nu_{2t} \sim IN(0, \tau^2)$$

(14*b*) $$y_t = \mu_1 + \varepsilon_{1t} \qquad \varepsilon_{1t} \sim IN(0, \omega_{11}),$$

where ν_{2t} is the error in the conditional model for z_t given y_t. Paralleling results for (5), $\mathcal{E}(\nu_{2t} \cdot y_t) = 0$ and $\mathcal{E}(\nu_{2t} \cdot \varepsilon_{1t}) = 0$. Symbolically, (13) is:

(15) $$F_{\boldsymbol{x}}(\boldsymbol{x}_t;\, \boldsymbol{\theta}) = F_{z|y}(z_t \mid y_t;\, \boldsymbol{\phi}_1) \cdot F_y(y_t;\, \boldsymbol{\phi}_2),$$

where the parameterization is $\boldsymbol{\phi} \equiv (\boldsymbol{\phi}_1', \boldsymbol{\phi}_2')' = \boldsymbol{h}(\boldsymbol{\theta})$, and $\boldsymbol{h}(\cdot)$ is a one-to-one function. Thus, via $\boldsymbol{g}(\cdot)$ and $\boldsymbol{h}(\cdot)$, there is a one-to-one mapping between $(\boldsymbol{\lambda}_1', \boldsymbol{\lambda}_2')$ and $(\boldsymbol{\phi}_1', \boldsymbol{\phi}_2')$. Specifically, the coefficient on y_t in (14a) is $d = b\omega_{22}/(\sigma^2 + b^2\omega_{22})$, which is not $1/b$ unless (5a) is non-stochastic (i.e., $\sigma^2 = 0$). Further, if z_t is super exogenous for b and σ^2, then d will vary as the marginal process for z_t varies (via ω_{22}) even though b remains constant. Inversion does not obtain the correct parameter for the inverted equation, and the parameter in the inverted model may be nonconstant even if the "uninverted" conditional model is constant.

Inversion may appear peculiar at first glance, but it is precisely what occurs when (e.g.) estimated money-demand functions are "inverted" to obtain prices as a function of money (common among macro-economists) or to obtain interest rates as a function of money (common among macro-modelers). For example, suppose a conditional money-demand function is estimated:

$$m = \kappa_1 p + \kappa_2 i - \kappa_3 R, \tag{16}$$

where m, p, and i are the logs of nominal money, the price level, and real income respectively, R is the interest rate, and the $\{\kappa_i\}$ are coefficients, all assumed positive. The two inversions above correspond to

$$p = (1/\kappa_1)m - (\kappa_2/\kappa_1)i + (\kappa_3/\kappa_1)R \tag{17a}$$

and

$$R = (-1/\kappa_3)m + (\kappa_1/\kappa_3)p + (\kappa_2/\kappa_3)i, \tag{17b}$$

where $b = \kappa_1$ in the first instance and $b = -\kappa_3$ in the second. From the analytical relationship between d and b above, the *empirically estimated* coefficient on m in (17a) or (17b) need not be at all close to $1/b$, even for large samples. The nonconstancy of the inverted model can be demonstrated empirically as well, as in Hendry (1985) and Hendry and Ericsson (1991a, 1991b).

Third, super exogeneity can identify parameters, in the sense of uniqueness, because any (nontrivial) combination of the conditional and marginal equations would be nonconstant; cf. Hendry (1987, p. 40). This contrasts with (e.g.) Cooley and Leroy (1981).

Fourth, Granger non-causality is neither necessary nor sufficient for policy analysis, and contrasts with a common approach to exogeneity. Lagged y in (11b) may influence current z (e.g., the Fed might pay attention to lagged money in setting the interest rate), yet the conditional model (11a) (e.g., a money-demand equation) would still be valid for policy analysis if z_t were super exogenous for $\boldsymbol{\lambda}_1$. Granger non-causality is relevant for strong exogeneity, but that concept is for forecasting, not policy analysis.

We now turn to the concept of cointegration.

3. Cointegration

Cointegration formalizes in statistical terms the property of a long-run relation between "integrated" economic variables. In this section, integration and cointegration are illustrated by (10), first as a first-order scalar autoregression (Example 4), then as a higher-order vector autoregression (Example 5). Specifically, the first-order bivariate (vector) autoregression is used to illustrate cointegration (Example 6), the relationship between cointegration and error-correction models (Example 7), and implications for weak exogeneity (Example 8). Then, a first-order trivariate vector autoregression shows how two cointegrating vectors might arise (Example 9).

Section 3.A describes integration and cointegration via the examples. Section 3.B summarizes several techniques for testing the order of integration and the existence of cointegration. Section 3.C draws upon Section 2 (on exogeneity) to discuss the implications of cointegration for policy analysis.

For the initial development of cointegration, see Granger (1981), Granger and Weiss (1983), the papers in Hendry (1986), and Engle and Granger (1987). For recent summaries and extensions, see Johansen (1988), Hylleberg and Mizon (1989), Engle and Yoo (1989), Dolado, Jenkinson, and Sosvilla-Rivero (1990), Johansen and Juselius (1990a), Campbell and Perron (1991), Phillips (1991), Phillips and Loretan (1991), Ericsson, Campos, and Tran (1991), Banerjee, Dolado, Galbraith, and Hendry (1991), and Johansen (1991a).

For the initial (and much earlier) development of error correction, see Phillips (1954, 1957), who *inter alia* discusses how error correction in policy makers' rules might help stabilize the economy. Important subsequent empirical and analytical contributions include Sargan (1964), Davidson, Hendry, Srba, and Yeo (1978), Salmon (1982), and Hendry, Pagan, and Sargan (1984).

A. Concepts and Structure

The essential concepts are integration and cointegration, which apply to individual time series and sets of time series respectively. A variable is integrated if it requires differencing to make it stationary. Many economic time series appear to be integrated; see Nelson and Plosser (1982). A set of integrated time series is cointegrated if some linear combination of those (non-stationary) series is stationary.

Example 4: integration. For a scalar (rather than bivariate) $\boldsymbol{x}_t$, (10) is a first-order autoregression:

$$x_t = \pi_1 x_{t-1} + \varepsilon_t, \tag{18}$$

which may be rewritten as:

$$\Delta x_t = \pi x_{t-1} + \varepsilon_t, \tag{19}$$

where $\pi = \pi_1 - 1$ by subtracting x_{t-1} from both sides of equation (18). If $\pi_1 = 1$ or equivalently $\pi = 0$, then x_t has a unit root and is said to be integrated of order one [denoted I(1)], meaning that x_t must be differenced once to achieve stationarity. In the simple case of (18), x_t is a random walk if it has a unit root. If $|\pi_1| < 1$, then x_t is stationary. For general autoregressive processes, (18) includes additional lags of x_t, and (so) (19) includes lags of Δx_t.

Example 5: cointegration. Equation (18) may be generalized to represent a vector of variables [as in (10)] and to include higher-order lags of x_t. Together, these result in:

$$\boldsymbol{x}_t = \sum_{i=1}^{\ell} \boldsymbol{\pi}_i \boldsymbol{x}_{t-i} + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t \sim IN(\mathbf{0}, \boldsymbol{\Omega}), \tag{20}$$

where ℓ is the maximum lag, and (20) may include a constant and dummies as well. In terms of the joint density in (1), the mean of $\boldsymbol{x}_t$ conditional on $\boldsymbol{X}_{t-1}$ is $\boldsymbol{\mu} = \sum_{i=1}^{\ell} \boldsymbol{\pi}_i \boldsymbol{x}_{t-i}$.

Following Johansen (1988) and Johansen and Juselius (1990a), (20) provides the basis for cointegration analysis. By adding and subtracting various lags of $\boldsymbol{x}_t$, (20) may be rewritten as:

$$\Delta \boldsymbol{x}_t = \boldsymbol{\pi} \boldsymbol{x}_{t-1} + \sum_{i=1}^{\ell-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{x}_{t-i} + \boldsymbol{\varepsilon}_t \tag{21}$$

where the $\{\boldsymbol{\Gamma}_i\}$ are

$$\boldsymbol{\Gamma}_i = -(\boldsymbol{\pi}_{i+1} + \ldots + \boldsymbol{\pi}_\ell) \qquad i = 1, \ldots, \ell - 1, \tag{22}$$

and

$$\boldsymbol{\pi} \equiv (\sum_{i=1}^{\ell} \boldsymbol{\pi}_i) - \boldsymbol{I}. \tag{23}$$

Equation (20) [and so (21)] simplifies to (10) for $\ell = 1$.[6] As in (19), $\boldsymbol{\pi}$ in (21) could be zero. If so, $\Delta \boldsymbol{x}_t$ in (21) depends upon $\boldsymbol{\varepsilon}_t$ and lags of $\Delta \boldsymbol{x}_t$ alone, all of which are I(0); so $\boldsymbol{x}_t$ is I(1).[7] If $\boldsymbol{\pi}$ is nonzero and of full rank with all the roots of an associated polynomial being within the unit circle, then all the $\boldsymbol{x}_t$ are I(0), paralleling $|\pi| < 1$ in the univariate case. However, because $\boldsymbol{\pi}$ is a matrix in (21) rather than a scalar, $\boldsymbol{\pi}$ may be of less than full rank, but of rank greater than zero. If so, each of the variables in $\boldsymbol{x}_t$ could be I(1), but with some linear combinations of those variables being I(0). The variables in $\boldsymbol{x}_t$ are then said to be cointegrated.

To show how cointegration can occur, denote the dimension of $\boldsymbol{x}_t$ as $p \times 1$ and the polynomial $(\sum_{i=1}^{\ell} \boldsymbol{\pi}_i \mathrm{z}^i) - \boldsymbol{I}_p$ as $\boldsymbol{\pi}(\mathrm{z})$, where z is the argument of the polynomial. Note that $\boldsymbol{\pi} = \boldsymbol{\pi}(1)$, from (23). The three possibilities for rank($\boldsymbol{\pi}$) are as follows.

(i) rank($\boldsymbol{\pi}$) $= p$. For $\boldsymbol{\pi}$ to have full rank, none of the roots of $|\boldsymbol{\pi}(\mathrm{z}^{-1})| = 0$ can be unity. Provided $|\boldsymbol{\pi}(\mathrm{z}^{-1})| = 0$ has all its $\ell \cdot p$ roots strictly inside the unit circle, $\boldsymbol{x}_t$ is stationary because $\boldsymbol{\pi}(L)$ can be inverted to give an infinite moving average representation of $\boldsymbol{x}_t$.

(ii) rank($\boldsymbol{\pi}$) $= 0$. Because $\boldsymbol{\pi} = 0$, equation (21) is in differences only, and there are p unit roots in $|\boldsymbol{\pi}(\mathrm{z}^{-1})| = 0$.

(iii) $0 <$ rank($\boldsymbol{\pi}$) $\equiv r < p$. In this case, $\boldsymbol{\pi}$ can be expressed as the outer product of two (full column rank) $p \times r$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$:

$$\boldsymbol{\pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'; \tag{24}$$

and there are $p - r$ unit roots in $|\boldsymbol{\pi}(\mathrm{z}^{-1})| = 0$.[8]

In (24), $\boldsymbol{\beta}'$ is the matrix of cointegrating vectors, and $\boldsymbol{\alpha}$ is the matrix of "weighting elements". Substituting (24) into (21) gives:

$$\Delta \boldsymbol{x}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \boldsymbol{x}_{t-1} + \sum_{i=1}^{\ell-1} \boldsymbol{\Gamma}_i \Delta \boldsymbol{x}_{t-i} + \boldsymbol{\varepsilon}_t. \tag{25}$$

Each $1 \times p$ row $\boldsymbol{\beta}'_i$ in $\boldsymbol{\beta}'$ is an individual cointegrating vector, as is required for "balance" to make each cointegrating relation $\boldsymbol{\beta}'_i \boldsymbol{x}_{t-1}$ an I(0) process in (25). Each $1 \times r$ row $\boldsymbol{\alpha}_j$ of $\boldsymbol{\alpha}$ is the set of weights for the r cointegrating terms appearing in the jth equation. Thus, the rank r is also the number of cointegrating vectors in the system. While $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ themselves are not unique, $\boldsymbol{\beta}$ uniquely defines the cointegration space, and suitable normalizations for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are available.

---

[6] Johansen (1988), Johansen and Juselius (1990a), and some others write (21) with the level of $\boldsymbol{x}$ entering at the ℓth lag rather than at the first lag. Doing so does not alter the coefficient on the lagged level (which is $\boldsymbol{\pi}$) although it does change the coefficients on the lagged values of $\Delta \boldsymbol{x}_t$. Since the analysis of cointegration concerns the properties of $\boldsymbol{\pi}$ alone, the choice of lag on $\boldsymbol{x}$ is irrelevant in this context.

[7] Here, we do not consider situations in which $\boldsymbol{x}_t$ is of an order of integration greater than unity.

[8] For example, for the first-order univariate model (18), $\boldsymbol{\pi}(\mathrm{z})$ is $\pi_1 \mathrm{z} - 1$, so the root of $|\boldsymbol{\pi}(\mathrm{z}^{-1})| = 0$ is π_1 itself. Possibilities (i) and (ii) correspond to $|\pi_1| < 1$ and $\pi_1 = 1$ respectively; and (iii) is not possible because $p = 1$.

In essence, $\boldsymbol{\alpha\beta}'\boldsymbol{x}_{t-1}$ in (25) contains all the long-run (levels) information on the process for $\boldsymbol{x}_t$: the only other observables in (25) are current and lagged $\Delta\boldsymbol{x}_t$. The vector $\boldsymbol{\beta}'\boldsymbol{x}_{t-1}$ measures the extent to which actual data deviate from the long-run relationship(s) among the variables in $\boldsymbol{x}_{t-1}$.

Engle and Granger (1987) establish an isomorphism between cointegration and error correction models: models with valid ECMs entail cointegration and, conversely, cointegrated series imply an error-correction representation for the econometric model. [For an exposition and extension, see Granger (1986).] To illustrate these and related issues, consider (20) as a first-order bivariate VAR.

Example 6: a single cointegrating vector. As a first-order bivariate VAR, (20) is:

$$\Delta y_t = \pi_{(11)}y_{t-1} + \pi_{(12)}z_{t-1} + \varepsilon_{1t} \tag{26a}$$

$$\Delta z_t = \pi_{(21)}y_{t-1} + \pi_{(22)}z_{t-1} + \varepsilon_{2t}, \tag{26b}$$

when expressed as (21) with $\boldsymbol{\pi} = \{\pi_{(ij)}\}$, and noting that $\boldsymbol{x}_t = (y_t, z_t)'$.[9] If there is one cointegrating vector ($r = 1$), then $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$ are 2×1 and 1×2 vectors, which may be denoted $(\alpha_1, \alpha_2)'$ and (β_1, β_2) respectively. Without loss of generality, $\boldsymbol{\beta}$ may be normalized with $\beta_1 = 1$. For convenience, denote the normalized $\boldsymbol{\beta}'$ vector as $(1, -\delta)$. Thus, (26) may be rewritten as:

$$\Delta y_t = \alpha_1(y_{t-1} - \delta z_{t-1}) + \varepsilon_{1t} \tag{27a}$$

$$\Delta z_t = \alpha_2(y_{t-1} - \delta z_{t-1}) + \varepsilon_{2t}, \tag{27b}$$

where $\alpha_1 = \pi_{(11)}$, $\alpha_2 = \pi_{(21)}$, and $\delta = -\pi_{(12)}/\pi_{(11)} = -\pi_{(22)}/\pi_{(21)}$. The cointegrating relation $\boldsymbol{\beta}'\boldsymbol{x}_{t-1}$ is $(y_{t-1} - \delta z_{t-1})$. Equations (27a) and (27b) express the growth rate of each variable in terms of a past disequilibrium and a random error. If lags of $\Delta\boldsymbol{x}_t$ appear in (21) (i.e., $\ell > 1$), then lagged values of Δy_t and Δz_t will be present in (27a) and (27b).

Example 7: cointegration and error-correction models. Together, (27a) and (27b) correspond to the joint distribution of $\boldsymbol{x}_t$ conditional on its past ($\boldsymbol{X}_{t-1}$), as described in Section 2. Equations (27a)-(27b) may be factorized into the conditional distribution of y_t given z_t and lags of both variables, and the marginal distribution of z_t (also given lags of both variables):

$$\Delta y_t = \gamma_1\Delta z_t + \gamma_2(y_{t-1} - \delta z_{t-1}) + \nu_{1t} \tag{28a}$$

$$\Delta z_t = \alpha_2(y_{t-1} - \delta z_{t-1}) + \varepsilon_{2t}, \tag{28b}$$

where $\gamma_1 = \omega_{12}/\omega_{22}$ and $\gamma_2 = \alpha_1 - (\omega_{12}/\omega_{22})\alpha_2$. Equation (28a) is also the ECM in (12), and the marginal equation (28b) is the same as (27b). Error-correction models imply and are implied by cointegration.

Example 8: cointegration and weak exogeneity. The parameters in the conditional and marginal models (28a) and (28b) are $(\gamma_1, \gamma_2, \delta, \sigma^2)'$ and $(\alpha_2, \delta, \omega_{22})'$ respectively, and have been denoted $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ previously. For cointegrated variables, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are (in general) linked via δ and α_2. The parameter δ enters both $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ directly; α_2 enters $\boldsymbol{\lambda}_2$ directly and $\boldsymbol{\lambda}_1$ via γ_2, which is $\alpha_1 - (\omega_{12}/\omega_{22})\alpha_2$. Thus, z_t is not (in general) weakly exogenous for the cointegrating vector $\boldsymbol{\beta}'$, or (hence) for δ.

---

[9] The notation $\pi_{(ij)}$ distinguishes this *(i,j)*th element of $\boldsymbol{\pi}$ from the *(i,j)*th element of $\boldsymbol{\pi}_1$, which is denoted π_{ij} [as in (11b)].

Weak exogeneity of z_t for $\boldsymbol{\beta}'$ is obtained when $\alpha_2 = 0$, in which case (28) becomes:

(29*a*) $$\Delta y_t = \gamma_1 \Delta z_t + \gamma_2 (y_{t-1} - \delta z_{t-1}) + \nu_{1t}$$

(29*b*) $$\Delta z_t = \varepsilon_{2t},$$

where $\gamma_2 = \alpha_1$, $\boldsymbol{\lambda}_1 = (\gamma_1, \gamma_2, \delta, \sigma^2)'$, and $\boldsymbol{\lambda}_2 = (\omega_{22})$. Then, (29a) alone is sufficient for fully efficient inference about $\boldsymbol{\beta}'$, i.e., about δ. Johansen's (1990) test for weak exogeneity is a test of $\alpha_2 = 0$.

Example 9: two cointegrating vectors. Multiple cointegrating vectors for integrated processes exist between only three or more series. For expositional convenience, consider just three time series $\boldsymbol{x}_t = (y_t, z_t, w_t)'$ with two cointegrating vectors:

(30) $$\boldsymbol{\beta}' = \begin{bmatrix} 1 & -\delta_1 & 0 \\ 0 & 1 & -\delta_2 \end{bmatrix}.$$

The unit coefficients are normalizations and are without loss of generality. The zero restrictions are *with* loss of generality, and are included for ease of exposition only. From (21), the VAR representation of $\boldsymbol{x}_t$ is:

(31*a*) $$\Delta y_t = \alpha_{11}(y_{t-1} - \delta_1 z_{t-1}) + \alpha_{12}(z_{t-1} - \delta_2 w_{t-1}) + \varepsilon_{1t}$$

(31*b*) $$\Delta z_t = \alpha_{21}(y_{t-1} - \delta_1 z_{t-1}) + \alpha_{22}(z_{t-1} - \delta_2 w_{t-1}) + \varepsilon_{2t}$$

(31*c*) $$\Delta w_t = \alpha_{31}(y_{t-1} - \delta_1 z_{t-1}) + \alpha_{32}(z_{t-1} - \delta_2 w_{t-1}) + \varepsilon_{3t},$$

where $\boldsymbol{\alpha} = \{\alpha_{ij}\}$. In (31), much more complicated interactions may exist between disequilibria [i.e., $(y_{t-1} - \delta_1 z_{t-1})$ and $(z_{t-1} - \delta_2 w_{t-1})$] and the variables themselves than in (27). With $\ell > 1$, lags of Δy_t, Δz_t, and Δw_t enter (31a)-(31c), further enriching the dynamics.

Empirical analysis of data with multiple cointegrating vectors is harder than with a single vector, but some such studies already exist. Hendry and Mizon (1989) detect two cointegrating vectors between money, prices, income, and interest rates for the United Kingdom, with one vector being a money demand function and the other being the relation between observed and potential income, where the latter is proxied by a trend. Johansen and Juselius (1990b) find both purchasing power parity and uncovered interest rate parity in data on the UK effective exchange rate, the UK wholesale price index, a trade-weighted foreign price index, and UK and Eurodollar interest rates.

B. Inference

The presence of unit roots complicates inference because some associated limiting distributions are non-standard. Dickey and Fuller (1979) have tabulated the critical values for the least squares estimator of π and its t ratio for the univariate process (19). The presence of lagged Δx_t does not affect their limiting distributions under the null hypothesis of one unit root; cf. Dickey and Fuller (1981).

Numerous system-based test procedures have been proposed, with the conceptually most straightforward being that of Johansen (1988) and Johansen and Juselius (1990a).

First, Johansen and Juselius develop a maximum likelihood-based testing procedure for determining the value of r, and tabulate the (asymptotic) critical values of the likelihood ratio (LR) statistic as a function of $p - r$. This statistic generalizes the Dickey-Fuller statistic to the multivariate context. Further, noting that rank($\boldsymbol{\pi}$) is the number of nonzero eigenvalues in a determinantal equation closely related to estimating $\boldsymbol{\pi}$, the LR test ties back directly to $\boldsymbol{\pi}$ by testing how many of those eigenvalues are zero. Additionally, the cointegrating vectors in $\boldsymbol{\beta}'$ are a subset of the eigenvectors, being those associated with the nonzero eigenvalues. Two variants of the LR statistic exist, one using the maximal eigenvalue over a subset of smallest eigenvalues (the "maximal eigenvalue statistic"), the other using all eigenvalues in that subset (the "trace statistic"). These tests and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$ are computed in several of the papers in this special issue.

Second, Johansen and Juselius develop procedures for testing hypotheses about $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}'$, such as zero restrictions. Certain zero restrictions on $\boldsymbol{\alpha}$ correspond to weak exogeneity, and so may be tested (as in Example 8 above). Conversely, weak exogeneity of z_t for the cointegrating vectors $\boldsymbol{\beta}$ is lost if one of those cointegrating vectors appears in both the conditional and marginal densities (i.e., it has nonzero weights in both). Johansen (1990) proposes an ingenious likelihood-based test of weak exogeneity pertaining to the cointegrating vectors. Conditional ECMs by themselves *assume* weak exogeneity, thereby excluding the same cointegrating vector from appearing in both the conditional and marginal processes. If weak exogeneity is valid, cointegration analysis can proceed on the conditional model without loss of information, and Johansen (1990) shows how to do so. Related tests of weak exogeneity have been developed by Boswijk (1991) and Urbain (1991).

Prior to Johansen (1988), Engle and Granger (1987) proposed the use of and established the consistency of unit-root tests in the context of cointegration. Specifically, Engle and Granger proposed testing whether or not an error u_t (defined as $\boldsymbol{\beta}'\boldsymbol{x}_t$) is I(0) by testing whether or not an autoregression in u_t had its roots within the unit circle. Test statistics include the Dickey-Fuller statistic and the Durbin-Watson statistic, using bounds in Sargan and Bhargava (1983) for the latter. Engle and Granger proposed estimating $\boldsymbol{\beta}$ by least squares in a static regression of the variables in $\boldsymbol{x}_t$. Stock (1987), Phillips (1987), and Phillips and Durlauf (1986) derived the asymptotic distribution of that estimator, showing that it is "super-consistent", converging to $\boldsymbol{\beta}$ at $O_p(T^{-1})$ rather than the usual $O_p(T^{-1/2})$.

Nevertheless, the computationally simple Engle-Granger technique suffers from several problems. Inference about $\boldsymbol{\beta}$ depends upon nuisance parameters; and, as Banerjee, Dolado, Hendry, and Smith (1986) demonstrate, large finite-sample biases can result when estimating $\boldsymbol{\beta}$ by a static regression. Unit-root tests applied directly to u_t usually lack power relative to Johansen's test because the latter conditions on the dynamics of the system whereas the former ignores much of that dynamics; cf. Kremers, Ericsson, and Dolado (1989). The number of cointegrating vectors is often of interest, but the Engle-Granger approach lacks means to estimate that number. The choice of normalization in regression affects the finite-sample properties of the Engle-Granger technique. Finally, many hypotheses of interest relate to the complete conditional model specification, and concern speeds of adjustment and the constancy of $\boldsymbol{\beta}$ over time.

C. Policy Implications

Changes in policy-maker's "rules" or reaction functions [such as (28b)] may change the cointegration and/or exogeneity properties of the system. For instance, if a policy maker reacts to the same cointegrating vector as appears in the economic agents' conditional model [e.g., (28a)], weak exogeneity for that cointegrating vector is lost. If a cointegrating vector appears in only the reaction function [e.g., $\gamma_2 = 0$ and $\alpha_2 \neq 0$ in (28)], and the policy maker decides to ignore that disequilibrium information (e.g., changing α_2 to zero), that cointegrating vector disappears from the system. Nevertheless, growth rate (short-run) effects might still be present: for higher-order VARs, lags of Δy_t and Δz_t could enter (28b) even if $\alpha_2 = 0$. More generally, changes in policy-maker's "rules" may identify conditional models as "structural" by demonstrating the conditional models' invariance to switches in policy. Many of the tests developed and applied by the papers in this issue aim to address precisely these issues.

4. This Special Issue on Cointegration and Exogeneity

This issue is divided into two parts: (I) applications, and (II) parameter constancy and predictive accuracy. Papers in the first part test for cointegration and exogeneity on a range of macro-economic data. All of these papers except Johansen's (which is more illustrative) develop and evaluate the specification of their models as well.

Hendry models the demand for TV advertising expenditure in the UK. Both the presence and amounts of advertising on the various TV channels are regulated by the government, so the demand for advertising expenditure has been a highly political issue. Cointegration analysis clarifies the long-run relationships in the data. Modeling the dynamics of prices and quantities on quarterly data obtains a large (absolute) price elasticity and strong within-year feedbacks between prices and quantities. Previous estimated price elasticities were obtained from annual data, and were all small in absolute value. The within-year feedbacks in the quarterly model imply simultaneity bias in the annual estimates, thereby explaining the difference in price elasticities. Since the price elasticity directly affects advertising revenues, which accrue to the commercial TV companies and to a government agency, policy implications are immediate.

The next three papers analyze money demand for the UK, Argentina, and Norway respectively. Johansen describes his (1990) test of weak exogeneity for cointegrating vectors, shows how cointegration analysis is feasible for x_t that are integrated of order two [I(2)], and applies both developments to UK money demand data; cf. Johansen (1991b). Nominal money and prices appear to be I(2), and cointegrate as real money to become I(1). Real money in turn cointegrates with real total final expenditure (the scale variable), interest rates, and inflation to generate an I(0) linear combination. Prices, income, and interest rates appear weakly exogenous for the single cointegrating vector in the system, whereas money is clearly not exogenous.

Ahumada models the demand for notes and coin in Argentina. She applies tests of cointegration to find a single cointegrating vector. Modeling from a general specification and simplifying, Ahumada obtains a parsimonious ECM that satisfies a battery of diagnostic tests. She applies both types of super exogeneity tests, and finds that prices, income, and interest rates appear super exogenous for the parameters of her money-demand equation, in spite of dramatic changes in the processes for those variables.

Bårdsen models the demand for narrow money in Norway. Paralleling the approach in Ahumada's paper, Bårdsen develops a data-coherent, parsimonious ECM for narrow money, which is robust to financial innovation and policy changes during his sample.

Juselius models the domestic and foreign effects on prices in a small open economy, Denmark. Danish inflation may be influenced by deviations from several markets' long-run solutions: purchasing power parity (in the goods market), uncovered interest rate parity (in the assets market), money demand (via a portfolio effect), and the real wage share (via the labor market). Too few observations are available to analyze all markets jointly; and, even if enough observations were available, system analysis of several cointegrating vectors jointly and subsequent system modeling of short-run dynamics appears very difficult in practice. Thus, Juselius analyzes each market separately to determine the corresponding cointegrating vectors, and then models inflation in terms of all four cointegrating vectors, plus short-run dynamics. In brief, foreign effects dominate the determination of Danish inflation, and this result is robust across changes in government and government policy.

Nymoen models the relationship between wages and prices in Finland, distinguishing between real-wage flexibility (i.e., the responsiveness of nominal wages to changes in the price level and productivity) and hysteresis (whereby real wages may be more or less responsive to other determinants, such as unemployment). Theoretically and empirically, this distinction resolves several conflicting results in the literature. Tests of cointegration help identify the existence and extent of real-wage flexibility, and further single-equation analysis determines the extent of hysteresis present. Nymoen's model encompasses previous empirical models, while none of the latter can encompass Nymoen's model.

Hunter describes his recent concept of "cointegrating exogeneity", which has implications for long-run forecasting of cointegrated variables. Hunter illustrates tests of weak and cointegrating exogeneity using Johansen and Juselius's (1990b) data on prices, interest rates, and the exchange rate for the United Kingdom.

The four papers in the second part of this issue contribute directly to the use of tests of parameter constancy and predictive accuracy in policy analysis. At first blush, these papers may appear tangential to cointegration and exogeneity. However, as shown above, that is not so since tests of constancy are critical in testing for super exogeneity and in establishing meaningful long-run relationships. Further, empirically constant parameters are an important element of an economically interpretable model.

Ericsson begins with an exposition of the statistical criteria for model evaluation and design, including various criteria based on forecasts. Using Hendry and Richard's (1982, 1983) taxonomy for model criteria, Ericsson resolves a debate between modelers emphasizing parameter constancy and those running "horse races" based on mean square forecast errors (MSFE). Further, the taxonomy implies a new test statistic, model forecast encompassing, which Ericsson applies to two models of UK money demand. Properties of several of the forecast-based tests are affected by the presence of I(1) and cointegrated variables. Ericsson's Table 2 categorizes numerous model evaluation and design criteria according to Hendry and Richard's taxonomy. Many of those criteria are used by other papers in this issue.

Granger and Deutsch develop tests from the conditional and unconditional forecasts of a model in which the conditioning variable (z_t above) is under policy control. If the conditional model is well-specified and policy "matters" in the conditional model, then the conditional forecasts should be more accurate than the unconditional ones. Specifically, Granger and Deutsch develop a test for whether or not the conditional MSFE is smaller than the unconditional MSFE. If the inequality does not hold, the conditional model is mis-specified, the policy considered is not influential in the conditional model, and/or (in finite samples) the test may lack power. To illustrate these and other forecast-based tests, Granger and Deutsch examine Barro and Rush's Natural Rate/Rational Expectations model of unemployment, Goldfeld and Kane's and Dutkowsky's models of the demand for borrowed reserves, and Hendry and Ericsson's model of UK money demand.

Hansen explains the basis for his recent, general tests of parameter constancy as applied to linear models. Applications include a univariate model of GNP and an error-correction model of consumers' expenditure, both with US data.

Campos derives the confidence intervals for linear combinations of forecasts from general dynamic econometric models. From the associated formulae, tests of parameter constancy are developed. The forecast confidence intervals as well as the constancy of the associated model may be important inputs to the policy-making process. To demonstrate the techniques developed, confidence intervals and test statistics are calculated for models of an oil price and of the Venezuelan CPI.

Appendix

Table of Contents
for the Special Issue of the *Journal of Policy Modeling* entitled
Cointegration, Exogeneity, and Policy Analysis

References

Banerjee, A., J.J. Dolado, J.W. Galbraith, and D.F. Hendry (1991) *Co-integration, Error-correction and the Econometric Analysis of Non-stationary Data*, Oxford: Oxford University Press, forthcoming.

Banerjee, A., J.J. Dolado, D.F. Hendry, and G.W. Smith (1986) "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence", *Oxford Bulletin of Economics and Statistics*, 48, 3, 253-277.

Barndorff-Nielsen, O. (1978) *Information and Exponential Families: In Statistical Theory*, Chichester: John Wiley.

Boswijk, H.P. (1991) "The LM-test for Weak Exogeneity in Error Correction Models", paper presented at the European Meeting of the Econometric Society, Cambridge, England, September.

Campbell, J.Y. and P. Perron (1991) "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots", Technical Working Paper No. 100, Cambridge, Massachusetts: National Bureau of Economic Research.

Chow, G.C. (1960) "Tests of Equality between Sets of Coefficients in Two Linear Regressions", *Econometrica*, 28, 3, 591-605.

Cooley, T.F. and S.F. LeRoy (1981) "Identification and Estimation of Money Demand", *American Economic Review*, 71, 5, 825-844.

Davidson, J.E.H., D.F. Hendry, F. Srba, and S. Yeo (1978) "Econometric Modelling of the Aggregate Time-series Relationship between Consumers' Expenditure and Income in the United Kingdom", *Economic Journal*, 88, 352, 661-692.

Dickey, D.A. and W.A. Fuller (1979) "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, 74, 366, 427-431.

Dickey, D.A. and W.A. Fuller (1981) "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49, 4, 1057-1072.

Dolado, J.J., T. Jenkinson, and S. Sosvilla-Rivero (1990) "Cointegration and Unit Roots", *Journal of Economic Surveys*, 4, 3, 249-273.

Engle, R.F. and C.W.J. Granger (1987) "Co-integration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55, 2, 251-276.

Engle, R.F. and D.F. Hendry (1989) "Testing Super Exogeneity and Invariance", Discussion Paper No. 89-51, San Diego, California: Department of Economics, University of California at San Diego.

Engle, R.F., D.F. Hendry, and J.-F. Richard (1983) "Exogeneity", *Econometrica*, 51, 2, 277-304.

Engle, R.F. and B.S. Yoo (1989) "Cointegrated Economic Time Series: A Survey with New Results", Discussion Paper 87-26R, San Diego, California: Department of Economics, University of California at San Diego; forthcoming in R.F. Engle and C.W.J. Granger (eds.) *Modelling Long-run Economic Relationships: Readings in Cointegration*, Oxford: Oxford University Press.

Ericsson, N.R., J. Campos, and H.-A. Tran (1991) "PC-GIVE and David Hendry's Econometric Methodology", International Finance Discussion Paper No. 406, Washington, D.C.: Board of Governors of the Federal Reserve System; forthcoming in the *Revista de Econometria*.

Ericsson, N.R. and D.F. Hendry (1989) "Encompassing and Rational Expectations: How Sequential Corroboration Can Imply Refutation", International Finance Discussion Paper No. 354, Washington, D.C.: Board of Governors of the Federal Reserve System.

Favero, C. and D.F. Hendry (1989) "Testing the Lucas Critique: A Review", mimeo, Oxford: Nuffield College.

Florens, J.-P. and M. Mouchart (1985a) "Conditioning in Dynamic Models", *Journal of Time Series Analysis*, 6, 1, 15-34.

Florens, J.-P. and M. Mouchart (1985b) "A Linear Theory for Noncausality", *Econometrica*, 53, 1, 157-175.

Granger, C.W.J. (1981) "Some Properties of Time Series Data and Their Use in Econometric Model Specification", *Journal of Econometrics*, 16, 1, 121-130.

Granger, C.W.J. (1986) "Developments in the Study of Cointegrated Economic Variables", *Oxford Bulletin of Economics and Statistics*, 48, 3, 213-228.

Granger, C.W.J. and A.A. Weiss (1983) "Time Series Analysis of Error-correction Models" in S. Karlin, T. Amemiya, and L.A. Goodman (eds.) *Studies in Econometrics, Time Series, and Multivariate Statistics: In Honor of Theodore W. Anderson*, New York: Academic Press, 255-278.

Henderson, J.M. and R.E. Quandt (1971) *Microeconomic Theory: A Mathematical Approach*, New York: McGraw-Hill, Second Edition.

Hendry, D.F. (1979) "Predictive Failure and Econometric Modelling in Macroeconomics: The Transactions Demand for Money", Chapter 9 in P. Ormerod (ed.) *Economic Modelling*, London: Heinemann Education Books, 217-242.

Hendry, D.F. (1983) "Econometric Modelling: The 'Consumption Function' in Retrospect", *Scottish Journal of Political Economy*, 30, 3, 193-220.

Hendry, D.F. (1985) "Monetary Economic Myth and Econometric Reality", *Oxford Review of Economic Policy*, 1, 1, 72-84.

Hendry, D.F. (ed.) (1986) *Economic Modelling with Cointegrated Variables, Oxford Bulletin of Economics and Statistics, Special Issue*, 48, 3.

Hendry, D.F. (1987) "Econometric Methodology: A Personal Perspective", Chapter 10 in T.F. Bewley (ed.) *Advances in Econometrics: Fifth World Congress*, Cambridge: Cambridge University Press, Volume 2, 29-48.

Hendry, D.F. (1988) "The Encompassing Implications of Feedback versus Feedforward Mechanisms in Econometrics", *Oxford Economic Papers*, 40, 1, 132-149.

Hendry, D.F. (1989) *PC-GIVE: An Interactive Econometric Modelling System*, Version 6.0/6.01, Oxford: Institute of Economics and Statistics and Nuffield College, University of Oxford.

Hendry, D.F. (1990) *Lectures on Econometric Methodology* (with D. Qin and C. Favero), Oxford: Oxford University Press, forthcoming.

Hendry, D.F. and N.R. Ericsson (1991a) "An Econometric Analysis of U.K. Money Demand in *Monetary Trends in the United States and the United Kingdom* by Milton Friedman and Anna J. Schwartz", *American Economic Review*, 81, 1, 8-38.

Hendry, D.F. and N.R. Ericsson (1991b) "Modeling the Demand for Narrow Money in the United Kingdom and the United States", *European Economic Review*, 35, 4, 833-881.

Hendry, D.F. and G.E. Mizon (1989) "Evaluating Dynamic Econometric Models by Encompassing the VAR", mimeo, Oxford: Nuffield College; forthcoming in P.C.B. Phillips and V.B. Hall (eds.) *Models, Methods and Applications of Econometrics*.

Hendry, D.F., A.R. Pagan, and J.D. Sargan (1984) "Dynamic Specification", Chapter 18 in Z. Griliches and M.D. Intriligator (eds.) *Handbook of Econometrics*, Amsterdam: North-Holland, Volume 2, 1023-1100.

Hendry, D.F. and J.-F. Richard (1982) "On the Formulation of Empirical Models in Dynamic Econometrics", *Journal of Econometrics*, 20, 1, 3-33; reprinted as Chapter 14 in C.W.J. Granger (ed.) (1990) *Modelling Economic Series: Readings in Econometric Methodology*, Oxford: Oxford University Press, 304-334.

Hendry, D.F. and J.-F. Richard (1983) "The Econometric Analysis of Economic Time Series", *International Statistical Review*, 51, 2, 111-163 (with discussion).

Hylleberg, S. and G.E. Mizon (1989) "Cointegration and Error Correction Mechanisms", *Economic Journal*, 99, 395, Supplement, 113-125.

Johansen, S. (1988) "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 2/3, 231-254.

Johansen, S. (1990) "Cointegration in Partial Systems and the Efficiency of Single Equation Analysis", mimeo, Copenhagen, Denmark: Institute of Mathematical Statistics, University of Copenhagen; forthcoming in the *Journal of Econometrics*.

Johansen, S. (1991a) "Estimating Systems of Trending Variables", invited lecture at the European Meeting of the Econometric Society, Cambridge, England, September; forthcoming in *Econometric Reviews*.

Johansen, S. (1991b) "A Statistical Analysis of Cointegration for I(2) Variables", Research Report No. 77, Helsinki, Finland: Department of Statistics, University of Helsinki.

Johansen, S. and K. Juselius (1990a) "Maximum Likelihood Estimation and Inference on Cointegration – With Applications to the Demand for Money", *Oxford Bulletin of Economics and Statistics*, 52, 2, 169-210.

Johansen, S. and K. Juselius (1990b) "Some Structural Hypotheses in a Multivariate Cointegration Analysis of the Purchasing Power Parity and the Uncovered Interest Parity for UK", Preprint No. 1990:1, Copenhagen, Denmark: Institute of Mathematical Statistics, University of Copenhagen; forthcoming in the *Journal of Econometrics.*

Kendall, M. and A. Stuart (1977) *The Advanced Theory of Statistics: Distribution Theory*, London: Charles Griffin, Fourth Edition, Volume 1.

Koopmans, T.C. (1950) "When Is an Equation System Complete for Statistical Purposes?", Chapter 17 in T.C. Koopmans (ed.) *Statistical Inference in Dynamic Economic Models*, New York: John Wiley, 393-409.

Kremers, J.J.K., N.R. Ericsson, and J.J. Dolado (1989) "The Power of Cointegration Tests", mimeo, Washington, D.C.: Board of Governors of the Federal Reserve System; presented at the World Congress of the Econometric Society, Barcelona, Spain, August 1990.

Lucas, Jr., R.E. (1976) "Econometric Policy Evaluation: A Critique" in K. Brunner and A.H. Meltzer (eds.) *Carnegie-Rochester Conference Series on Public Policy: The Phillips Curve and Labor Markets*, Volume 1, *Journal of Monetary Economics*, Supplement, 19-46.

Nelson, C.R. and C.I. Plosser (1982) "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications", *Journal of Monetary Economics*, 10, 2, 139-162.

Phillips, A.W. (1954) "Stabilisation Policy in a Closed Economy", *Economic Journal*, 64, 254, 290-323.

Phillips, A.W. (1957) "Stabilisation Policy and the Time-Forms of Lagged Responses", *Economic Journal*, 67, 266, 265-277.

Phillips, P.C.B. (1987) "Time Series Regression with a Unit Root", *Econometrica*, 55, 2, 277-301.

Phillips, P.C.B. (1991) "Optimal Inference in Cointegrated Systems", *Econometrica*, 59, 2, 283-306.

Phillips, P.C.B. and S.N. Durlauf (1986) "Multiple Time Series Regression with Integrated Processes", *Review of Economic Studies*, 53, 4, 473-495.

Phillips, P.C.B. and M. Loretan (1991) "Estimating Long-run Economic Equilibria", *Review of Economic Studies*, 58, 3, 407-436.

Richard, J.-F. (1980) "Models with Several Regimes and Changes in Exogeneity", *Review of Economic Studies*, 47, 1, 1-20.

Salmon, M. (1982) "Error Correction Mechanisms", *Economic Journal*, 92, 367, 615-629.

Sargan, J.D. (1964) "Wages and Prices in the United Kingdom: A Study in Econometric Methodology" in P.E. Hart, G. Mills, and J.K. Whitaker (eds.) *Econometric Analysis for National Economic Planning*, Colston Papers, Volume 16, London: Butterworths, 25-63 (with discussion); reprinted as Chapter 10 in D.F. Hendry and K.F. Wallis (eds.) (1984) *Econometrics and Quantitative Economics*, Oxford: Basil Blackwell, 275-314.

Sargan, J.D. and A. Bhargava (1983) "Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk", *Econometrica*, 51, 1, 153-174.

Spanos, A. (1986) *Statistical Foundations of Econometric Modelling*, Cambridge: Cambridge University Press.

Stock, J.H. (1987) "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors", *Econometrica*, 55, 5, 1035-1056.

Suits, D.B. (1955) "An Econometric Model of the Watermelon Market", *American Journal of Agricultural Economics* (formerly the *Journal of Farm Economics*), 37, 2, 237-251.

Tinbergen, J. (1931) "Ein Schiffbauzyklus?", *Weltwirtschaftliches Archiv*, 34, 152-164; reprinted as "A Shipbuilding Cycle?" in L.H. Klaassen, L.M. Koyck, and H.J. Witteveen (eds.) (1959) *Jan Tinbergen: Selected Papers*, Amsterdam: North-Holland, 1-14.

Urbain, J.-P. (1991) "On Weak Exogeneity in Error Correction Models", paper presented at the European Meeting of the Econometric Society, Cambridge, England, September.

International Finance Discussion Papers

IFDP NUMBER	TITLES 1991	AUTHOR(s)
415	Cointegration, Exogeneity, and Policy Analysis: An Overview	Neil R. Ericsson
414	The Usefulness of P^* Measures for Japan and Germany	Linda S. Kole Michael P. Leahy
413	Comments on the Evaluation of Policy Models	Clive W.J. Granger Melinda Deutsch
412	Parameter Constancy, Mean Square Forecast Errors, and Measuring Forecast Performance: An Exposition, Extensions, and Illustration	Neil R. Ericsson
411	Explaining the Volume of Intraindustry Trade: Are Increasing Returns Necessary?	Donald Davis
410	How Pervasive is the Product Cycle? The Empirical Dynamics of American and Japanese Trade Flows	Joseph E. Gagnon Andrew K. Rose
409	Anticipations of Foreign Exchange Volatility and Bid-Ask Spreads	Shang-Jin Wei
408	A Re-assessment of the Relationship Between Real Exchange Rates and Real Interest Rates: 1974 - 1990	Hali J. Edison B. Dianne Pauls
407	Argentina's Experience with Parallel Exchange Markets: 1981-1990	Steven B. Kamin
406	PC-GIVE and David Hendry's Econometric Methodology	Neil R. Ericsson Julia Campos Hong-Anh Tran
405	EMS Interest Rate Differentials and Fiscal Policy: A Model with an Empirical Application to Italy	R. Sean Craig
404	The Statistical Discrepancy in the U.S. International Transactions Accounts: Sources and Suggested Remedies	Lois E. Stekler
403	In Search of the Liquidity Effect	Eric M. Leeper David B. Gordon
402	Exchange Rate Rules in Support of Disinflation Programs in Developing Countries	Steven B. Kamin
401	The Adequacy of U.S. Direct Investment Data	Lois E. Stekler Guy V.G. Stevens

---

Please address requests for copies to International Finance Discussion Papers, Division of International Finance, Stop 24, Board of Governors of the Federal Reserve System, Washington, D.C. 20551.

International Finance Discussion Papers

IFDP NUMBER	TITLES	AUTHOR(s)
	1991	
400	Determining Foreign Exchange Risk and Bank Capital Requirements	Michael P. Leahy
399	Precautionary Money Balances with Aggregate Uncertainty	Wilbur John Coleman II
398	Using External Sustainability to Forecast the Dollar	Ellen E. Meade Charles P. Thomas
397	Terms of Trade, The Trade Balance, and Stability: The Role of Savings Behavior	Michael Gavin
396	The Econometrics of Elasticities or the Elasticity of Econometrics: An Empirical Analysis of the Behavior of U.S. Imports	Jaime Marquez
395	Expected and Predicted Realignments: The FF/DM Exchange Rate during the EMS	Andrew K. Rose Lars E. O. Svensson
394	Market Segmentation and 1992: Toward a Theory of Trade in Financial Services	John D. Montgomery
	1990	
393	*Post* Econometric Policy Evaluation A Critique	Beth Ingram Eric M. Leeper
392	Mercantilism as Strategic Trade Policy: The Anglo-Dutch Rivalry for the East India Trade	Douglas A. Irwin
391	Free Trade at Risk? An Historical Perspective	Douglas A. Irwin
390	Why Has Trade Grown Faster Than Income?	Andrew K. Rose
389	Pricing to Market in International Trade: Evidence from Panel Data on Automobiles and Total Merchandise	Joseph E. Gagnon Michael M. Knetter
388	Is the EMS the Perfect Fix? An Empirical Exploration of Exchange Rate Target Zones	Robert P. Flood Andrew K. Rose Donald J. Mathieson
387	Estimating Pass-through: Structure and Stability	William R. Melick
386	International Capital Mobility: Evidence from Long-Term Currency Swaps	Helen Popper
385	Is National Treatment Still Viable? U.S. Policy in Theory and Practice	Sydney J. Key
384	Three-Factor General Equilibrium Models: A Dual, Geometric Approach	Douglas A. Irwin