Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 354

June 1989

ENCOMPASSING AND RATIONAL EXPECTATIONS: HOW SEQUENTIAL CORROBORATION CAN IMPLY REFUTATION

Neil R. Ericsson and David F. Hendry

ABSTRACT

Even though pieces of empirical evidence individually may corroborate an economic theory, their *joint* existence may refute that same theory. We discuss examples concerning testing for omitted variables, simultaneity, and rational expectations in the context of general-to-simple versus simple-to-general modeling. The proposition in the first sentence strongly favors the building of empirical models which are consistent with *all* available evidence.

Encompassing and Rational Expectations:
How Sequential Corroboration Can Imply Refutation

Neil R. Ericsson and David F. Hendry[1]

1. **Introduction**

Often an economic theory is not tested directly. Rather, empirical evidence is presented as corroborating (or being consistent with) a given theory. Sometimes a single piece of evidence is sufficient to refute a theory (or at least its empirical implementation): e.g., over-identifying restrictions are rejected or coefficient estimates are of the wrong sign. However, the implications of a *set* of evidence can be subtler. Several pieces of empirical evidence may be presented, each of which corroborates a theory, but the *joint* presence of those very pieces may refute that same theory.[2] This proposition, while surprising at first sight, is more obvious upon closer examination; and it has substantive implications for econometric modeling.

Section 2 states and proves the applicable theorem; Sections 3–5 apply that theorem to three areas: omitted variables, simultaneity, and (relatedly) expectations-based models. The theorem (and so the examples) argue for the importance of accounting for a wide variety of evidence (via the encompassing principle) and of accounting for the evidence *as a whole* (congruency).[3] Despite their importance, issues of statistical inference from *finite*

---

[1]The first author is a staff economist in the International Finance Division, Federal Reserve Board. The second author is Professor of Economics at Nuffield College, Oxford, England, and Visiting Research Professor at the Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina. This paper represents the views of the authors and should not be interpreted as reflecting those of the Board of Governors of the Federal Reserve System or other members of its staff. This research was supported in part by UK E.S.R.C. grant R231184. We are grateful for helpful comments from and discussions with Julia Campos, Frank Diebold, Eric Leeper, Jaime Marquez, Doug McManus, and Adrian Neale.

[2]We make no claim to the originality of this proposition, and in fact it appears well-known in the natural sciences; cf. footnote 5. Our interest is in its applications to applied econometrics and so in its implications for econometric methodology.

[3]See Mizon and Richard (1986) on encompassing and Hendry and Richard (1982) on congruence. For detailed discussions of the role of corroboration in a progressive research stragegy, see Popper (1959, Section 82), Lakatos (1970), Boland (1982, ch. 1), and White (1988, ch. 11).

samples are ignored in order to facilitate focusing on the *logical* implications of evidence, given the evidence itself.

2. A Statement of the Theorem

The theorem and proof are so simple as to hardly require formalization: the importance of the theorem lies in its interpretation rather than in the theorem itself. Even so, the formalization clarifies what assumptions are necessary for its application, and what the logical properties of the theorem are (e.g., one of existence rather than inevitability). For ease of exposition, both theories and evidence on theories are viewed as restrictions on (and hence sets in) some observation space.

Consider two theories, A and B, and n pieces of evidence, $\{W_i, i=1,\ldots,n\}$, all interpreted as sets. The relationship between theories and evidence is of interest, so let the operator ϵ_c denote "is consistent with" or equivalently "is corroborated by". Thus, $A \epsilon_c W_i$ ("theory A is corroborated by the evidence W_i") means logically that a subset of theory A lies in the evidence set W_i, or mathematically that the intersection of the sets W_i and A is not empty.

An Existence Theorem. Suppose that, for the n sets $\{W_i, i=1,\ldots,n\}$, (a) their intersection $W^* \equiv (\cap_i W_i)$ is not the empty set, and further that (b) W^* is a *proper* subset of each of the W_i. Then there exist non-empty sets A and B such that:

$$A \ \epsilon_c \ W_i \quad (i=1,\ldots,n)\ , \tag{i}$$

$$B \ \epsilon_c \ W_i \quad (i=1,\ldots,n)\ , \tag{ii}$$

but

$$A \ \epsilon_c \ W^* \tag{iii}$$

and

$$B \ \not\epsilon_c \ W^*\ . \tag{iv}$$

Proof. The proof follows straightforwardly, and by construction. First, for all W_i, let B intersect the part of W_i *not* included in the intersection W^*. (The set A may do so as well.) This is feasible because $(W_i \setminus W^*) \neq \emptyset \ \forall i$ (i.e., for all i, there are elements in W_i not in W^*). Second, let A (but not B) intersect the intersection W^*. This is feasible because W^* is non-empty. By the definitions of the operator ϵ_c and of W^*, A and B satisfy (i)–(iv). QED

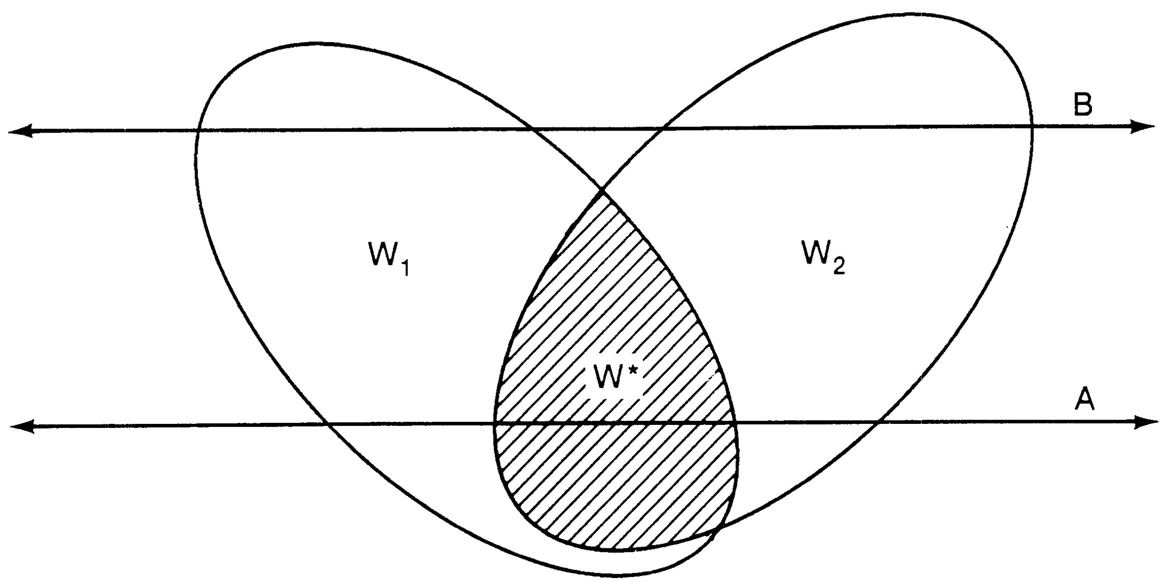


Figure 1. An illustration of the existence theorem for two pieces of evidence (n=2).

Individually, the W_i's corroborate both theories A and B; but observing *all* the W_i's reduces the range of feasible theories, and that range may not include one (or conceivably either) of the theories. Thus, the joint observation of the W_i's may corroborate A and *refute* B.[4] Figure 1 illustrates the theorem for n=2.

The two assumptions serve to exclude trivial solutions. Assumption (a) means that there *are* theories which are non-empty and which are consistent with all the evidence. Assumption (b) means that, for each W_i, at least one of the remaining W_j ($j \neq i$) offers some *additional* information about what theories are acceptable or not. That is, no single piece of evidence implies all the other pieces of evidence, making them redundant. Otherwise, only that single piece of evidence need be considered; and since the theorem is about implications with at least two "distinct" pieces of evidence, the theorem would not apply.

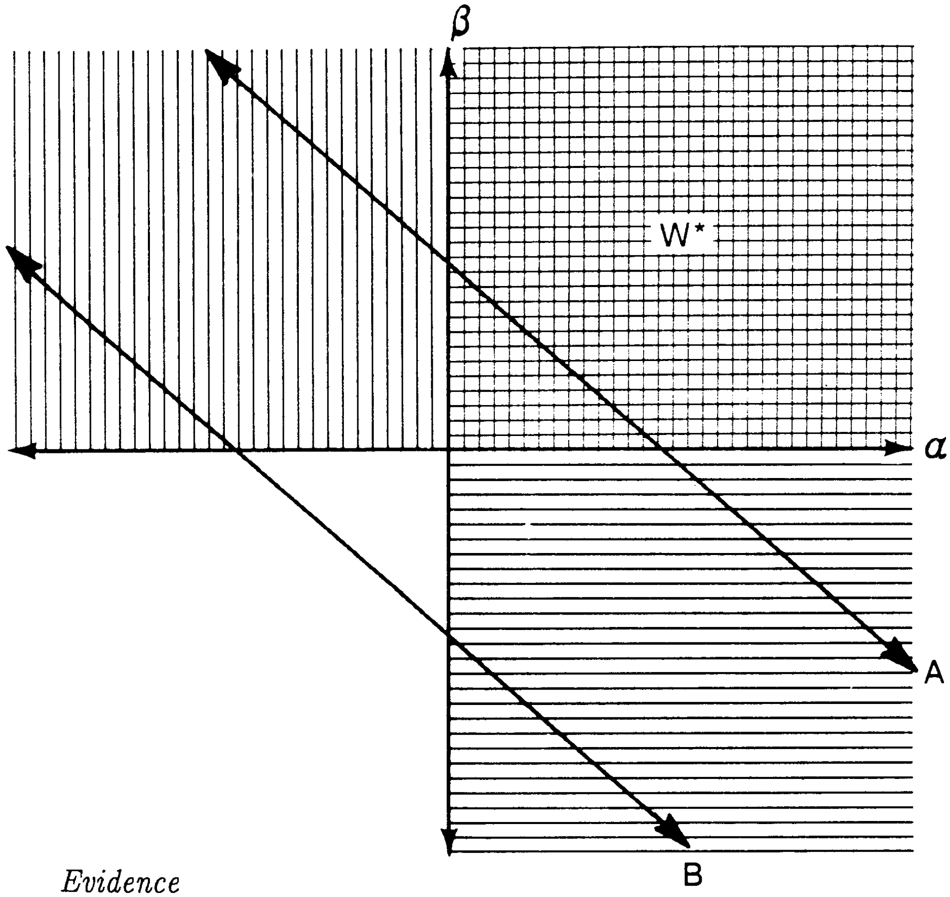
A simple example serves to clarify the theorem.

Example 1: Evidence on a linear restriction. Suppose there are two theories A and B about the parameters α and β such that $A = \{(\alpha,\beta) : \alpha+\beta=1\}$ and $B = \{(\alpha,\beta) : \alpha+\beta=-1\}$, and the pieces of evidence on α and β are $W_\alpha = \{\alpha : \alpha > 0\}$ and $W_\beta = \{\beta : \beta > 0\}$. Observing either W_α or W_β is corroborating evidence for the theory A, and also for theory B. However, observing both W_α and W_β is inconsistent with theory B (for which at least one of α and β must be negative) but corroborates theory A. Figure 2 portrays the relationship between A, B, W_α, and W_β; theory B does not intersect the joint evidence $W^* = W_\alpha \cap W_\beta$ but theory A does. The evidence W_α and W_β in no way proves that theory A is "correct"; it simply does not refute theory A. Further evidence could refute theory A as well (e.g., $W_{\beta*} = \{\beta : \beta > 2\}$).[5]

---

[4]Keynes (1921) *inter alia* attempted to build a probability theory based upon sequential corroboration, but by most accounts failed because of David Hume's "problem of induction"; cf. Boland (1982). Our theorem illustrates why such a theory is not feasible.

[5]Popper (1959, Appendix *IX, ¶5) discusses a related (essentially inverted) example. In particle physics, Shimony (1988) considers two observations on parallel photons. Each observation is consistent both with quantum theory and with local hidden-variables models. Jointly they refute *every* local hidden-variables model, but corroborate quantum theory.



Evidence

$W_\alpha \;=\; \{\, \alpha : \alpha > 0 \,\}$

$W_\beta \;=\; \{\, \beta : \beta > 0 \,\}$

$W^* \;=\; W_\alpha \cap W_\beta$

Theories

$A \;=\; \{\, (\alpha,\beta) : \alpha + \beta = 1 \,\}$

$B \;=\; \{\, (\alpha,\beta) : \alpha + \beta = -1 \,\}$

Figure 2. The relationship between W_α, W_β, W^*, and the theories A and B (Example 1: evidence on a linear restriction).

3. An Application to Testing for Omitted Variables

Because model evaluation (in the form of diagnostic tests) assesses the validity of ignoring certain information in a given model, the problem identified in Section 2 can arise when diagnostic tests are implemented against specific alternatives (and hence specific information sets) and not also against the general alternative which imbeds the specific alternatives.[6] Often, these information sets can be expressed in terms of data, so we consider the situation in which more than one variable is omitted but diagnostic tests are performed for only individual omitted variables. Cf. Hendry and Mizon (1978) for an analysis in the context of autocorrelated residuals, and Hendry (1987) for a taxonomy of test statistics according to the information sets which generate them.

Example 2: Evidence on conditional means. Suppose that the data generation process is:

$$y_t = \alpha x_{1t} + \beta(x_{2t} \cdot x_{3t}) + u_t \qquad u_t \sim \text{NID}(0,\sigma^2) \ , \tag{3.1}$$

where, for simplicity, each x_{it} is normally and independently distributed with zero mean and finite variance and is independent of u_t and the other x_{it}'s. The econometrician posits the following instead of (3.1):

$$y_t = \gamma x_{1t} + v_t \qquad v_t \sim \text{NID}(0,\omega^2) \ , \tag{3.2}$$

and tests for the significance of each of x_{2t} and x_{3t} in determining the conditional mean of y_t. Neither x_{2t} nor x_{3t} individually influence the conditional mean of y_t:

$$E(y_t \,|\, [x_{1t}, x_{it}]) = \alpha x_{1t} \qquad i = 2, 3 \ , \tag{3.3}$$

so no tests based on either x_{2t} or x_{3t} can provide evidence of either being important in explaining the conditional mean of y_t. However, the conditional mean of y_t given all the x_{it}'s is not that given in (3.3), but:

---

[6]This *modus operandi* is one aspect of simple-to-general modeling, a procedure whereby an empirical model is specified, "specific" diagnostic tests are run on the model, and corrections to the model are made in light of those tests. By contrast, in *general-to-simple* modeling, each model is a simplification of the general (maintained) model, so tests of the validity of the implied reduction are always against the common (and joint) information set defined by the general model.

$$E(y_t | [x_{1t}, x_{2t}, x_{3t}]) = \alpha x_{1t} + \beta(x_{2t} \cdot x_{3t}) . \tag{3.4}$$

Thus, the information contained in x_{2t} and x_{3t} jointly (W^*) can refute the hypothesis that y_t depends upon x_{1t} alone (theory B), whereas the information in either x_{2t} or x_{3t} alone (W_1 and W_2) cannot. Figure 1 (also used above) portrays this relationship between theory and evidence.

Davidson, Hendry, Srba, and Yeo's (1978) empirical model of consumers' expenditure in the UK illustrates Example 2. For their model, they show that liquidity is an insignificant determinant of expenditure, maintaining a restriction of long-run unit homogeneity on income. Separately, they show that the restriction of long-run unit homogeneity on income is not rejected when liquidity is excluded. However, Hendry and von Ungern-Sternberg (1981) find that the liquidity-to-income ratio *is* a determinant in a more general model, so showing the importance of examining evidence on income and liquidity jointly rather than just separately.

4. An Application to Testing for Simultaneity

In the standard simultaneous equations framework, estimation by ordinary least-squares is inconsistent, and the degree of inconsistency depends upon the inter-correlations of the equations' disturbances and on the process determining the included endogenous variables. These features of simultaneous equations entail that several sorts of evidence can be helpful in determining whether simultaneity bias actually exists.

Example 3: Evidence on simultaneity. Consider the following system of equations for a single variable y_t (defining the equation of interest) and a vector of variables $\mathbf{x}_t$. Unless otherwise indicated, bold characters are vectors (lower case) or matrices (upper case); lower case non-bold characters are scalars.

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + e_t \tag{4.1}$$

$$\mathbf{x}_t = \boldsymbol{\Pi}\mathbf{z}_t + \mathbf{v}_t \tag{4.2}$$

The parameter of interest is $\boldsymbol{\beta}$, $E(\mathbf{z}_t\mathbf{v}_t')=\mathbf{0}$, $E(\mathbf{z}_t e_t)=\mathbf{0}$, and $\boldsymbol{\Pi}\neq\mathbf{0}$. The first expectation defines $\boldsymbol{\Pi}$ as the matrix of reduced form coefficients for $\mathbf{x}_t$; the second expectation, along with $\Pi \neq 0$, ensures that $\mathbf{z}_t$ is a valid set of instruments for estimating β. Trivially, the instrumental variables (here, 2SLS) estimator is consistent:

$$\delta \equiv \operatorname{plim} \beta_{IV} = \beta , \tag{4.3}$$

provided $\Pi E(\mathbf{z}_t \mathbf{z}_t')\Pi'$ is non-singular (i.e., there are enough valid non-redundant instruments). The least-squares estimator of β may or may not be consistent, depending upon the covariance between $\mathbf{v}_t$ and e_t:

$$\begin{aligned} \gamma \equiv \operatorname{plim} \beta_{LS} &= \beta + [E(\mathbf{x}_t \mathbf{x}_t')]^{-1} \cdot E(\mathbf{x}_t e_t) \\ &= \beta + [\Pi \mathbf{M}_{zz} \Pi' + \Sigma_{vv}]^{-1} \cdot \Sigma_{ve} , \end{aligned} \tag{4.4}$$

where $\mathbf{M}_{zz} = E(\mathbf{z}_t \mathbf{z}_t')$, $\Sigma_{vv} = E(\mathbf{v}_t \mathbf{v}_t')$, and $\Sigma_{ve} = E(\mathbf{v}_t e_t)$; see Bronfenbrenner (1953). Hausman's test statistic compares β_{IV} and β_{LS} to see whether they are equal (and hence whether $\Sigma_{ve}=0$), and so whether there is any simultaneity bias with least squares. However, additional evidence may exist on whether or not simultaneity is an issue.

Equation (4.2) is a "reduced form" for $\mathbf{x}_t$, so Π and Σ_{vv} may well be non-constant over time, being (potentially) complicated functions of economic, policy, and institutional parameters. To the extent that Π and/or Σ_{vv} change, γ also will change if simultaneity bias exists. To link this to the issue of corroboration, the evidence on the constancy of various parameters is categorized as follows.

W_1: γ is constant.

W_2: δ is constant.

W_3: Π and/or Σ_{vv} are non-constant.

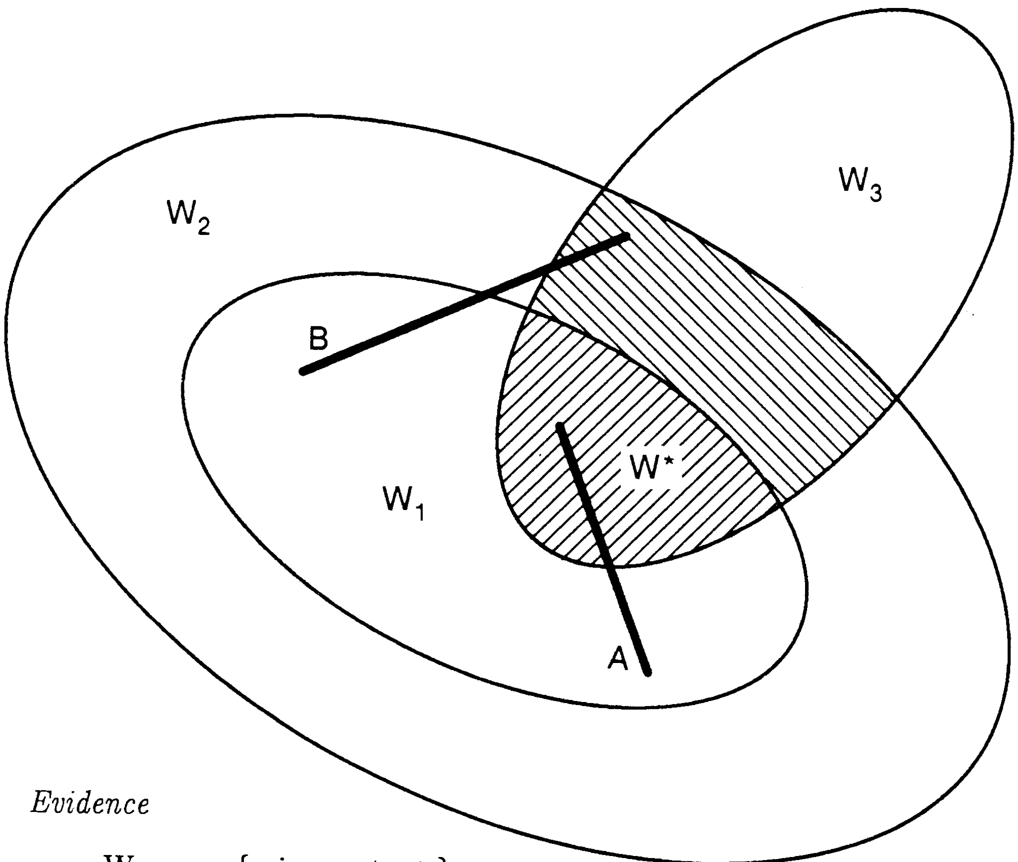
The theories of interest are:

A: y_t and $\mathbf{x}_t$ are *not* simultaneously determined ($\Sigma_{ve}=0$), and

B: y_t and $\mathbf{x}_t$ *are* simultaneously determined ($\Sigma_{ve} \neq 0$).

For both theories, β is the parameter of interest and is assumed constant throughout.

The W_i individually corroborate both theories. For theory A, W_1 is a necessary condition, W_2 is implied by W_1 and the validity of the instruments $\mathbf{z}_t$, and W_3 is irrelevant. For theory B, W_2 is implied by the constancy of β, W_3 does not violate any assumptions of B, and W_1 could occur if (e.g.) $(y_t, \mathbf{x}_t', \mathbf{z}_t')$ were jointly stationary. In the last instance,



Evidence

$W_1 = \{\gamma \text{ is constant.}\}$

$W_2 = \{\delta \text{ is constant.}\}$

$W_3 = \{\Pi \text{ and/or } \mathbf{\Sigma}_{vv} \text{ are non-constant.}\}$

$(W_2 \cap W_3) \setminus W_1$ (The Lucas critique is confirmed.)

$W_1 \cap W_2 \cap W_3$ (The Lucas critique is refuted.)

Theories

A = $\{y_t$ and $\mathbf{x}_t$ are *not* simultaneously determined $(\mathbf{\Sigma}_{ve}=0)$.$\}$
B = $\{y_t$ and $\mathbf{x}_t$ *are* simultaneously determined $(\mathbf{\Sigma}_{ve}\neq 0)$.$\}$

Figure 3. The relationship between W_1, W_2, W_3, W^*, and specific A and B (Example 3: evidence on simultaneity).

estimation (whether OLS or IV) would generate constant coefficients, regardless of which variables were included or excluded, because both estimators are functions of the sample data moments (which converge to respective population moments), and the population moments are constant. Thus, with (y_t, x_t', z_t') jointly stationary, *both* W_1 and W_2 follow (but W_3 cannot).

The implications of *joint* occurrences of the W_i follow straightforwardly. Either $(W_1 \cap W_2)$ or $(W_2 \cap W_3)$ corroborates B. However, if $(W_2 \cap W_3)$ occurs, then B implies that W_1 cannot occur, i.e., OLS is inconsistent and its inconsistency varies as Π and Σ_{vv} vary.[7] Thus, W_1, W_2, and W_3 jointly *refute* B whereas individually they corroborate it. Conversely, $(W_1 \cap W_2 \cap W_3)$ corroborates A. Thus, the constancy or otherwise of the parameters in (4.1)–(4.2) can have implications for whether or not, as a *logical* issue, y_t and $\mathbf{x}_t$ could have been generated simultaneously. Figure 3 illustrates Example 3 (and 4a).

A simple illustration of Example 3 helps clarify what is happening.

Example 3: Evidence on simultaneity with an AR(1) x_t process.* Suppose that $\mathbf{x}_t$ is a univariate AR(1) process:

$$x_t = \pi x_{t-1} + v_t \qquad v_t \sim \mathrm{NI}(0, \sigma_{vv}) \ , \tag{4.2a}$$

where $0<|\pi|<1$. If $z_t = x_{t-1}$, (4.4) becomes:

$$\gamma = \beta + (1-\pi^2)\cdot(\sigma_{ve}/\sigma_{vv}) \ . \tag{4.4a}$$

If $\sigma_{ve} \neq 0$ and π and/or σ_{vv} change, then γ changes and so W_1 would not be the case. Conversely, if γ remains constant in spite of π and/or σ_{vv} changing, then $\sigma_{ve}=0$ and there is no simultaneity.

5. Applications to Testing Feedback *versus* Feedforward Models

The final set of illustrations turns on Hendry's (1988) proposal for testing feedback (conditional) versus feedforward (expectations-based) models by applying the encompassing principle to evidence on the constancy or otherwise of the conditional model and of the marginal process for the conditioning variables. We begin by interpreting a process with

[7] There is a set of variations in Π and Σ_{vv} such that their effects just cancel each other. We ignore this set because it is of measure zero.

expectations as being part of a simultaneous equations system, in which case the formulae in Section 4 for the inconsistency of OLS apply. Examples with the univariate AR(1) $\mathbf{x}_t$ process motivate and clarify the more general proposition.

Example 4: Evidence on expectations. The feedforward framework can be characterized in the following manner. Agents make decisions about a variable y_t in light of their expectations about future values of some strictly exogenous variables $\mathbf{x}_t$ and about future values of y_t itself. Expectations are formed, conditional upon an information set (denoted $\mathbf{w}_{t-1}$) which typically includes lagged values of $\mathbf{x}_t$ and y_t.

$$\sum_{i=0}^{\infty} \kappa_i E(y_{t+i} | \mathbf{w}_{t-1}) = \sum_{i=0}^{\infty} \lambda_i' E(\mathbf{x}_{t+i} | \mathbf{w}_{t-1}) \tag{5.1}$$

$\kappa_0=1$, the λ_i's and the remaining κ_i's are the "deep" parameters of the agents' behavior (or direct functions of them), and typically $E(y_t | \mathbf{w}_{t-1}) \equiv y_t$. For illustrative purposes, we have assumed linearity, but the issues raised also apply to nonlinear models involving expectations. By repeated substitution of (5.1) into itself, the conditional expectation of y_t can be expressed in terms of the expectations of the $\mathbf{x}_{t+i}$'s alone:

$$E(y_t | \mathbf{w}_{t-1}) = \sum_{i=0}^{\infty} \delta_i' E(\mathbf{x}_{t+i} | \mathbf{w}_{t-1}) , \tag{5.2}$$

where δ_i depends upon $\{\lambda_i, \kappa_i\}$ and $\sum_{i=0}^{\infty} \delta_i'$ is finite. For completeness (and, e.g., estimation), it is necessary to specify the process for $\mathbf{x}_t$, which is:

$$E(\mathbf{x}_t | \mathbf{w}_{t-1}) = \Pi \mathbf{x}_{t-1} , \tag{5.3}$$

noting the strict exogeneity of $\mathbf{x}_t$ and again assuming linearity.[8] Equations (5.2) and (5.3) can be written in "model form" as:

$$y_t = \sum_{i=0}^{\infty} \delta_i' E(\mathbf{x}_{t+i} | \mathbf{w}_{t-1}) + \epsilon_t \qquad E(\epsilon_t \cdot \mathbf{w}_{t-1})=\mathbf{0} \tag{5.4}$$

$$\mathbf{x}_t = \Pi \mathbf{x}_{t-1} + \mathbf{v}_t \qquad E(\mathbf{v}_t \cdot \mathbf{x}_{t-1}')=\mathbf{0} . \tag{5.5}$$

Equation (5.5) parallels the reduced form for $\mathbf{x}_t$ in (4.2) with $\mathbf{z}_t=\mathbf{x}_{t-1}$. As noted above, the expectation of y_t in (5.2) and its realization frequently are taken to be the same; but

[8]If $\mathbf{x}_t$ depends upon several of its own lags, rather than just one, that dependence can be rewritten as (5.3) by "stacking" the lags and redefining $\mathbf{x}_t$.

we allow possible discrepancies through non-zero ϵ_t. For expositional convenience, we assume $(\epsilon_t, \mathbf{v}_t)$ is independently and identically distributed, normal.

Historically, many macro-econometric equations have been estimated by ordinary least-squares and with actual values of $\mathbf{x}_t$ rather than expectations of its current and future values, thus assuming that it is valid to condition on $\mathbf{x}_t$ itself and that future expectations are unimportant. That implies a conditional model (i.e., conditional on observed $\mathbf{x}_t$):

$$y_t = \gamma' \mathbf{x}_t + \nu_t \qquad E(\nu_t \cdot \mathbf{x}_t) = \mathbf{0} \ , \tag{5.6}$$

where γ is the parameter of interest. However, assuming (5.4) and (5.5), γ is a derived parameter and is a complicated function of $\{\delta_i\}$ and Π. To express y_t in (5.4) explicitly in terms of the observed $\mathbf{x}_t$ requires two steps: repeated substitution of (5.5) into itself to obtain $E(\mathbf{x}_{t+i} | \mathbf{w}_{t-1}) = \Pi^i E(\mathbf{x}_t | \mathbf{w}_{t-1}) = \Pi^i(\mathbf{x}_t - \mathbf{v}_t)$, and then direct solution of (5.4):

$$y_t = \left(\sum_{i=0}^{\infty} \delta_i' \Pi^i\right) \cdot \mathbf{x}_t + \left[\epsilon_t - \left(\sum_{i=0}^{\infty} \delta_i' \Pi^i\right) \cdot \mathbf{v}_t\right] \ , \tag{5.7}$$

assuming that $\mathbf{x}_t$ is stationary. With (5.5), this representation parallels (4.1)–(4.2), in which $\beta' = \Sigma_0^{\infty} \delta_i' \Pi^i$, $e_t = [\epsilon_t - (\Sigma_0^{\infty} \delta_i' \Pi^i) \cdot \mathbf{v}_t]$, and $\mathbf{z}_t = \mathbf{x}_{t-1}$. Thus, (4.4) is appropriate for calculating the value to which the least-squares estimator of the coefficient on $\mathbf{x}_t$ in (5.6) converges.

$$\gamma = \left(\sum_{i=0}^{\infty} \{\Pi'\}^i \delta_i\right) + [\mathbf{M}_{xx}]^{-1} \cdot \left[\Sigma_{v\epsilon} - \Sigma_{vv}\left(\sum_{i=0}^{\infty} \{\Pi'\}^i \delta_i\right)\right] \tag{5.8}$$

$\mathbf{M}_{xx}^{\vartheta} \equiv E(\mathbf{x}_t \mathbf{x}_t')^{\vartheta} = (I - \Pi \otimes \Pi)^{-1} \Sigma_{vv}^{\vartheta}$, ϑ is the column vectoring operator, and $\otimes$ is the corresponding Kronecker product.[9] The coefficient γ is constant if Π, Σ_{vv}, $\Sigma_{v\epsilon}$, and the δ_i's are. However, as changes occur in the process generating $\mathbf{x}_t$ (e.g., Π varies over time), γ also will change and the conditional model (5.6) will "break down". The Lucas critique

---

[9] If (5.1) were a nonlinear rather than a linear difference equation in expectations, then (5.8) would include an approximation error due to the linearization of that difference equation by (5.2). The least-squares estimator for (5.6) still would have a probability limit similar in form to γ in (5.8), but with an additional term introduced by the approximation error; cf. White (1980). Likewise, the conditional expectation for $\mathbf{x}_t$ in (5.3) might be nonlinear, in which case (5.5) with Π would be the least-squares approximation to that nonlinear function. Such nonlinearities *per se* could not induce non-constancy in either γ or Π although, in finite samples, apparent non-constancy might be detected if (e.g.) $\mathbf{x}_t$ were very slowly changing (e.g., very autoregressive). Additional data would clarify that there was constancy.

applies, and (5.6) fails to isolate the δ_i's (and so the underlying structural parameters); cf. Lucas (1976). Equation (5.1) (and so (5.4)) remains constant in spite of (5.5) evolving. Conversely, if γ is constant in spite of changes in the process for $\mathbf{x}_t$, y_t could not have been generated by (5.1) with *constant* "deep" parameters. In this case, the Lucas critique is "refuted". That is, because of (5.8), (5.1) is inconsistent with the observations that γ is constant *and* that Π and/or Σ_{vv} have changed.

To relate this to Sections 2 and 4, the evidence is categorized as follows.

W_1: γ is constant.

W_2: $\{\lambda_i\}$ and $\{\kappa_i\}$ (and so $\{\delta_i\}$) are constant.

W_3: Π and/or Σ_{vv} are non-constant.

The theories of interest are:

A: y_t is determined conditional upon $\mathbf{x}_t$, and

B: y_t is determined via expectations of future y_t and of current and future $\mathbf{x}_t$.

Unlike Section 4, the parameters of interest for the two hypotheses generally are not the same, being γ for one and the deep structural parameters for the other. However, under each hypothesis, the corresponding parameters of interest generally are claimed to be invariant to changes in the distribution of $\mathbf{x}_t$ (e.g., changes in policy rules); otherwise, the model would not be valid for two main purposes, forecasting and policy simulation, nor would the parameters have an "economic" interpretation.

Given (5.7) and with the choice of $\{W_i\}$, the discussion of corroboration and refutation for simultaneity applies to the expectations model, with one difference: in general W_1 does not imply W_2, so $(W_1 \cap W_3)$ need not intersect W_2.[10] The following examples

---

[10]For example, suppose the conditional model is correct (with constant γ (W_1)) and x_t is univariate AR(1):

$$y_t = \gamma x_t + \nu_t$$
$$x_t = \pi x_{t-1} + v_t \quad ,$$

but that an expectations model is estimated with the expectations of a single future x_t:

$$y_t = \delta_r E(x_{t+r} | w_{t-1}) + \epsilon_t \quad ,$$

where $w_{t-1} = x_{t-1}$. In that case, $E(x_{t+r} | w_{t-1}) = \pi^r (x_t - v_t)$, so substitution gives:

$$y_t = (\gamma \pi^{-r}) E(x_{t+r} | w_{t-1}) + (\nu_t + \gamma v_t) \quad .$$

Thus, $\delta_r = \gamma \pi^{-r}$ because $E([\nu_t + \gamma v_t] | w_{t-1}) = 0$, so δ_r changes as π changes. Further, if π changes (W_3) and δ_r is constant (W_2), then r=0. That exception is Example 4a.

illustrate the implications for special cases of (5.4) and (5.5), namely, when only certain δ_i are non-zero and when $\mathbf{x}_t$ is univariate AR(1).

Example 4a: Evidence on expectations with $\delta_0 \neq 0$, $\delta_i = 0\ \forall i > 0$. In this case, (5.4) simplifies to:

$$y_t = \delta_0' E(\mathbf{x}_t | \mathbf{w}_{t-1}) + \epsilon_t \qquad E(\epsilon_t \cdot \mathbf{w}_{t-1}) = \mathbf{0} . \tag{5.4a}$$

By substituting (5.5) into (5.4a), we obtain:

$$y_t = \delta_0' \mathbf{x}_t + (\epsilon_t - \delta_0' \mathbf{v}_t) . \tag{5.7a}$$

Thus, the expectations model defined by (5.4a) and (5.5) is equivalent to the simultaneous equations model (4.1)–(4.2), with $e_t = (\epsilon_t - \delta_0' \mathbf{v}_t)$, $\mathbf{z}_t = \mathbf{x}_{t-1}$, and $\beta = \delta_0$. From (4.4), simultaneity bias arises in estimating δ_0 in (5.4a) by using realized values of $\mathbf{x}_t$ rather than its expectation.

$$\gamma = \delta_0 + [\mathbf{M}_{xx}]^{-1} \cdot (\Sigma_{v\epsilon} - \Sigma_{vv} \delta_0) , \tag{5.8a}$$

which is not δ_0 even if $\mathbf{v}_t$ and ϵ_t are uncorrelated. See Hendry (1988) for details. With $\sigma_{v\epsilon} = 0$ and x_t a stationary univariate AR(1) process, (5.8a) simplifies:

$$\gamma = \pi^2 \delta_0 . \tag{5.8a$'$}$$

(Nb. x_{t-1} is not a valid instrument if π is zero.)

Example 4b: Evidence on expectations with $\delta_r \neq 0$, $\delta_i = 0\ \forall i \neq r$, $i, r \geq 0$. In this case, (5.4) simplifies to:

$$y_t = \delta_r' E(\mathbf{x}_{t+r} | \mathbf{w}_{t-1}) + \epsilon_t \qquad E(\epsilon_t \cdot \mathbf{w}_{t-1}) = \mathbf{0} , \tag{5.4b}$$

and so:

$$y_t = \delta_r' \Pi^r \mathbf{x}_t + (\epsilon_t - \delta_r' \Pi^r \mathbf{v}_t) . \tag{5.7b}$$

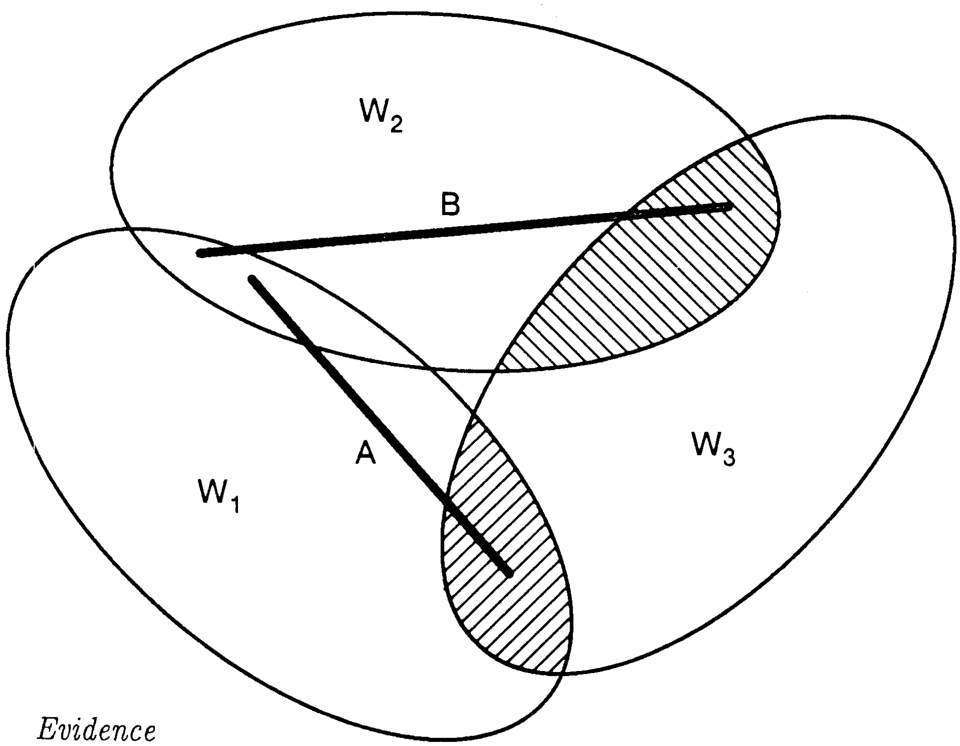
This again parallels (4.1), so the coefficient defined by $E(y_t | \mathbf{x}_t)$ is:

$$\gamma = (\Pi')^r \delta_r + [\mathbf{M}_{xx}]^{-1} \cdot [\Sigma_{v\epsilon} - \Sigma_{vv} (\Pi')^r \delta_r] . \tag{5.8b}$$

With $\sigma_{v\epsilon} = 0$ and a stationary univariate AR(1) process for $\mathbf{x}_t$:

$$\gamma = \pi^{r+2} \delta_r . \tag{5.8b$'$}$$

Figure 4 portrays Example 4 in general, and 4b for r>0. Table 1 summarizes the formulae for γ as a function of the (simultaneous or expectations-based) process generating the data.



Evidence

$W_1 = \{\gamma$ is constant.$\}$
$W_2 = \{\{\lambda_i\}$ and $\{\kappa_i\}$ (and so $\{\delta_i\}$) are constant.$\}$
$W_3 = \{\Pi$ and/or Σ_{vv} are non-constant.$\}$

$(W_2 \cap W_3) \setminus W_1$ (The Lucas critique is confirmed.)

$W_1 \cap W_3$ (The Lucas critique is refuted.)

Theories

$A = \{y_t$ is determined conditional upon $\mathbf{x}_t$.$\}$

$B = \{y_t$ is determined via expectations of future y_t and of current and future $\mathbf{x}_t$.$\}$

Figure 4. The relationship between W_1, W_2, W_3, and specific A and B (Example 4: evidence on expectations).

Table 1: Formulae for γ ($\equiv$ plim β_{LS}) as a function of the data generation process.

	The process generating $\mathbf{x}_t$[a]	
Example	vector autoregressive	univariate AR(1)
	Simultaneity	
3	$\beta + [\Pi \mathbf{M}_{zz} \Pi' + \boldsymbol{\Sigma}_{vv}]^{-1} \cdot \boldsymbol{\Sigma}_{ve}$	$\beta + (1-\pi^2) \cdot (\sigma_{ve}/\sigma_{vv})$
	Expectations	
4	$(\sum_{i=0}^{\infty} \{\Pi'\}^i \delta_i) + [\mathbf{M}_{xx}]^{-1} \cdot [\boldsymbol{\Sigma}_{v\epsilon} - \boldsymbol{\Sigma}_{vv}(\sum_{i=0}^{\infty} \{\Pi'\}^i \delta_i)]$	$\pi^2 \sum_{i=0}^{\infty} \pi^i \delta_i$
4a[b]	$\delta_0 + [\mathbf{M}_{xx}]^{-1} \cdot (\boldsymbol{\Sigma}_{v\epsilon} - \boldsymbol{\Sigma}_{vv} \delta_0)$	$\pi^2 \delta_0$
4b[c]	$(\Pi')^r \delta_r + [\mathbf{M}_{xx}]^{-1} \cdot [\boldsymbol{\Sigma}_{v\epsilon} - \boldsymbol{\Sigma}_{vv} (\Pi')^r \delta_r]$	$\pi^{r+2} \delta_r$

Notes:

a. The second moment matrix for $\mathbf{x}_t$ (denoted $\mathbf{M}_{xx}$) can be expressed as $[\Pi \mathbf{M}_{zz} \Pi' + \boldsymbol{\Sigma}_{vv}]$, and is equal to $\sigma_{vv}/(1-\pi^2)$ when $\mathbf{x}_t$ is univariate AR(1). For the expectations examples with $\mathbf{x}_t$ being univariate AR(1), $\sigma_{v\epsilon}$ is set to zero.

b. For Example 4a, $\delta_0 \neq 0$ and $\delta_i = 0$ $\forall i > 0$.

c. For Example 4b, $\delta_r \neq 0$ and $\delta_i = 0$ $\forall i \neq r$, $i, r \geq 0$.

Hendry's (1985) empirically constant conditional model of the demand for M_1 in the UK illustrates both Examples 3 and 4. Cuthbertson (1988) seeks to reinterpret Hendry's model as a reduced form of a forward-looking process for money demand. However, Hendry (1988) establishes that the marginal expectations processes for income, prices, and the interest rate are not constant over the sample, so Cuthbertson's interpretation is precluded. More generally, Nickell (1985) and Campbell and Shiller (1988) *inter alia* note a possible isomorphism between conditional error-correction models and rational expectations models. As implied by the results above, that equivalence does not hold when the marginal process changes over time: the *data* can resolve the interpretation of a constant conditional model. Cf. Hendry and Ericsson (1988) on the demand for money and Campos and Ericsson (1988) on consumers' expenditure for other recent empirical applications in which expectations-based models logically could not encompass the results obtained.

Before concluding, three remarks are in order. First, complete encompassing of conditional models by expectations models could be used instead of the "limited" encompassing proposed by Hendry (1988). For instance, given values for π and δ_0 in Example 4a, a prediction of γ could be constructed from (5.8a′) and compared with the observed estimate of γ from (5.6). Although intuitively appealing and well-founded theoretically, such complete encompassing can prove exceedingly difficult when $\mathbf{x}_t$ is determined by a complicated (and unknown) process. Even so, fully efficient estimation of the expectations model requires properly specifying that process for $\mathbf{x}_t$. Conditional models only require estimation of the conditional equation for efficiency.

Second, and relatedly, one advantage of Hendry's proposal for refuting the Lucas critique is that the full process for $\mathbf{x}_t$ in (4.2) need not be specified: identifying a subset of the $\mathbf{z}_t$ is sufficient. A proof appears in Hendry (1988), and the standard formula for omitted variables bias provides the intuition. If some additional set of variables $\mathbf{z}_t^*$ is required in (4.2) to make it complete:

$$\mathbf{x}_t = \Psi_1\mathbf{z}_t + \Psi_2\mathbf{z}_t^* + \mathbf{v}_t^* , \tag{5.9}$$

then the least-squares estimator of Π in (4.2) is subject to omitted variables bias. By assumption, Π changes (W_3). From the bias formula, it could have done so for one (or more) of three reasons: the underlying coefficient Ψ_1 on $\mathbf{z}_t$ changed; the coefficient Ψ_2 on $\mathbf{z}_t^*$ changed; or the correlation (Φ, say) between $\mathbf{z}_t^*$ and $\mathbf{z}_t$ changed. Because γ is a function of Π, which in turn is a function of Ψ_1, Ψ_2, and Φ, then γ will change as any of Ψ_1, Ψ_2, and Φ change, excepting coincidental cancellation due to equivalent variations in parameters.

Third, in practice, W_1, W_2, and W_3 are not known, but the corresponding coefficients can be tested for constancy, e.g., using Chow's (1960) statistic in a recursive framework or Hoffman and Pagan's (1988) and Ghysels and Hall's (1988) statistic (generalizing upon Chow) for the GMM estimator. Thus, actual inferences about empirical models in light of evidence implicitly or explicitly will have varying degrees of uncertainty associated with them.

6. Conclusions

Because a sequence of apparently confirming evidence can actually refute a theory, it is important to examine *all* available evidence on an empirical model *jointly* rather than simply corroborate a subset of the implications of a theory.[11] Only well-tested theories that have successfully weathered tests outside the control of their proponents and can explain the gestalt of existing empirical evidence seem likely to provide a useful basis for applied economic analysis and policy. That means *encompassing* the evidence with a *congruent* empirical model. We cannot do better than cite Milton Friedman in support of this view.

> It is one of our chief defects that we place all too much emphasis on the derivation of hypotheses and all too little on testing their validity. This distortion of emphasis is frequently unavoidable, resulting from the absence of widely accepted and objective criteria for testing the validity of hypotheses in the social sciences. But this is not the whole story. Because we cannot adequately test the validity of many hypotheses, we have fallen into the habit of not trying to test the validity of hypotheses even when we can do so. We examine evidence, reach a conclusion, set it forth, and rest content, neither asking ourselves what evidence might contradict our hypothesis nor seeking to find out whether it does. Friedman (1951, p. 107)

[11]This appears closely related to why "ordinary" encompassing is not transitive but parsimonious encompassing is; cf. Hendry and Richard (1987).

> As I shall argue at greater length below, the only relevant test of the *validity* of a hypothesis is comparison of its predictions with experience. The hypothesis is rejected if its predictions are contradicted ("frequently" or more often than predictions from an alternative hypothesis); it is accepted if its predictions are not contradicted; great confidence is attached to it if it has survived many opportunities for contradiction. Factual evidence can never "prove" a hypothesis; it can only fail to disprove it, which is what we generally mean when we say, somewhat inexactly, that the hypothesis has been "confirmed" by experience. Friedman (1953, pp. 8–9) (italics in original)

To avoid potential apparent paradoxes such as those discussed in Sections 2–5, it seems crucial to conduct inference within the framework of general to simple, at least implicitly so by always testing a conjectured model against the most unrestricted model that is logically entailed by the evidence.[12]

[12]For a similar argument, see Pagan (1987, p. 6).

References

Boland, L.A. (1982) *The Foundations of Economic Method*, London, George Allen and Unwin.

Bronfenbrenner, J. (1953) "Sources and Size of Least-squares Bias in a Two-equation Model", Chapter 9 in Wm.C. Hood and T.C. Koopmans (eds.) *Studies in Econometric Method*, New Haven, Yale University Press, 221–235.

Campbell, J.Y. and R.J. Shiller (1988) "Interpreting Cointegrated Models", *Journal of Economic Dynamics and Control*, 12, 2/3, 505–522.

Campos, J. and N.R. Ericsson (1988) "Econometric Modeling of Consumers' Expenditure in Venezuela", International Finance Discussion Paper No. 325, Board of Governors of the Federal Reserve System, Washington, D.C.

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

Cuthbertson, K. (1988) "The Demand for M_1: A Forward Looking Buffer Stock Model", *Oxford Economic Papers*, 40, 1, 110–131.

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.

Friedman, M. (1951) "Comment" in G. Haberler (ed.) *Conference on Business Cycles*, New York, National Bureau of Economic Research, 107–114.

Friedman, M. (1953) "The Methodology of Positive Economics" in M. Friedman *Essays in Positive Economics*, Chicago, University of Chicago Press.

Ghysels, E. and A.D. Hall (1988) "A Test for Structural Stability of Euler Conditions Parameters Estimated via GMM Estimator", paper presented at the European Meetings of the Econometric Society, Bologna, Italy.

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

Hendry, D.F. (1987) "Econometric Methodology: A Personal Perspective", Chapter 10 in T.F. Bewley (ed.) *Advances in Econometrics*, 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. and N.R. Ericsson (1988) "An Econometric Analysis of UK Money Demand in *Monetary Trends in the United States and the United Kingdom* by Milton Friedman and Anna J. Schwartz", mimeo, Board of Governors of the Federal Reserve System, Washington, D.C.; forthcoming, *American Economic Review.*

Hendry, D.F. and G.E. Mizon (1978) "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England", *Economic Journal*, 88, 351, 549–563.

Hendry, D.F. and J.-F. Richard (1982) "On the Formulation of Empirical Models in Dynamic Econometrics", *Journal of Econometrics*, 20, 1, 3–33.

Hendry, D.F. and J.-F. Richard (1987) "Recent Developments in the Theory of Encompassing", C.O.R.E. Discussion Paper No. 8722, Louvain-la-Neuve, Belgium, forthcoming in *CORE Twentieth Anniversary Volume*, Cambridge, MIT Press.

Hendry, D.F. and Th. von Ungern-Sternberg (1981) "Liquidity and Inflation Effects on Consumers' Expenditure", Chapter 9 in A.S. Deaton (ed.) *Essays in the Theory and Measurement of Consumer Behaviour*, Cambridge, Cambridge University Press, 237–260.

Hoffman, D. and A. Pagan (1988) "Post-sample Prediction Tests for Generalized Method of Moments Estimators", Working Paper No. 129, University of Rochester, Rochester, N.Y.

Keynes, J.M. (1921) *A Treatise on Probability*, London, Macmillan.

Lakatos, I. (1970) "Falsification and the Methodology of Scientific Research Programmes" in I. Lakatos and A. Musgrave (eds.) *Criticism and the Growth of Knowledge*, Cambridge, Cambridge University Press, 91–196.

Lucas, Jr., R.E. (1976) "Econometric Policy Evaluation: A Critique" in K. Brunner and A.H. Meltzer (eds.) *Carnegie-Rochester Conferences on Public Policy*, Volume 1, *Journal of Monetary Economics*, 19–46.

Mizon, G.E. and J.-F. Richard (1986) "The Encompassing Principle and Its Application to Testing Non-nested Hypotheses", *Econometrica*, 54, 3, 657–678.

Nickell, S. (1985) "Error Correction, Partial Adjustment and All That: An Expository Note", *Oxford Bulletin of Economics and Statistics*, 47, 2, 119–129.

Pagan, A. (1987) "Three Econometric Methodologies: A Critical Appraisal", *Journal of Economic Surveys*, 1, 1, 3–24.

Popper, K.R. (1959) *The Logic of Scientific Discovery*, London, Hutchinson.

Shimony, A. (1988) "The Reality of the Quantum World", *Scientific American*, 258, 1, 46–53.

White, H. (1980) "Using Least Squares to Approximate Unknown Regression Functions", *International Economic Review*, 21, 1, 149–170.

White, H. (1988) *Specification Analysis in Econometrics*, Cambridge, Cambridge University Press, forthcoming.

International Finance Discussion Papers

IFDP NUMBER	TITLES 1989	AUTHOR(s)
354	Encompassing and Rational Expectations: How Sequential Corroboration Can Imply Refutation	Neil R. Ericsson David F. Hendry
353	The United States as a Heavily Indebted Country	David H. Howard
352	External Debt and Developing Country Growth	Steven B. Kamin Robert B. Kahn Ross Levine
351	An Algorithm to Solve Dynamic Models	Wilbur John Coleman II
350	Implications of the U.S. Current Account Deficit	David H. Howard
349	Financial Integration in the European Community	Sydney J. Key
348	Exact and Approximate Multi-Period Mean-Square Forecast Errors for Dynamic Econometric Models	Neil R. Ericsson Jaime R. Marquez
347	Macroeconomic Policies, Competitiveness, and U.S. External Adjustment	Peter Hooper
346	Exchange Rates and U.S. External Adjustment in the Short Run and the Long Run	Peter Hooper
345	U.S. External Adjustment: Progress and Prospects	William L. Helkie Peter Hooper
344	Domestic and Cross-Border Consequences of U.S. Macroeconomic Policies	Ralph C. Bryant John Helliwell Peter Hooper
343	The Profitability of U.S. Intervention	Michael P. Leahy
342	Approaches to Managing External Equilibria: Where We Are, Where We Might Be Headed, and How We Might Get There	Edwin M. Truman
341	A Note on "Transfers"	David B. Gordon Ross Levine
340	A New Interpretation of the Coordination Problem and its Empirical Significance	Matthew B. Canzoneri Hali J. Edison

---

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)
	1988	
339	A Long-Run View of the European Monetary System	Hali J. Edison Eric Fisher
338	The Forward Exchange Rate Bias: A New Explanation	Ross Levine
337	Adequacy of International Transactions and Position Data for Policy Coordination	Lois Stekler
336	Nominal Interest Rate Pegging Under Alternative Expectations Hypotheses	Joseph E. Gagnon Dale W. Henderson
335	The Dynamics of Uncertainty or The Uncertainty of Dynamics: Stochastic J-Curves	Jaime Marquez
334	Devaluation, Exchange Controls, and Black Markets for Foreign Exchange in Developing Countries	Steven B. Kamin
333	International Banking Facilities	Sydney J. Key Henry S. Terrell
332	Panic, Liquidity and the Lender of Last Resort: A Strategic Analysis	R. Glen Donaldson
331	Real Interest Rates During the Disinflation Process in Developing Countries	Steven B. Kamin David F. Spigelman
330	International Comparisons of Labor Costs in Manufacturing	Peter Hooper Kathryn A. Larin
329	Interactions Between Domestic and Foreign Investment	Guy V.G. Stevens Robert E. Lipsey
328	The Timing of Consumer Arrivals in Edgeworth's Duopoly Model	Marc Dudey
327	Competition by Choice	Marc Dudey