ENCOMPASSING AND RATIONAL EXPECTATIONS:
HOW SEQUENTIAL CORROBORATION CAN IMPLY REFUTATION

Neil R. Ericsson and David F. Hendry

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.
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. 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. 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). Despite their importance, issues of statistical inference from finite

---

1The 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.

2We 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.

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,...,n\}, all interpreted as sets. The relationship between theories and evidence is of interest, so let the operator \(\varepsilon_c\) denote "is consistent with" or equivalently "is corroborated by". Thus, \(A \varepsilon_c W_i\) ("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,...,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:

\[
\begin{align*}
A & \varepsilon_c W_i \quad (i=1,...,n) , \\
B & \varepsilon_c W_i \quad (i=1,...,n) , \\
& \text{but} \\
A & \varepsilon_c W^* \\
B & \not\varepsilon_c W^* .
\end{align*}
\]

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 \(\varepsilon_c\) 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. 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 $\alpha$ and $\beta$ such that $A = \{(\alpha, \beta) : \alpha + \beta = 1\}$ and $B = \{(\alpha, \beta) : \alpha + \beta = -1\}$, and the pieces of evidence on $\alpha$ and $\beta$ are $W_\alpha = \{\alpha : \alpha > 0\}$ and $W_\beta = \{\beta : \beta > 0\}$. Observing either $W_\alpha$ or $W_\beta$ is corroborating evidence for the theory A, and also for theory B. However, observing both $W_\alpha$ and $W_\beta$ is inconsistent with theory B (for which at least one of $\alpha$ and $\beta$ must be negative) but corroborates theory A. Figure 2 portrays the relationship between A, B, $W_\alpha$, and $W_\beta$; theory B does not intersect the joint evidence $W^* = W_\alpha \cap W_\beta$ but theory A does. The evidence $W_\alpha$ and $W_\beta$ 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., $W_{\beta^*} = \{\beta : \beta > 2\}$).

---

4Keynes (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.

5Popper (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 \( W_\alpha, W_\beta, W^* \), 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. 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:

\[ y_t = \alpha_{1t} + \beta(x_{2t} \cdot x_{3t}) + u_t \]  
\[ u_t \sim \text{NID}(0, \sigma^2) \],  \hspace{1cm} (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):

\[ y_t = \gamma x_{1t} + v_t \]  
\[ v_t \sim \text{NID}(0, \omega^2) \],  \hspace{1cm} (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 \( y_t \):

\[ \text{E}(y_t | [x_{1t} \cdot x_{it}]) = \alpha x_{1t} \]  
i = 2, 3 , \hspace{1cm} (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:

---

6This 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.
\[ E(y_t | x_{1t}, x_{2t}, x_{3t}) = \alpha x_{1t} + \beta (x_{2t}, x_{3t}) \].

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 \( x_t \). Unless otherwise indicated, bold characters are vectors (lower case) or matrices (upper case); lower case non-bold characters are scalars.

\[
\begin{align*}
y_t &= \beta' x_t + e_t \\
x_t &= \Pi z_t + v_t
\end{align*}
\]

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

\[
\delta = \text{plim } \beta_{IV} = \beta ,
\]

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

\[
\gamma = \text{plim } \beta_{LS} = \beta + \left[ \mathbf{E}(\mathbf{x}_t \mathbf{x}_t')^{-1} \mathbf{E}(\mathbf{x}_t \epsilon_t) \right] = \beta + \left[ \Pi \mathbf{M}_{zz} \Pi' + \Sigma_{vv} \right]^{-1} \mathbf{e}_v ,
\]

where \( \mathbf{M}_{zz} = \mathbf{E}(\mathbf{z}_t \mathbf{z}_t') \), \( \Sigma_{vv} = \mathbf{E}(\mathbf{v}_t \mathbf{v}_t') \), and \( \Sigma_{ve} = \mathbf{E}(\mathbf{v}_t \epsilon_t) \); see Bronfenbrenner (1953). Hausman's test statistic compares \( \beta_{IV} \) and \( \beta_{LS} \) 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 \( \Pi \) and \( \Sigma_{vv} \) may well be non-constant over time, being (potentially) complicated functions of economic, policy, and institutional parameters. To the extent that \( \Pi \) and/or \( \Sigma_{vv} \) change, \( \gamma \) 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:** \( \gamma \) is constant.
- **W_2:** \( \delta \) is constant.
- **W_3:** \( \Pi \) and/or \( \Sigma_{vv} \) are non-constant.

The theories of interest are:

- **A:** \( \mathbf{y}_t \) and \( \mathbf{x}_t \) are not simultaneously determined (\( \Sigma_{ve} = 0 \)), and
- **B:** \( \mathbf{y}_t \) and \( \mathbf{x}_t \) are simultaneously determined (\( \Sigma_{ve} \neq 0 \)).

For both theories, \( \beta \) is the parameter of interest and is assumed constant throughout.

The \( W_1 \) 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 \( \beta \), \( W_3 \) does not violate any assumptions of B, and \( W_1 \) could occur if (e.g.) \( (\mathbf{y}_t, \mathbf{x}_t, \mathbf{z}_t') \) were jointly stationary. In the last instance,
Evidence

\[ W_1 = \{ \gamma \text{ is constant.} \} \]

\[ W_2 = \{ \delta \text{ is constant.} \} \]

\[ W_3 = \{ \Pi \text{ and/or } \Sigma_{vv} \text{ are non-constant.} \} \]

\[ (W_2 \cap W_3) \setminus W_1 \quad (\text{The Lucas critique is confirmed.}) \]

\[ W_1 \cap W_2 \cap W_3 \quad (\text{The Lucas critique is refuted.}) \]

Theories

\[ A = \{ y_t \text{ and } x_t \text{ are not simultaneously determined (} \Sigma_{ve}=0.\} \} \]

\[ B = \{ y_t \text{ and } x_t \text{ are simultaneously determined (} \Sigma_{ve}\neq0). \} \]

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 \(\Pi\) and \(\Sigma_{vv}\) 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 \(x_t\) 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 \(x_t\) is a univariate AR(1) process:

\[
x_t = \pi x_{t-1} + v_t \quad \quad \quad \quad \quad v_t \sim \text{NI}(0, \sigma_{vv}) \quad (4.2a)
\]

where \(0 < |\pi| < 1\). If \(z'_t = x_{t-1}\), (4.4) becomes:

\[
\gamma = \beta + (1-\pi^2) \cdot (\sigma_{ve}/\sigma_{vv}) \quad (4.4a)
\]

If \(\sigma_{ve} \neq 0\) and \(\pi\) and/or \(\sigma_{vv}\) change, then \(\gamma\) changes and so \(W_1\) would not be the case. Conversely, if \(\gamma\) remains constant in spite of \(\pi\) and/or \(\sigma_{vv}\) 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 \(\Pi\) and \(\Sigma_{vv}\) 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) $x_t$ 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 $x_t$ and about future values of $y_t$ itself. Expectations are formed, conditional upon an information set (denoted $w_{t-1}$) which typically includes lagged values of $x_t$ and $y_t$.

$$\sum_{i=0}^{\infty} \kappa_i E(y_{t+i} | w_{t-1}) = \sum_{i=0}^{\infty} \lambda'_i E(x_{t+i} | w_{t-1})$$

(5.1)

$k_0=1$, the $\lambda_i$'s and the remaining $\kappa_i$'s are the "deep" parameters of the agents' behavior (or direct functions of them), and typically $E(y_t | w_{t-1})=y_t$. 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 $x_{t+i}$'s alone:

$$E(y_t | w_{t-1}) = \sum_{i=0}^{\infty} \delta'_i E(x_{t+i} | w_{t-1})$$

(5.2)

where $\delta_i$ 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 $x_t$, which is:

$$E(x_t | w_{t-1}) = \Pi x_{t-1}$$

(5.3)

noting the strict exogeneity of $x_t$ and again assuming linearity.\(^8^8\) Equations (5.2) and (5.3) can be written in "model form" as:

$$y_t = \sum_{i=0}^{\infty} \delta'_i E(x_{t+i} | w_{t-1}) + \epsilon_t$$

$$x_t = \Pi x_{t-1} + v_t$$

(5.4)

(5.5)

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

\(^8^8\)If $x_t$ 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 $x_t$. 

we allow possible discrepancies through non-zero $\epsilon_t$. For expository convenience, we assume $(\epsilon_t, v_t)$ is independently and identically distributed, normal.

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

$$y_t = \gamma' x_t + \nu_t \quad \text{E}(\nu_t \cdot x_t) = 0,$$

(5.6)

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

$$y_t = (\sum_{i=0}^{\infty} \delta_i' \Pi^i) \cdot x_t + \left[ \epsilon_t - (\sum_{i=0}^{\infty} \delta_i' \Pi^i) \cdot v_t \right],$$

(5.7)

assuming that $x_t$ is stationary. With (5.5), this representation parallels (4.1)–(4.2), in which $\beta' = \Sigma_0' \delta_1 \Pi^i, \epsilon_t = [\epsilon_t - (\sum_0^{\infty} \delta_i' \Pi^i) \cdot v_t]$, and $z_t = x_{t-1}$. Thus, (4.4) is appropriate for calculating the value to which the least-squares estimator of the coefficient on $x_t$ in (5.6) converges.

$$\gamma = (\sum_{i=0}^{\infty} \{\Pi^i\}' \delta_i) + [M_{xx}]^{-1} \cdot [\Sigma_{ve} - \Sigma_{vv} (\sum_{i=0}^{\infty} \{\Pi^i\}' \delta_i)]$$

(5.8)

$M_{xx}^\theta \equiv E(x_t x_t')^\theta = (I - \Pi \otimes \Pi)^{-1} \Sigma_{vv}^\theta$, $\theta$ is the column vectoring operator, and $\otimes$ is the corresponding Kronecker product. The coefficient $\gamma$ is constant if $\Pi, \Sigma_{vv}, \Sigma_{ve}$, and the $\delta_i$'s are. However, as changes occur in the process generating $x_t$ (e.g., $\Pi$ varies over time), $\gamma$ also will change and the conditional model (5.6) will "break down". The Lucas critique

---

9If (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 $\gamma$ in (5.8), but with an additional term introduced by the approximation error; cf. White (1980). Likewise, the conditional expectation for $x_t$ in (5.3) might be nonlinear, in which case (5.5) with $\Pi$ would be the least-squares approximation to that nonlinear function. Such nonlinearities per se could not induce non-constancy in either $\gamma$ or $\Pi$ although, in finite samples, apparent non-constancy might be detected if (e.g.) $x_t$ were very slowly changing (e.g., very autoregressive). Additional data would clarify that there was constancy.
applies, and (5.6) fails to isolate the $\delta_i$'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 $\gamma$ is constant in spite of changes in the process for $x_t, y_t$ 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 $\gamma$ is constant and that $\Pi$ and/or $\Sigma_{\gamma\gamma}$ have changed.

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

- **W1**: $\gamma$ is constant.
- **W2**: $\{\lambda_i\}$ and $\{\kappa_i\}$ (and so $\{\delta_i\}$) are constant.
- **W3**: $\Pi$ and/or $\Sigma_{\gamma\gamma}$ are non-constant.

The theories of interest are:

- **A**: $y_t$ is determined conditional upon $x_t$, and
- **B**: $y_t$ is determined via expectations of future $y_t$ and of current and future $x_t$.

Unlike Section 4, the parameters of interest for the two hypotheses generally are not the same, being $\gamma$ 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 $x_t$ (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 $\{W_1\}$, 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 $W_2$.\(^\text{10}\)

The following examples

\(^{10}\)For example, suppose the conditional model is correct (with constant $\gamma$ ($W_1$)) and $x_t$ is univariate AR(1):
\[
y_t = \gamma x_t + \nu_t \\
x_t = \pi x_{t-1} + \nu_t
\]
but that an expectations model is estimated with the expectations of a single future $x_t$:
\[
y_t = \delta_t E(x_{t+r}|w_{t-1}) + \epsilon_t
\]
where $w_{t-1} = x_{t-1}$. In that case, $E(x_{t+r}|w_{t-1}) = \pi^r(x_t - \nu_t)$, so substitution gives:
\[
y_t = (\gamma \pi^r) E(x_{t+r}|w_{t-1}) + (\nu_t + \gamma \nu_{t})
\]
Thus, $\delta_t = \gamma \pi^r$ because $E([\nu_t + \gamma \nu_{t}]|w_{t-1}) = 0$, so $\delta_t$ changes as $\pi$ changes. Further, if $\pi$ changes ($W_3$) and $\delta_t$ 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 \( \delta_i \) are non-zero and when \( x_t \) is univariate AR(1).

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

\[
y_t = \delta_0' E(x_t | w_{t-1}) + \epsilon_t \quad \text{E}(\epsilon_t \cdot w_{t-1}) = 0. \quad (5.4a)
\]

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

\[
y_t = \delta_0' x_t + (\epsilon_t - \delta_0' v_t). \quad (5.7a)
\]

Thus, the expectations model defined by (5.4a) and (5.5) is equivalent to the simultaneous equations model (4.1)–(4.2), with \( \epsilon_t = (\epsilon_t - \delta_0' v_t) \), \( z_t = x_{t-1} \), and \( \beta = \delta_0 \). From (4.4), simultaneity bias arises in estimating \( \delta_0 \) in (5.4a) by using realized values of \( x_t \) rather than its expectation.

\[
\gamma = \delta_0 + [M_{xx}]^{-1} (\Sigma_{ve} - \Sigma_{vv} \delta_0), \quad (5.8a)
\]

which is not \( \delta_0 \) even if \( v_t \) and \( \epsilon_t \) are uncorrelated. See Hendry (1988) for details. With \( \sigma_{ve} = 0 \) and \( x_t \) a stationary univariate AR(1) process, (5.8a) simplifies:

\[
\gamma = \pi^2 \delta_0. \quad (5.8a')
\]

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

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

\[
y_t = \delta_r' E(x_{t+r} | w_{t-1}) + \epsilon_t \quad \text{E}(\epsilon_t \cdot w_{t-1}) = 0, \quad (5.4b)
\]

and so:

\[
y_t = \delta_r' \Pi^r x_t + (\epsilon_t - \delta_r' \Pi^r v_t). \quad (5.7b)
\]

This again parallels (4.1), so the coefficient defined by \( E(y_t | x_t) \) is:

\[
\gamma = (\Pi')^r \delta_r + [M_{xx}]^{-1} \{ \Sigma_{ve} - \Sigma_{vv} (\Pi')^r \delta_r \} \quad (5.8b')
\]

With \( \sigma_{ve} = 0 \) and a stationary univariate AR(1) process for \( x_t \):

\[
\gamma = \pi^{r+2} \delta_r \quad (5.8b')
\]

Figure 4 portrays Example 4 in general, and 4b for \( r > 0 \). Table 1 summarizes the formulae for \( \gamma \) as a function of the (simultaneous or expectations-based) process generating the data.
\[ \begin{align*}
W_1 & = \{ \gamma \text{ is constant.} \} \\
W_2 & = \{ \{ \lambda_i \} \text{ and } \{ \kappa_i \} \text{ (and so } \{ \delta_i \} \text{) are constant.} \} \\
W_3 & = \{ \Pi \text{ and/or } \Sigma_{\tau \tau} \text{ are non-constant.} \} \\
\square & \quad (W_2 \cap W_3) \setminus W_1 \quad (\text{The Lucas critique is confirmed.}) \\
\blacksquare & \quad W_1 \cap W_3 \quad (\text{The Lucas critique is refuted.})
\end{align*} \]

**Theories**

\[ \begin{align*}
A & = \{ y_t \text{ is determined conditional upon } x_{t'} \} \\
B & = \{ y_t \text{ is determined via expectations of future } y_t \text{ and of current and future } x_{t'} \}
\end{align*} \]

**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 $\gamma (\equiv \plimm \beta_{LS})$ as a function of the data generation process.

<table>
<thead>
<tr>
<th>Example</th>
<th>The process generating $x_t$</th>
<th>vector autoregressive</th>
<th>univariate AR(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$\beta + \left[ \Pi \Sigma_{xx} \Pi' + \Sigma_{vv} \right]^{-1} \Sigma_{ve}$</td>
<td>(\beta + (1-\pi^2)\cdot(\sigma_{ve}/\sigma_{vv}))</td>
<td></td>
</tr>
<tr>
<td>Expectations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$\left( \sum_{i=0}^{\infty} { \Pi' }^i \delta_i \right) + \left[ \Sigma_{ve} - \Sigma_{vv} \right] \left( \sum_{i=0}^{\infty} { \Pi' }^i \delta_i \right)$</td>
<td>$\pi^2 \sum_{i=0}^{\infty} \pi^i \delta_i$</td>
<td></td>
</tr>
<tr>
<td>4a b</td>
<td>$\delta_0 + \left[ \Sigma_{xx} \right]^{-1} \left( \Sigma_{ve} - \Sigma_{vv} \delta_0 \right)$</td>
<td>$\pi^2 \delta_0$</td>
<td></td>
</tr>
<tr>
<td>4b c</td>
<td>$(\Pi')^i \delta_i + \left[ \Sigma_{xx} \right]^{-1} \left[ \Sigma_{ve} - \Sigma_{vv} (\Pi')^i \delta_i \right]$</td>
<td>$\pi^r \delta_r$</td>
<td></td>
</tr>
</tbody>
</table>

Notes:

a. The second moment matrix for $x_t$ (denoted $\Sigma_{xx}$) can be expressed as $\left[ \Pi \Sigma_{xx} \Pi' + \Sigma_{vv} \right]$, and is equal to $\sigma_{vv}/(1-\pi^2)$ when $x_t$ is univariate AR(1). For the expectations examples with $x_t$ being univariate AR(1), $\sigma_{ve}$ is set to zero.

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

c. For Example 4b, $\delta_i \neq 0$ 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 $\pi$ and $\delta_0$ in Example 4a, a prediction of $\gamma$ could be constructed from (5.8a') and compared with the observed estimate of $\gamma$ from (5.6). Although intuitively appealing and well-founded theoretically, such complete encompassing can prove exceedingly difficult when $x_t$ is determined by a complicated (and unknown) process. Even so, fully efficient estimation of the expectations model requires properly specifying that process for $x_t$. 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 $x_t$ in (4.2) need not be specified: identifying a subset of the $z_t$ 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 $z_t^*$ is required in (4.2) to make it complete:

$$x_t = \Psi_1 z_t + \Psi_2 z_t^* + v_t^* ,$$

(5.9)
then the least-squares estimator of $\Pi$ in (4.2) is subject to omitted variables bias. By assumption, $\Pi$ changes ($W_3$). From the bias formula, it could have done so for one (or more) of three reasons: the underlying coefficient $\Psi_1$ on $z_t$ changed; the coefficient $\Psi_2$ on $z_t^*$ changed; or the correlation ($\Phi$, say) between $z_t^*$ and $z_t$ changed. Because $\gamma$ is a function of $\Pi$, which in turn is a function of $\Psi_1$, $\Psi_2$, and $\Phi$, then $\gamma$ will change as any of $\Psi_1$, $\Psi_2$, and $\Phi$ 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. 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)

---

11This 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 inexactlly, 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

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


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

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