

# ENCOMPASSING AND RATIONAL EXPECTATIONS: HOW SEQUENTIAL CORROBORATION CAN IMPLY REFUTATION

Neil R. Ericsson and David F. Hendry  
Revised: July 15, 1997

## ABSTRACT

Even though pieces of empirical evidence individually may corroborate an economic theory, their joint existence may refute that same theory. Testing of rational expectations models provides a concrete illustration of this principle. Surprisingly, empirical refutation of a rational expectations model may occur without having to estimate that model, and the refutation may be for a large class of expectations-based models and not just for a particular model specification. Narrow money demand in the United Kingdom illustrates such refutation. The general proposition concerning corroboration and refutation strongly favors the building of empirical models that are consistent with all available evidence.

*Keywords and phrases:* conditional models, congruence, corroboration, encompassing, feedback, feedforward, Lucas critique, rational expectations, refutation, statistical inference.

*JEL classifications:* C52, E13.

All correspondence should be addressed to the first author.

Neil R. Ericsson  
Stop 24, International Finance Division  
Federal Reserve Board  
2000 C Street, N.W.  
Washington, D.C. 20551 U.S.A.  
1 (202) 452-3709 (office)  
1 (202) 736-5638 (fax)  
ericsson@frb.gov (Internet)

David F. Hendry  
Nuffield College  
Oxford OX1 1NF  
England  
44 (1865) 278587 (office)  
44 (1865) 278557 (fax)  
david.hendry@nuffield.oxford.ac.uk (Internet)

# ENCOMPASSING AND RATIONAL EXPECTATIONS: HOW SEQUENTIAL CORROBORATION CAN IMPLY REFUTATION

Neil R. Ericsson and David F. Hendry\*

## 1 INTRODUCTION

Economic theories are rarely 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 a given empirical implementation thereof. Examples include rejection of an over-identifying restriction or a coefficient estimate having the wrong sign. Implications from a set of evidence can be subtler. Several pieces of empirical evidence may exist, each of which corroborates a theory, but jointly those very pieces of evidence may refute that same theory. While surprising at first sight, this proposition is more obvious upon closer examination. It has substantive implications for econometric modeling and, in particular, for testing hypotheses about the empirical role of expectations.

Section 2 illustrates the proposition and states the applicable theorem. Section 3 applies that theorem to evidence on model constancy when testing expectations-based models. Models with expectations are of particular interest to economists: they have an intuitive appeal, they are ubiquitous, in fair part because of the Lucas (1976) critique, and yet the empirical evidence in favor of the Lucas critique is minimal; see Ericsson and Irons (1995). Surprisingly, empirical refutation of an expectations-based model may occur without having to estimate that model, and refutation may be for a large class of expectations-based models and not just for a particular model specification. While Section 3 focuses on expectations-based models, the theorem applies to evidence from a wide range of tests, including those of omitted variables and simultaneity; see Ericsson and Hendry (1989). Section 4 discusses how the theorem and its application favor accounting for a wide variety of evidence (via the encompassing principle) and accounting for the evidence as a whole (congruency). Evaluation of a theory in light of

---

\*The first author is a staff economist in the Division of International Finance, Federal Reserve Board, Washington, D.C., U.S.A. The second author is Leverhulme Personal Research Professor of Economics at Nuffield College, Oxford, England, and was a Visiting Research Professor at the Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina when this research was begun. The authors may be reached on the Internet at [ericsson@frb.gov](mailto:ericsson@frb.gov) and [david.hendry@nuffield.oxford.ac.uk](mailto:david.hendry@nuffield.oxford.ac.uk) respectively. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting those of the Board of Governors of the Federal Reserve System, Norges Bank, or other members of their staff. The authors gratefully acknowledge the generous hospitality of Norges Bank (where they revised some of the material herein) and financial support (to DFH) from the U.K. Economic and Social Research Council under grants L116251015 and R000234954. We wish to thank Lisa Barrow and Neva Kerbeshian for research assistance; and Julia Campos, Frank Diebold, Eric Leeper, Jaime Marquez, Doug McManus, Adrian Neale, Baldev Raj, Peter Urbach, John Worrall, and two anonymous referees for helpful comments. All numerical results were obtained using PcGive Professional Version 9.00; see Doornik and Hendry (1996).

individual pieces of evidence is an important aspect of empirical modeling. However, the full set of evidence can have implications additional to the inferences from the individual pieces of evidence: the whole can be greater than the sum of the parts. Section 5 shows empirically how the theorem resolves a debate on the interpretation of an error correction model for narrow money demand in the United Kingdom.

Four notes are germane. First, the proposition of corroboration leading to refutation may not be original, and in fact it seems known in the natural sciences.<sup>1</sup> Our interest is in its application to empirical economic modeling and so in its implications for econometric methodology. Second, others such as Hendry (1988), Favero and Hendry (1992), Engle and Hendry (1993), and Hendry (1995, Chapter 14) also discuss how evidence on model constancy can be used to test expectations-based models. The current paper complements those papers in that it focuses on the sequential accumulation of economic evidence and on the contrasting inferences arising from the individual pieces of evidence and from the evidence taken jointly. The cited papers neglect the sequential and corroborating aspects of the evidence involved. The sequential aspect is particularly important in empirical economics: economic evidence typically accrues over time and by different researchers, and pre-existing empirical results are often ignored in new research. Even different pieces of evidence within a given paper are not always considered jointly, as the second example in Section 2.1 shows. Third, issues of statistical inference from a *finite* sample are ignored. While important, such issues are set aside below in order to facilitate focusing on the logical rather than the statistical implications of evidence, given the evidence itself. Put somewhat differently, the results below hold even in large samples, so they apply to small samples as well. Finally, for ease of exposition, both theories and evidence on theories are viewed as restrictions on (and hence sets in) some observation space.

## 2 TWO ILLUSTRATIONS AND THE THEOREM

Section 2.1 motivates the theorem with a commonplace event and with the Marshall-Lerner condition from international trade. Section 2.2 presents the theorem itself. While the theorem and its proof are simple, their formalization clarifies the implications of the theorem and the assumptions necessary for its application (Section 2.3).

### 2.1 TWO ILLUSTRATIONS

Two examples demonstrate how sequential corroboration can imply refutation. The first example is non-economic but conveys the essence of the problem. The second example concerns evidence on the Marshall-Lerner condition.

---

<sup>1</sup>Popper (1959, Appendix \*IX, pp. 397–398) 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, the observations refute *every* local hidden-variables model, but corroborate quantum theory.

Consider three friends,  $X$ ,  $Y$ , and  $Z$ , who have checked their coats at a restaurant. As they are leaving, they go to retrieve their coats. The manager apologizes — their coat checks have been lost — and hands out the three coats randomly to  $X$ ,  $Y$ , and  $Z$ . The three friends decide to test the hypothesis that each one of them has someone else’s coat: that is,  $X$  has  $Y$  or  $Z$ ’s coat,  $Y$  has  $X$  or  $Z$ ’s coat, and  $Z$  has  $X$  or  $Y$ ’s coat — a complete mix-up of coats.  $X$  looks at the coat that she’s been given: it is  $Y$ ’s coat. This piece of evidence corroborates the “complete mix-up” hypothesis. Furthermore, if  $X$  had found that she had been given her own coat, that evidence would have refuted the complete mix-up hypothesis.  $Y$  now looks at the coat that he’s received: it is  $X$ ’s coat. By itself,  $Y$ ’s evidence also corroborates the complete mix-up hypothesis, and it could have refuted that hypothesis as well. Individually,  $X$ ’s evidence and  $Y$ ’s evidence corroborate the complete mix-up hypothesis. However, together,  $X$ ’s and  $Y$ ’s evidence imply that  $Z$  has  $Z$ ’s own coat, thereby refuting the complete mix-up hypothesis. Sequential corroboration can imply refutation.

Empirical evidence on the Marshall-Lerner condition also illustrates this point. Briefly, the Marshall-Lerner condition posits that the price elasticities for export and import demand equations of a given country are negative and that the sum of those elasticities is less than  $-1$ . That is, for export and import price elasticities  $a$  and  $b$ , the Marshall-Lerner condition requires that  $a < 0$ ,  $b < 0$ , and  $a + b < -1$ . If the Marshall-Lerner condition holds, then an increase in the price of imports relative to the price of exports will improve the trade balance (exports minus imports); see Krugman and Obstfeld (1994, pp. 476–478). Whether or not the Marshall-Lerner condition is satisfied has direct consequences for economic policy; again, see Krugman and Obstfeld (1994).

Table 1 lists Taylor’s (1989, Tables 2D–2 and 2D–3) estimated price elasticities for two countries, Italy and the United Kingdom. For Italy, the elasticities are  $a = -0.27$  and  $b = -0.32$ . Each elasticity by itself corroborates the Marshall-Lerner condition. Jointly, however, the two elasticities violate the Marshall-Lerner condition: the sum of elasticities  $a + b$  is  $-0.59$ , which is greater than  $-1$ . For the United Kingdom, a similar situation arises. The individual elasticities are negative ( $a = -0.43$  and  $b = -0.18$ ), but their sum ( $-0.61$ ) is greater than  $-1$ . For each country, sequential corroboration of the Marshall-Lerner condition by individual elasticities leads to refutation by the pair of elasticities.

The results have a straightforward graphical interpretation: Figure 1 plots the Marshall-Lerner condition and the Italian evidence thereon. Each line corresponding to a piece of evidence intersects the shaded area  $A$ , which represents values satisfying the Marshall-Lerner condition. However the intersection of the two lines at  $(a, b) = (-0.27, -0.32) \equiv W^*$  lies outside the shaded area.

This subsection sets up some notation and presents the general theorem. Then, the examples from Section 2.1 are re-expressed in terms of the theorem's notation.

Consider a theory, denoted  $A$ ,  $n$  pieces of evidence  $\{W_i, i = 1, \dots, n\}$ , and the intersection of the evidence  $W^* (\equiv \cap_{i=1}^n W_i)$ , all interpreted as sets. The relationship between the theory and the evidence is of interest, so let the operator  $\in_c$  denote “is consistent with” or, equivalently, “is corroborated by”. The statement  $A \in_c W_i$  (“theory  $A$  is corroborated by the evidence  $W_i$ ”) has two interpretations. As a statement in logic,  $A \in_c W_i$  means that a subset of theory  $A$  lies in the evidence set  $W_i$ . As a mathematical statement,  $A \in_c W_i$  means that the intersection of the sets  $A$  and  $W_i$  is not empty.

*Theorem.* Suppose that, for the  $n$  sets  $\{W_i, i = 1, \dots, n\}$ ,

- (a) their intersection  $W^*$  is not the empty set, and further that
- (b)  $W^*$  is a proper subset of each of the  $W_i$ .

Then, there exists a non-empty set  $A$  such that:

- (i)  $A \in_c W_i$  ( $i = 1, \dots, n$ ), and
- (ii)  $A \notin_c W^*$ .

Additionally, for the same set of evidence  $\{W_i, i = 1, \dots, n\}$ , there exists another non-empty set  $B$  such that:

- (iii)  $B \in_c W_i$  ( $i = 1, \dots, n$ ), and
- (iv)  $B \in_c W^*$ .

The proof appears in Appendix A.

Assumptions (a) and (b) are the conditions necessary for the theorem to hold. Under those assumptions, corroboration of a theory [(i)] can imply its refutation [(ii)], even though another theory (here, denoted  $B$ ) can be consistent with all the evidence jointly.

The examples in Section 2.1 relate to the theorem as follows. In the coat example, the two pieces of evidence are:

$$W_1 : \{X \text{ has } Y\text{'s coat}\}$$

$$W_2 : \{Y \text{ has } X\text{'s coat}\},$$

so their intersection  $W^* (= W_1 \cap W_2)$  is:

$$W^* : \{X \text{ has } Y\text{'s coat, } Y \text{ has } X\text{'s coat, and so } Z \text{ has } Z\text{'s coat}\}.$$

Theory  $A$  is the complete mix-up hypothesis:

$$A : \{X \text{ has } Y \text{ or } Z\text{'s coat, } Y \text{ has } X \text{ or } Z\text{'s coat, and } Z \text{ has } X \text{ or } Y\text{'s coat}\}.$$

Theory  $B$  is unspecified, but could be:

$$B : \{\text{one and only one person has his or her own coat}\},$$

which is consistent with  $W^*$ . Because just three people are involved,  $W^*$  actually confirms theory  $B$ .

For the Marshall-Lerner condition, the evidence is on the Italian export and import

price elasticities  $a$  and  $b$ :

$$W_1 : \{a = -0.27\}$$

$$W_2 : \{b = -0.32\},$$

and their intersection is:

$$W^* : \{a = -0.27, b = -0.32\}.$$

Theory  $A$  is the Marshall-Lerner condition:

$$A : \{a < 0, b < 0, a + b < -1\}.$$

Theory  $B$  is unspecified, but could be (for instance):

$$B : \{a < 0, b < 0, a + b > -1\},$$

which is consistent with  $W^*$ . In each of the two examples, only one “theory” is of interest. In discussing rational expectations below, substantive theories correspond to both  $A$  and  $B$ .

### 2.3 REMARKS ON THE THEOREM

Several remarks elucidate the nature of the theorem. First, the assumptions serve to exclude irrelevant situations. Assumption (a) entails that theories exist that are not empty and that are consistent with all the evidence. Assumption (b) means that, for any piece of evidence  $W_i$ , at least one of the remaining pieces of evidence  $W_j$  ( $j \neq i$ ) 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 two or more distinct pieces of evidence, the theorem would not apply.

Second, sequential corroboration is often nontrivial in that a single piece of evidence may refute the theory being investigated. In the coat example,  $X$  could have discovered that she had her own coat; likewise, for  $Y$ . In examining the Marshall-Lerner condition, a positive estimate for either elasticity could have been obtained, and positivity by itself would have violated the Marshall-Lerner condition. That said, the definition of corroboration in Section 2.2 is weaker than the one usually used in the philosophy of science, where corroboration necessarily implies the potential for refutation. The theorem holds for either definition of corroboration.

Third, each individual piece of evidence  $W_i$  corroborates both theories  $A$  and  $B$ . However, observing *all* the evidence  $\{W_i\}$  reduces the range of feasible theories, and that range excludes  $A$  while still including  $B$ . Thus, the joint observation of  $\{W_i\}$  refutes theory  $A$  while corroborating theory  $B$ . Different inferences about the theories arise, depending upon whether the evidence is considered piecemeal or as a whole.<sup>2</sup>

Fourth, the theorem is an existence theorem, rather than a theorem of inevitability.

---

<sup>2</sup>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, Chapter 1). The theorem illustrates why such a theory is not feasible.

The theorem states that there exists some pair of theories  $A$  and  $B$  that satisfies (i)–(iv). Any specific pair of theories would need to be considered directly in light of the evidence at hand. In particular, any specific theory might be corroborated by each individual piece of evidence and by the evidence as a whole.

### 3 TESTING FEEDBACK VERSUS FEEDFORWARD MODELS

In empirical analysis, contending models often differ in their treatment of expectations. Two common forms are feedforward (or expectations-based) models, such as those based on rational expectations, and feedback (or conditional) models. Empirical properties of such models often can clarify which model, if either, is consistent with the empirical evidence. Specifically, Hendry’s (1988) proposal for testing feedback versus feedforward models is a key illustration of the theorem in Section 2. In that proposal, the empirical constancy of a conditional model and the nonconstancy of an associated marginal model serve to rule out certain expectations-based models as potential explanations of the data. The expectations-based models are consistent with each piece of evidence on constancy, but not with them jointly. Section 2’s theorem helps clarify the interpretation and generality of this testing procedure, which has been subject to some confusion in the literature; cf. Cuthbertson (1988, 1991).

Evidence on model constancy is of more general interest as well. Constancy often implies invariance, which is required for valid counterfactual policy simulations. Rational expectations models are motivated as a means for obtaining parameters invariant to policy changes. And, many models are empirically nonconstant — that observation *inter alia* motivated the Lucas (1976) critique.

Section 3.1 examines evidence on constancy for models of a simple bivariate process and derives implications of that evidence. Section 3.2 discusses ramifications of these results. Section 4 then turns to the encompassing principle and its role in testing these models, and Section 5 provides an empirical example.

#### 3.1 A PAIR OF SIMPLE MODELS

To illustrate Hendry’s (1988) testing procedure, consider two models for a variable  $y$ . In one model,  $y$  depends on another variable  $x$  through expectations of  $x$ . In the other model,  $y$  depends on  $x$  directly. This subsection presents the two contending models, derives relationships between the models, and considers the models’ empirical constancy in light of the theorem in Section 2. The expectations-based model discussed is simple in its treatment of expectations and lags by being bivariate, one-period, and linear. However, these limitations are purely for expositional convenience and are not critical for the application of the theorem.

In the first model,  $y$  depends on the expectation of future  $x$ . Models with expectations, and those with rational expectations in particular, have been popular over the last two decades, in part because of the Lucas (1976) critique. In the expectations-

based model below, economic agents make decisions about a variable  $y$  in light of their expectations about next period's value of some strictly exogenous variable  $x$ , given an information set  $\mathcal{I}$ . Algebraically, this model is:

$$y_t = \delta \mathcal{E}(x_{t+1} | \mathcal{I}_{t-1}) + \varepsilon_t \quad \varepsilon_t \sim NI(0, \Sigma_{\varepsilon\varepsilon}), \quad (1)$$

where  $t$  is a time subscript, the coefficient  $\delta$  is regarded as a “deep structural parameter” of the agents’ behavior or a direct function thereof,  $\mathcal{E}(\cdot | \cdot)$  is the expectations operator,  $\varepsilon_t$  is the agents’ decision error, “ $\sim NI(\cdot, \cdot)$ ” denotes “is normally and independently distributed as”, and  $\Sigma_{\varepsilon\varepsilon}$  is the variance of  $\varepsilon_t$ . The information set  $\mathcal{I}_{t-1}$  often includes lagged  $x$  and lagged  $y$ . For convenience, assume  $\mathcal{I}_{t-1} = (x_{t-1}, y_{t-1})$ . The error  $\varepsilon_t$  is typically assumed to be an innovation with respect to  $\mathcal{I}_{t-1}$ , and so  $\mathcal{E}(\varepsilon_t | \mathcal{I}_{t-1}) = 0$ . Normality of  $\varepsilon_t$  is not a critical assumption.

The second model is a conditional model. 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. These models thus assume that it is valid to condition on  $x_t$  itself and that expectations of future  $x$  are unimportant. Economically, agents’ decisions depend upon observed outcomes of  $x$ , rather than expectations of future (and as yet unobserved) realizations. For expositional convenience,  $y$  is assumed to depend on the current  $x$  alone in the conditional model:

$$y_t = \gamma x_t + e_t \quad e_t \sim NI(0, \Sigma_{ee}), \quad (2)$$

where  $\gamma$  is the parameter of interest, and  $\Sigma_{ee}$  is the variance of the error  $e_t$ . Because (2) is a *conditional* model,  $e_t$  satisfies  $\mathcal{E}(e_t | x_t) = 0$ ,  $\mathcal{E}(e_t x_t | x_t) = 0$ , and hence  $\mathcal{E}(e_t x_t) = 0$ . By substitution from (2), the condition  $\mathcal{E}(e_t x_t) = 0$  implies that:

$$\gamma = \mathcal{E}(y_t x_t) / \mathcal{E}(x_t^2), \quad (3)$$

which is the usual least-squares formula, but using population rather than sample moments. Expressed slightly differently,  $\gamma$  is defined by taking expectations of  $y_t$  conditional on  $x_t$ :  $\mathcal{E}(y_t | x_t) = \gamma x_t$ . That contrasts with  $\delta$ , which is defined by taking expectations of  $y_t$  conditional on  $\mathcal{I}_{t-1}$ :  $\mathcal{E}(y_t | \mathcal{I}_{t-1}) = \delta \mathcal{E}(x_{t+1} | \mathcal{I}_{t-1})$ .

The implications of the constancy of  $\delta$  and  $\gamma$  now follow, first for the expectations-based model and then for the conditional model. To obtain the first set of implications, the expectations-based model is assumed to be correctly specified, in which case the conditional model is a reduced form of that expectations-based model. To appreciate this relation between the models, the process for  $x$  must be described. For instance, suppose that  $x$  is a stationary first-order autoregressive process:

$$x_t = \pi x_{t-1} + u_t \quad u_t \sim NI(0, \Sigma_{uu}), \quad (4)$$



where  $\pi$  is an autoregressive coefficient such that  $0 < |\pi| < 1$ , and  $u_t$  is an error term with mean zero and variance  $\Sigma_{uu}$ . If (4) and the expectations-based model (1) constitute the data generation process for  $(y_t, x_t)$ , then the conditional model (2) can be derived from those two equations, with  $\gamma$  being a function of  $\delta$  and  $\pi$ .

The derivation proceeds as follows. From (4):

$$\mathcal{E}(x_{t+1}|\mathcal{I}_{t-1}) = \pi^2 x_{t-1} = \pi x_t - \pi u_t. \quad (5)$$

Substituting (5) into (1),  $y_t$  can be re-expressed in terms of  $x_t$ :

$$y_t = \delta \pi x_t + (\varepsilon_t - \delta \pi u_t). \quad (6)$$

Thus, from the perspective of the expectations-based model, the conditional expectation of  $y_t$  with respect to  $x_t$  is:

$$\mathcal{E}(y_t|x_t) = \delta \pi x_t + \mathcal{E}(\varepsilon_t|x_t) - \delta \pi \mathcal{E}(u_t|x_t). \quad (7)$$

Equation (7) foreshadows the bias of the least squares estimator of  $\gamma$  for  $\delta$ . Substituting (6) into (3), the coefficient  $\gamma$  in (2) is:

$$\begin{aligned} \gamma &= \delta \pi + (M_{xx})^{-1}(\Sigma_{u\varepsilon} - \delta \pi \Sigma_{uu}) \\ &= \delta \pi^3 + (1 - \pi^2)(\Sigma_{u\varepsilon}/\Sigma_{uu}), \end{aligned} \quad (8)$$

where  $M_{xx}$  is the population second moment of  $x_t$ , and  $\Sigma_{u\varepsilon}$  is the covariance of  $u_t$  and  $\varepsilon_t$ . The second line in (8) follows because  $M_{xx} = \Sigma_{uu}/(1 - \pi^2)$  for a first-order autoregressive  $x$ . In the second line of (8), the first term on the right-hand side ( $\delta \pi^3$ ) convolutes the structural parameter  $\delta$  with the autoregressive coefficient  $\pi$  from the process of the forcing variable  $x$ . The second term,  $(1 - \pi^2)(\Sigma_{u\varepsilon}/\Sigma_{uu})$ , also involves aspects of both (1) and (4). Even if  $\Sigma_{u\varepsilon} = 0$ , the coefficient  $\gamma$  is an amalgam of the structural parameter  $\delta$  and the forcing variable parameter  $\pi$ .

This description of the models and their relationship under the assumption of the expectations hypothesis provides the basis for discussing evidence on empirical constancy. The coefficient  $\gamma$  is constant if  $\delta$ ,  $\pi$ ,  $\Sigma_{u\varepsilon}$ , and  $\Sigma_{uu}$  are constant. However, autoregressions such as (4) often are empirically nonconstant, in that estimates of  $\pi$  and  $\Sigma_{uu}$  over different subsamples are statistically and numerically significantly different from each other. Such nonconstancy is unsurprising in light of financial innovation, deregulation, changes in government, and changes in government policy. If the expectations-based model is correct, then the coefficient  $\gamma$  in the conditional model will vary as  $\pi$  or  $\Sigma_{uu}$  varies. That is, the conditional model (2) will “break down”, the Lucas critique will apply, and (2) will fail to isolate the underlying structural parameter  $\delta$ . By assumption,  $\delta$  in (1) will remain constant in spite of (4) evolving.

Empirical evidence on constancy sometimes differs from this well-known case. Specifically,  $\gamma$  might be constant in spite of changes in the process for  $x$ . In this case,  $y$  could not have been generated by (1) with a *constant* deep parameter  $\delta$ , and the Lucas critique is empirically refuted. That is, because of (8), the expectations-based model (1) is inconsistent with the observations that  $\gamma$  is constant *and* that  $\pi$  and/or  $\Sigma_{uu}$  have changed.<sup>3</sup>

To relate this result to sequential corroboration and the theorem in Section 2, the evidence on the constancy of various parameters is categorized as follows.

$W_1$ :  $\delta$  is constant.

$W_2$ :  $\gamma$  is constant.

$W_3$ :  $\pi$  and  $\Sigma_{uu}$  are nonconstant.

The theory  $A$  is the expectations-based model (1):

$A$ :  $y_t$  is determined via expectations of  $x_{t+1}$  with constant parameterization  $\delta$ .

Individually, the pieces of evidence ( $W_1$ ,  $W_2$ , and  $W_3$ ) corroborate theory  $A$ . The evidence  $W_1$  is a necessary condition for the theory,  $W_3$  does not violate any assumptions of that theory, and  $W_2$  could occur if  $(y_t, x_t)$  were jointly stationary. Under joint stationarity, standard estimation procedures (e.g., OLS, 2SLS, FIML, GMM) would generate constant coefficients, regardless of which lags or leads of variables were included or excluded in (2). Such estimators are functions of the sample data moments, which converge to their respective population moments; and the population moments are constant because the data are assumed stationary. Thus, with  $(y_t, x_t)$  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  $(W_1 \cap W_3)$  corroborates  $A$ . However,  $(W_2 \cap W_3)$  refutes  $A$ . If  $W_3$  occurs, then theory  $A$  implies that  $\gamma$  is nonconstant (via (8)), whereas  $W_2$  states that  $\gamma$  is constant. By the same token, theory  $A$  is inconsistent with the joint occurrence of all three pieces of evidence,  $(W_1 \cap W_2 \cap W_3)$ . Jointly,  $W_2$  and  $W_3$  refute theory  $A$  whereas individually they corroborate it. Thus, the constancy or otherwise of the parameters in (2) and (4) can have implications for whether or not, as a logical issue,  $y_t$  could have been generated by the expectations-based model (1).

To derive the implications of the evidence  $\{W_i\}$  for the conditional model, assume that (4) and the conditional model (2) generate the data, and then solve for  $\delta$  as a function of  $\gamma$  and  $\pi$ . Substitution of (5) into (2) gives:

$$y_t = (\gamma/\pi)\mathcal{E}(x_{t+1}|\mathcal{I}_{t-1}) + (e_t + \gamma u_t), \quad (9)$$

assuming  $\pi \neq 0$ . Thus,  $\delta = \gamma/\pi$  because  $\mathcal{E}([e_t + \gamma u_t]|\mathcal{I}_{t-1}) = 0$ . Implications of the evidence  $W_i$  for the conditional model (2) follow from (9). For clarity, the conditional

---

<sup>3</sup>There is a set of variations in  $\pi$  and  $\Sigma_{uu}$  such that their effects on  $\gamma$  just cancel each other. This set is ignored because it is such a restrictive special case.

model (2) is labeled as theory  $B$ :

$B$ :  $y_t$  is determined conditional upon  $x_t$  with constant parameterization  $\gamma$ .

As with theory  $A$ , the  $W_i$  individually corroborate theory  $B$ , and for similar reasons. The evidence  $W_2$  is a necessary condition for  $B$ ,  $W_3$  is irrelevant, and  $W_1$  would occur if the data were jointly stationary. Joint occurrences of evidence also parallel the results for  $A$ . Either  $(W_1 \cap W_2)$  or  $(W_2 \cap W_3)$  corroborates  $B$ . However,  $(W_1 \cap W_3)$  refutes  $B$ . If the process for  $x$  is nonconstant, then theory  $B$  implies that  $\delta$  is nonconstant (via (9)), whereas  $W_1$  states that  $\delta$  is constant. Theory  $B$ , like theory  $A$ , is inconsistent with the joint occurrence of all three pieces of evidence.

### 3.2 REMARKS

A few comments help clarify these results' ramifications. They touch on the comparison of models, the generality of the results, implications of refuting the Lucas critique, the role of nonconstancy, and finite sample inferences.

First, the implications of evidence for the expectations-based models is of primary interest empirically. These models typically are claimed to isolate the economic invariants of the process for  $y$ , yet empirical invariance is rarely examined. By contrast, a number of conditional models have remained empirically constant in spite of changes in the process for  $x$ . See Ericsson and Irons (1995), who summarize related empirical evidence for and against the Lucas critique.

Second, Nickell (1985) and Campbell and Shiller (1988) *inter alia* note a possible isomorphism between conditional error correction models and rational expectations models. That equivalence does not hold if the marginal process changes over time: the *data* can resolve the interpretation of a constant conditional model.

Third, the parameters of interest for the two hypotheses are not the same, being  $\delta$  for one and  $\gamma$  for the other. However, under each hypothesis, the corresponding parameter of interest is claimed to be invariant to changes in the distribution of  $x$  (e.g., changes in policy rules). Otherwise, the model would not be valid for two main economic purposes, forecasting and policy simulation, nor would the parameters have an economic interpretation.

Fourth, the models in Section 3.1 are simple. Variables are current-dated in the conditional model, occur at a single lead in the expectations-based model, or occur at a single lag in the model for  $x$ . Furthermore, the entire relationship determining  $y$  is bivariate and linear. Appendix B generalizes the results, allowing theory  $A$  to include multivariate relationships with multiple leads and lags. Qualitatively, the inferences about evidence remain unchanged, with a minor exception concerning expectations about current  $x$ ; see also Ericsson and Hendry (1989, Section 4). Nonlinear specifications also give qualitatively similar results; see Appendix B.

Fifth, and surprisingly, the expectations-based model (1) need not be estimated in order to refute it. As the theorem's application in Section 3.1 shows, the constancy

of the conditional model plus the nonconstancy of the model for  $x$  is inconsistent with (1). Furthermore, this set of evidence refutes not only (1) but a whole class of expectations-based models like it. This class includes all models with expectations of  $x$  and  $y$  at any horizons in the future (Appendix B). The intuition for generic refutation of expectations-based models is exactly as in Section 3.1 for the specific, one-period ahead, expectations-based model.

Sixth, the full process for  $x$  need not be specified in order to refute the Lucas critique: identifying a subset of  $x$ 's determinants is sufficient. If the simpler specification for  $x$  exhibits nonconstancy, then so must any suitably augmented process for  $x$ . A proof appears in Hendry (1988), and the standard formula for omitted variables bias provides the intuition. Suppose an additional variable  $z_t$  is required to make (4) well-specified:

$$x_t = \Psi_1 x_{t-1} + \Psi_2 z_t + u_{1t} \quad u_{1t} \sim NI(0, \Sigma_{11}) \quad (10)$$

$$z_t = \Phi x_{t-1} + u_{2t} \quad u_{2t} \sim NI(0, \Sigma_{22}), \quad (11)$$

where (11) is the marginal process defining how  $z$  behaves,  $u_{1t}$  and  $u_{2t}$  are mean-zero error terms with variances  $\Sigma_{11}$  and  $\Sigma_{22}$ , and  $\mathcal{E}(u_{1t}u_{2t}) = 0$  because (10) is an equation conditional on  $z_t$ . The least-squares estimator of  $\pi$  in (4) is subject to omitted variables bias. By assumption ( $W_3$ ),  $\pi$  changes. From the formula for omitted variables bias,  $\pi$  could have done so for one or more of three reasons: the underlying coefficient  $\Psi_1$  on  $x_{t-1}$  changed, the coefficient  $\Psi_2$  on  $z_t$  changed, or the coefficient  $\Phi$  relating  $z_t$  to  $x_{t-1}$  changed. Because the information set  $\mathcal{I}$  now includes  $z$ ,  $x$  in (4) can be redefined as  $(x, z)$ , with  $\pi$  a function of  $\Psi_1$ ,  $\Psi_2$ , and  $\Phi$ . The process for this augmented  $x$  must have changed because  $\Psi_1$ ,  $\Psi_2$ , or  $\Phi$  changed.

Seventh, nonconstancy somewhere in the economy is necessary for Hendry's (1988) type of test to work. If the conditional model is constant,  $x$  may be super exogenous, with implications for policy; see Engle, Hendry, and Richard (1983), Hendry (1988), and Engle and Hendry (1993). If the conditional model is nonconstant, its nonconstancy may arise for reasons other than the Lucas critique. Many forms of misspecification generate omitted-variables bias, which will alter as correlations between the included and excluded variables change. If the entire economic system is constant, then all regressions will have constant coefficients, at least in population, regardless of whether or not the regression represents a well-specified model.

Eighth, whether or not the Lucas critique applies is an empirical issue, not a theoretical one. For instance, the Lucas critique may be empirically refuted for certain sectors of an economy but not for others. Even with empirical refutation of expectations-based models, some forward-looking aspects may remain, such as data-based predictors in conditional models; see Campos and Ericsson (1988), Ericsson, Campos, and Tran (1990), Hendry and Ericsson (1991, pp. 864–866), and Ericsson and Irons (1995, p. 297). Relatedly, policy can affect agent behavior in conditional models, albeit through the

right-hand side variables themselves rather than through the parameters determining those variables.

Ninth, as with the evidence on the Marshall-Lerner condition, the individual pieces of evidence in Section 3.1 can be informative in themselves. For instance, the nonconstancy of  $\delta$  (the negation of  $W_1$ ) would immediately preclude the expectations-based model (1) as a satisfactory explanation of the data.

Tenth,  $W_1$ ,  $W_2$ , and  $W_3$  often are not known in practice, but the corresponding coefficients can be tested for constancy, e.g., using Chow's (1960) statistic in a recursive framework, Hoffman and Pagan's (1989) and Ghysels and Hall's (1990) statistic for the GMM estimator (generalizing on Chow), or Hansen's (1992a, 1992b) and Andrews's (1993) tests for nonconstancy with an unknown breakpoint. Thus, actual inferences about empirical models will have varying degrees of uncertainty associated with them.

## 4 ENCOMPASSING, CONGRUENCE, AND SEQUENTIAL CORROBORATION

The theorem's application in Section 3.1 points to the importance of encompassing evidence from alternative models. Encompassing a *set* of evidence augments standard testing procedures and corroboration from individual pieces of evidence. The current section discusses explicitly the relationships between encompassing, the theorem, and its application to expectations-based models.

Briefly, one model encompasses another if the first model can explain the properties of the second; see Mizon and Richard (1986) and Hendry and Richard (1989) for details. In Section 3.1, the expectations-based model fails to account for (or encompass) the constancy of the conditional model when the process for  $x$  is nonconstant. By construction, encompassing is a property that a model satisfies if the model is well-specified. While tests of encompassing do not require that either model be well-specified, the power of the tests may be affected if both models are mis-specified.

The theorem and its application in Section 3.1 stress the importance of examining the evidence as an entire set rather than piecemeal. That is, a model should be congruent with the data in the sense of Hendry and Richard (1982). Because empirical evidence often arises sequentially and from different authors, encompassing of other authors' results is critical to demonstrating that a model is congruent. In practice, encompassing frequently entails starting with a general model that nests all relevant existing empirical models and simplifying therefrom: general to specific modeling. Encompassing other authors' models one at a time could give rise to an apparently encompassing model not being able to encompass the other authors' models jointly, with (again) sequential corroboration implying refutation. This is closely related to why ordinary encompassing is not transitive but parsimonious encompassing is; cf. Hendry and Richard (1989).

Encompassing and congruence together are essential elements of a progressive re-

search strategy. White (1990, p. 381; 1994, Chapter 10) formalizes this in an econometric context; see also Pagan (1987, p. 6). Relatedly, Popper (1959, Section 82), Lakatos (1970), and Boland (1982, Chapter 1) discuss in detail the role of corroboration in a progressive research strategy.

Complete encompassing of conditional models by expectations-based models could be used instead of the limited encompassing proposed by Hendry (1988). For instance, given values for  $\pi$ ,  $\delta$ , and  $\Sigma_{u\varepsilon}/\Sigma_{uu}$ , a prediction of  $\gamma$  could be constructed from (8) and compared with the observed estimate of  $\gamma$  from (2). Although intuitively appealing and well-founded theoretically, such complete encompassing can prove difficult when  $x$  is determined by a complicated (and unknown) process. Even so, fully efficient estimation of the expectations-based model may well require a properly specified model for  $x$ . Conditional models with weakly exogenous regressors only require estimation of the conditional equation for efficiency. Relatedly, Govaerts, Hendry, and Richard (1994) show that the choice of completing model can affect the particulars of the encompassing test and its power. This may be particularly important for rational expectations models because of the leads and lags involved and possible gaps in dynamic structure.

## 5 AN EMPIRICAL EXAMPLE

This section illustrates the testing of expectations-based models by considering Hendry and Ericsson's (1991) empirically constant conditional model of the demand for narrow money ( $M_1$ ) in the United Kingdom. Cuthbertson (1988) sought to reinterpret an earlier version of this model [in Hendry (1985)] as a reduced form of a forward-looking process for money demand. However, Hendry (1988) established that the marginal processes for income, prices, and the interest rate were not constant over the sample. From Section 3, Cuthbertson's interpretation is thereby precluded. Hendry's (1985) evidence on the conditional model's constancy, followed by Hendry's (1988) evidence on the marginal models' nonconstancy, illustrates how sequential corroboration of forward-looking models by evidence on constancy can imply those models' refutation.

This section presents the conditional and marginal models and the corresponding evidence on constancy, and it relates that evidence to the theorem in Section 2. For ease of exposition, this section focuses on Hendry and Ericsson's (1991) model of money demand and their auxiliary models for prices and interest rates: their model of money demand is slightly simpler than the one in Hendry (1985) and spans a longer sample (1964–1989). The evidence on constancy is graphically conveyed, appearing in  $2 \times 2$  panels labeled  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ .

Hendry and Ericsson (1991, equation (6)) estimate the following conditional model for narrow money demand in the United Kingdom.

$$\begin{aligned}
\Delta(m-p)_t = & \underbrace{-0.69}_{(0.13)} \Delta p_t - \underbrace{0.17}_{(0.06)} \Delta(m-p-i)_{t-1} \\
& - \underbrace{0.093}_{(0.009)} (m-p-i)_{t-1} - \underbrace{0.630}_{(0.060)} R_t^{net} + \underbrace{0.023}_{(0.004)} \quad (12)
\end{aligned}$$

$$T = 100 [1964(3) - 1989(2)] \quad R^2 = 0.76 \quad \hat{\sigma} = 1.31\%$$

The variables are the stock of narrow money ( $M$ ), real total final expenditure ( $I$ ), the deflator thereof ( $P$ ), and a net interest rate ( $R^{net}$ ). Variables in lower case denote logarithms,  $\Delta$  is the first-difference operator,  $T$  is the sample size,  $t$  is the time index,  $R^2$  and  $\hat{\sigma}$  are the squared multiple correlation coefficient and the estimated equation standard error respectively, and estimated standard errors appear in parentheses ( . ). Data are quarterly, and details appear in Hendry and Ericsson (1991).

Equation (12) is a parsimonious error correction model, in which the growth of real money  $\Delta(m-p)_t$  depends on inflation  $\Delta p_t$ , lagged inverse velocity  $(m-p-i)_{t-1}$  and its change  $\Delta(m-p-i)_{t-1}$ , and the net interest rate  $R_t^{net}$ .<sup>4</sup> The long-run static solution is:

$$m-p-i = -6.78 R^{net} + 0.25, \quad (13)$$

which is obtained by setting  $\Delta m = \Delta p = \Delta i = 0$  in (12) and re-arranging terms. The long-run elasticity of money with respect to prices is unity, as is its elasticity with respect to income. Its long-run semi-elasticity with respect to the interest rate is nearly  $-7$ . Hendry and Ericsson (1991) discuss additional economic properties of (12).

Empirical evaluation of a model's constancy is easily performed with recursive estimation; see Doornik and Hendry (1996) for computational and statistical details. Figures 2a and 2b plot the recursive least-squares estimates for the coefficients on  $\Delta p_t$  and  $R_t^{net}$  and twice their estimated standard errors:  $\{\hat{\beta}_t \pm 2\text{ese}(\hat{\beta}_t)\}$  in a common notation. The recursive estimates vary little, especially relative to their estimated uncertainty; and the estimates become much more precise over time. Figure 2c plots the corresponding one-step residuals and plus-or-minus twice the estimated equation standard error: i.e.,  $\{y_t - \hat{\beta}_t' x_t\}$  and  $\{0 \pm 2\hat{\sigma}_t\}$ . The estimated equation standard error varies little over the sample, and few residuals are larger than twice the estimated equation standard error. Finally, Figure 2d plots the breakpoint Chow (1960) statistics for testing predictive failure, normalized by their one-off 5% critical values. The horizontal axis is the date of the breakpoint between estimation and forecast samples. No split of the sample results in a rejection of constancy. This constancy occurs in spite of important in-sample events: changes in government (1970, 1974, 1979), increases in oil prices by OPEC (1973, 1979), financial innovation (1985), and the introduction of Competition and Credit Control (1971).

---

<sup>4</sup>Technically speaking, (12) is an equilibrium correction model; see Hendry (1995).

Hendry and Ericsson (1991, equations (8) and (9)) also estimate marginal equations for inflation and the interest rate.<sup>5</sup>

$$\Delta p_t = \begin{matrix} 0.62 \\ (0.10) \end{matrix} \Delta p_{t-1} + \begin{matrix} 0.23 \\ (0.09) \end{matrix} \Delta p_{t-2} + \begin{matrix} 0.023 \\ (0.008) \end{matrix} D_t + \begin{matrix} 0.0030 \\ (0.0014) \end{matrix} \quad (14)$$

$$T = 100 [1964(3) - 1989(2)] \quad R^2 = 0.72 \quad \hat{\sigma} = 0.768\%$$

$$\Delta R_t^{net} = \begin{matrix} -0.095 \\ (0.037) \end{matrix} R_{t-2}^{net} + \begin{matrix} 0.0085 \\ (0.0035) \end{matrix} \quad (15)$$

$$T = 100 [1964(3) - 1989(2)] \quad R^2 = 0.06 \quad \hat{\sigma} = 0.0136$$

Both equations were derived as simplifications from fourth-order autoregressive models; and  $D$  is a zero-one impulse dummy for 1979(3) that aims to capture the one-off effect on inflation of an increase in VAT. Figures 3 and 4 plot recursive estimates and statistics for equations (14) and (15) respectively. Both figures point to the nonconstancy of those equations. In 1973–1974, the recursive estimates vary notably, with statistically significant changes in estimated coefficients. In Figures 4a and 4b in particular, the 95% confidence regions for the full-sample estimates lie well outside the 95% confidence regions for the initial subsample. The estimated equation standard errors for (14) and (15) increase by 38% and 135% respectively over the sample, indicating a marked worsening of fit, as detected by the recursive Chow statistics. Cuthbertson's (1991) concern about finite sample power appears unimportant here, as substantial numerical and statistical nonconstancy is present.

In the notation of the theorem, the evidence on constancy is as follows.

$W_2$ : the conditional money-demand model (12) has constant parameters (Figure 2).

$W_3$ : the marginal models for inflation and the net interest rate [(14) and (15)] have nonconstant parameters (Figures 3 and 4).

Theory  $A$  is the expectations-based model:

$A$ : money is determined via expectations of inflation and the interest rate with a constant parameterization.

The combination of empirical evidence is inconsistent with models of money based on forward-looking expectations, but it is consistent with a constant conditional model of money demand.

## 6 CONCLUSIONS

Because a sequence of apparently confirming evidence can refute a theory, it is important to examine all available evidence on an empirical model jointly, rather than

---

<sup>5</sup>No marginal model for total final expenditure is considered because that variable enters (12) only at a lag and not contemporaneously.



simply to 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)

Friedman additionally emphasizes the importance of seeking evidence that could potentially refute the model or theory being investigated, rather than just looking to support a model through corroborating evidence. To avoid potential apparent paradoxes such as those discussed in Sections 2–5, inference should be conducted within the framework of general to specific, at least implicitly so by always testing a conjectured model against the most unrestricted model that is logically entailed by the evidence. Recently developed econometric tools provide a potentially powerful way of testing rational expectations and conditional models, using evidence on model constancy.

## APPENDIX A: PROOF OF THE THEOREM

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

## APPENDIX B: EVIDENCE ON GENERAL RATIONAL EXPECTATIONS MODELS

This appendix derives the relationship between coefficients in a conditional model and the parameters of a general rational expectations model and ties evidence on model constancy to the theorem. In so doing, it generalizes related results in Hendry (1988), Favero and Hendry (1992) and Hendry (1995, Chapter 14).

The feedforward framework can be characterized in the following manner. Agents make decisions about a variable  $y$  in light of their expectations about future values of some strictly exogenous variables  $x$  and about future values of  $y$  itself. Expectations are formed, conditional upon an information set  $\mathcal{I}_{t-1}$ , which usually includes lagged values of  $x$  and  $y$ .

$$\sum_{i=0}^{\infty} \kappa_i \mathcal{E}(y_{t+i} | \mathcal{I}_{t-1}) = \sum_{i=0}^{\infty} \lambda'_i \mathcal{E}(x_{t+i} | \mathcal{I}_{t-1}) \quad (16)$$

In (16),  $\kappa_0 = 1$ , the  $\{\lambda_i\}$  and the remaining  $\{\kappa_i\}$  are the deep structural parameters of the agents' behavior or direct functions of them, and typically  $\mathcal{E}(y_t | \mathcal{I}_{t-1}) \equiv y_t - \varepsilon_t$ , where  $\varepsilon_t$  is an innovation with respect to  $\mathcal{I}_{t-1}$ . Under weak conditions, the expectation of  $y_t$  can be expressed in terms of expectations of the  $\{x_{t+i}\}$  alone:

$$\mathcal{E}(y_t | \mathcal{I}_{t-1}) = \sum_{i=0}^{\infty} \delta'_i \mathcal{E}(x_{t+i} | \mathcal{I}_{t-1}), \quad (17)$$

where each  $\delta_i$  depends upon  $\{\lambda_i, \kappa_i\}$ , and  $\sum_{i=0}^{\infty} \delta_i$  is finite.

For completeness (and, e.g., for estimation), it is necessary to specify the process for  $x$ , which is:

$$\mathcal{E}(x_t | \mathcal{I}_{t-1}) = \pi x_{t-1}, \quad (18)$$

again assuming linearity. If  $x_t$  depends upon several of its own lags, rather than just one, that dependence can be rewritten as (18) by stacking the lags and redefining  $x_t$ . Lagged  $y$  is excluded from (18), as expectations-based models often assume strictly exogenous  $x$ . Lags of  $y$  in (18) complicate the algebraic relations below, but the conclusions remain qualitatively the same, whether or not  $x$  is strictly exogenous.

Equations (17) and (18) can be written in model form as:

$$y_t = \sum_{i=0}^{\infty} \delta'_i \mathcal{E}(x_{t+i} | \mathcal{I}_{t-1}) + \varepsilon_t \quad \varepsilon_t \sim NI(0, \Sigma_{\varepsilon\varepsilon}) \quad (19)$$

$$x_t = \pi x_{t-1} + u_t \quad u_t \sim NI(0, \Sigma_{uu}), \quad (20)$$

where  $\mathcal{E}(\varepsilon_t | \mathcal{I}_{t-1}) = 0$  and  $\mathcal{E}(u_t | \mathcal{I}_{t-1}) = 0$ . For expositional convenience,  $(\varepsilon_t, u_t)$  is assumed to be independently and identically distributed, normal. The conditional model is:

$$y_t = \gamma' x_t + e_t \quad e_t \sim NI(0, \Sigma_{ee}), \quad (21)$$

where  $\gamma$  is the parameter of interest, and  $e_t$  satisfies  $\mathcal{E}(e_t | \mathcal{I}_{t-1}) = 0$ . However, assuming (19) and (20),  $\gamma$  is a derived parameter and is a complicated function of  $\{\delta_i\}$  and  $\pi$ .

To express  $y_t$  in (19) explicitly in terms of the observed  $x_t$  requires two steps. First, repeated substitution of (20) into itself obtains:

$$\mathcal{E}(x_{t+i} | \mathcal{I}_{t-1}) = \pi^i \mathcal{E}(x_t | \mathcal{I}_{t-1}) = \pi^i (x_t - u_t). \quad (22)$$

Second, direct solution of (19) with (22) gives:

$$y_t = \left( \sum_{i=0}^{\infty} \delta'_i \pi^i \right) x_t + \left[ \varepsilon_t - \left( \sum_{i=0}^{\infty} \delta'_i \pi^i \right) u_t \right], \quad (23)$$

assuming that  $x_t$  is stationary. From (3), the least-squares estimator of the coefficient on  $x_t$  in (21) converges to:

$$\gamma = \left( \sum_{i=0}^{\infty} \{\pi'\}^i \delta_i \right) + (M_{xx})^{-1} \left[ \Sigma_{u\varepsilon} - \Sigma_{uu} \left( \sum_{i=0}^{\infty} \{\pi'\}^i \delta_i \right) \right], \quad (24)$$

where  $M_{xx}^\nu \equiv \mathcal{E}(x_t x_t')^\nu = (I - \pi \otimes \pi)^{-1} \Sigma_{uu}^\nu$ ,  $\nu$  is the column vectoring operator, and  $\otimes$  is the corresponding Kronecker product.

The coefficient  $\gamma$  is constant if  $\pi$ ,  $\Sigma_{uu}$ ,  $\Sigma_{u\varepsilon}$ , and the  $\{\delta_i\}$  are. However, as changes occur in the process generating  $x$  (e.g., as  $\pi$  varies over time),  $\gamma$  also will change and the conditional model (21) will break down. The Lucas critique applies, and (21) fails to isolate the  $\{\delta_i\}$ . Equation (16) (and so (19)) remains constant in spite of (20) evolving. Conversely, if  $\gamma$  is constant in spite of changes in the process for  $x$ , then  $y$  could not have been generated by (16) with constant deep parameters. In this case, the Lucas critique is refuted empirically. That is, because of the relationship (24), the expectations-based model (16) is inconsistent with the observations that  $\gamma$  is constant *and* that  $\pi$  and/or  $\Sigma_{uu}$  have changed.

To relate this to Section 2, the evidence is categorized as follows.

$W_1$ :  $\{\lambda_i\}$  and  $\{\kappa_i\}$  (and so  $\{\delta_i\}$ ) are constant.

$W_2$ :  $\gamma$  is constant.

$W_3$ :  $\pi$  and/or  $\Sigma_{uu}$  are nonconstant.

The theories of interest correspond to constant-parameter models in (16) and (21).

*A*:  $y_t$  is determined via expectations of future  $y$  and of current and future  $x$ .

*B*:  $y_t$  is determined conditional upon  $x_t$ .

The implications of evidence (both individually and jointly) for these more complicated models are the same as those described in Section 3.1 for the simple bivariate models.

Equation (16) assumes linearity, but the issues raised also apply to nonlinear models involving expectations. If (16) were a nonlinear rather than a linear difference equation in expectations, then (24) would include an approximation error due to the linearization of that difference equation by (17). The least-squares estimator for (21) still would have a probability limit similar in form to  $\gamma$  in (24), but with an additional term introduced by the approximation error; cf. White (1980). The least-squares bias from the approximation error would be a function of population data moments involving  $x_t$ , so finding a constant conditional model would refute the claimed nonlinear expectations-based model when (20) is nonconstant. Likewise, the conditional expectation for  $x_t$  in (18) might be nonlinear, in which case (20) with  $\pi$  would be the least-squares approximation to that nonlinear function. Again, such nonlinearities would induce nonconstancy in  $\gamma$  under the hypothesis that (16) and a nonlinear version of (18) were the data generation process.

- Andrews, D. W. K. (1993) "Tests for Parameter Instability and Structural Change with Unknown Change Point", *Econometrica*, 61, 4, 821–856.
- Boland, L. A. (1982) *The Foundations of Economic Method*, London, George Allen and Unwin.
- 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., June.
- 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 M1: A Forward Looking Buffer Stock Model", *Oxford Economic Papers*, 40, 1, 110–131.
- Cuthbertson, K. (1991) "The Encompassing Implications of Feedforward versus Feedback Mechanisms: A Reply to Hendry", *Oxford Economic Papers*, 43, 2, 344–350.
- Doornik, J. A., and D. F. Hendry (1996) *PcGive Professional 9.0 for Windows*, London, International Thomson Business Press.
- Engle, R. F., and D. F. Hendry (1993) "Testing Super Exogeneity and Invariance in Regression Models", *Journal of Econometrics*, 56, 1/2, 119–139.
- Engle, R. F., D. F. Hendry, and J.-F. Richard (1983) "Exogeneity", *Econometrica*, 51, 2, 277–304.
- Ericsson, N. R., J. Campos, and H.-A. Tran (1990) "PC-GIVE and David Hendry's Econometric Methodology", *Revista de Econometria*, 10, 1, 7–117.
- Ericsson, N. R., and D. F. Hendry (1989) "Encompassing and Rational Expectations: How Sequential Corroboration Can Imply Refutation", International Finance Discussion Paper No. 354, Board of Governors of the Federal Reserve System, Washington, D.C., June.
- Ericsson, N. R., and J. S. Irons (1995) "The Lucas Critique in Practice: Theory Without Measurement", Chapter 8 in K. D. Hoover (ed.) *Macroeconometrics: Developments, Tensions, and Prospects*, Boston, Kluwer Academic Publishers, 263–312.
- Favero, C., and D. F. Hendry (1992) "Testing the Lucas Critique: A Review", *Econometric Reviews*, 11, 3, 265–306.
- Friedman, M. (1951) "Comment" in G. Haberler (ed.) *Conference on Business Cycles*, New York, National Bureau of Economic Research, 107–114.
- Ghysels, E., and A. Hall (1990) "A Test for Structural Stability of Euler Conditions Parameters Estimated via the Generalized Method of Moments Estimator", *International Economic Review*, 31, 2, 355–364.
- Govaerts, B., D. F. Hendry, and J.-F. Richard (1994) "Encompassing in Stationary Linear Dynamic Models", *Journal of Econometrics*, 63, 1, 245–270.

- Hansen, B. E. (1992a) “Testing for Parameter Instability in Linear Models”, *Journal of Policy Modeling*, 14, 4, 517–533.
- Hansen, B. E. (1992b) “Tests for Parameter Instability in Regressions with I(1) Processes”, *Journal of Business and Economic Statistics*, 10, 3, 321–335.
- Hendry, D. F. (1985) “Monetary Economic Myth and Econometric Reality”, *Oxford Review of Economic Policy*, 1, 1, 72–84.
- Hendry, D. F. (1988) “The Encompassing Implications of Feedback versus Feedforward Mechanisms in Econometrics”, *Oxford Economic Papers*, 40, 1, 132–149.
- Hendry, D. F. (1995) *Dynamic Econometrics*, Oxford, Oxford University Press.
- Hendry, D. F., and N. R. Ericsson (1991) “Modeling the Demand for Narrow Money in the United Kingdom and the United States”, *European Economic Review*, 35, 4, 833–886 (with discussion).
- 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 (1989) “Recent Developments in the Theory of Encompassing”, Chapter 12 in B. Cornet and H. Tulkens (eds.) *Contributions to Operations Research and Economics: The Twentieth Anniversary of CORE*, Cambridge, Massachusetts, MIT Press, 393–440.
- Hoffman, D., and A. R. Pagan (1989) “Post-sample Prediction Tests for Generalized Method of Moments Estimators”, *Oxford Bulletin of Economics and Statistics*, 51, 3, 333–343.
- Keynes, J. M. (1921) *A Treatise on Probability*, London, Macmillan.
- Krugman, P. R., and M. Obstfeld (1994) *International Economics: Theory and Policy*, New York, HarperCollins College Publishers, Third Edition.
- 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.) *The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy*, Volume 1, *Journal of Monetary Economics*, supplementary issue, 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. R. (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.

- Taylor, J. B. (1989) “The Current Account and Macroeconomic Policy: An Econometric Analysis”, Chapter 2 in A. E. Burger (ed.) *U.S. Trade Deficit: Causes, Consequences, and Cures*, Boston, Kluwer Academic Publishers, 131–191 (with discussion).
- White, H. (1980) “Using Least Squares to Approximate Unknown Regression Functions”, *International Economic Review*, 21, 1, 149–170.
- White, H. (1990) “A Consistent Model Selection Procedure Based on  $m$ -testing”, Chapter 16 in C. W. J. Granger (ed.) *Modelling Economic Series: Readings in Econometric Methodology*, Oxford, Oxford University Press, 369–383.
- White, H. (1994) *Estimation, Inference and Specification Analysis*, New York, Cambridge University Press.