On Congruent Econometric Relations

David F. Hendry
Nuffield College,
Oxford *

March 30, 2008

Abstract

The paper responds to the claimed criticisms in Faust and Whiteman (1997) of the 'LSE' approach to empirical macro-econometrics. Their bold attempt to inter-relate 'alternative methodologies' is undone by misconceptions which lead them to incorrect conclusions. To clarify the analysis, the notion of structure is reviewed to underpin sustainable econometric relations, then identification and exogeneity are reconsidered. The theory of reduction and modelling strategies are examined, as is the basis for policy. The results rebut their claims, and confirm the need for a progressive research strategy which combines the best of both theory and evidence.

1 Introduction

Jon Faust and Charles Whiteman have devoted considerable effort to investigating the 'LSE' approach to dynamic econometric modeling. Since criticism is the lifeblood of science, I welcome their discussion. One must applaud their attempt to inter-relate 'alternative methodologies' for obtaining evidence in empirical economics. There are important areas of agreement between us, as in our joint doubts about 'specific-to-general' modeling strategies that lack pre-assigned stopping rules, and the recommendation that "Finding improved ways to exploit economic theory while thoroughly examining the empirical properties of the models would surely be beneficial". Equally, there are major areas of disagreement, as in our views about the identification of 'LSE' models, and sustainable bases for policy analyses. Such an exchange of ideas is doubtless part of the process foreseen in the opening address of *Econometrica* by Schumpeter (1933): "Theoretic and 'factual' research will of themselves find their right proportions ... by positive achievement".

Some of the disagreements derive from misconceptions in their paper. Specifically, they misunderstand exogeneity, identification of rival econometric models, and the conditions required for viable policy analyses using empirical econometric models. Further, they confound the theory of modeling with its empirical applications (i.e., they ignore the population/sample dichotomy), and mix the initial general model with simplifications thereof. These errors interact to lead them to a series of false conclusions about identification and the role of structural breaks, which are overturned on a more careful analysis.

The paper focuses on eight aspects of their discussion. First, it reconsiders their general views about RBC and VAR approaches in comparison to the 'LSE' approach. Secondly, it analyzes the notion of structure (invariance under extensions of the information set over time, across regimes, and for new

*Prepared for the Carnegie–Rochester Conference Series on Public Policy, 1996. Financial support from the UK Economic and Social Research Council under grant R000234954 is gratefully acknowledged. It is a pleasure to acknowledge helpful discussions with Jurgen Doornik, Neil Ericsson, Grayham Mizon and Jean-Francois Richard.

sources of information), and discusses how to find structure in empirical research, to develop sustainable econometric relations. It then turns to identification, its application when rival models exist, and the identifiability of structure, thereby elucidating the role which structural breaks play in modeling. Because of the central role of weak exogeneity in most econometric approaches, the concept is clarified, and its relation to simultaneity, and 'efficient estimation' analyzed. The theory of reduction and empirical modeling strategies are then examined, together with the basis for drawing policy implications in economics. Finally, I comment on some rhetorical devices used by Faust and Whiteman, rebut their citation claims, and conclude.

2 Alternative modeling approaches

The characterization of RBC and VAR approaches by Faust and Whiteman as "1) start with ...; 2) stop", is seriously misleading. It crucially ignores the later reformulation and reapplication of those approaches to essentially the same data, with modifications induced by earlier empirical mis-matches or theoretical advances. Their apparent belief that there is some qualitative difference between "within study" and "between study" revisions has no substance. Indeed, ignoring the dimension of repeated evaluation on a common database camouflages incipient "data mining": each study may be start–stop, but the sequence is not. Conversely, if they were to wait for sufficient independent data to sustain new analyses, their progress would be a small fraction of that resulting from the methods which they seek to criticize. Rather, the advantage of a later study should be genuine tests of previous models on new evidence — precisely what is emphasized in the 'LSE' approach — not starting an unrelated new search for empirical structure. Since the latter aims to develop congruent models, namely ones that are coherent with all the sources of information and sustainable over time, there are many potential cross-checks on consistency.

Because of this lacuna, Faust and Whiteman fail to realize that many of the 'LSE' criticisms apply to sequences of RBC and VAR models. For example, any 'test' that affects decisions about later model specification becomes an index of design adequacy whatever the theory starting point (see e.g., Pagan, 1987). The concept of "a priori" ceases to be well defined when the current economic theory formulation is closely dependent on previous empirical findings; there is a far smaller status difference than they realize between their views and the 'LSE' approach, which utilizes such evidence as well as 'institutional' knowledge. Thus, the final sentence of their paper may have reached the correct conclusion for the wrong reason!

From the outset, they have missed the main message of the 'LSE' approach: embed empirical econometric modeling in a progressive research strategy where all the available evidence and theoretical insights are utilized in a mutually interactive approach. Their implicit attitude that only a single 'one-off' study occurs is a travesty of empirical modeling in economics. If they dislike the fact that science is a progressive sequence of learning about the coherence of theories and evidence by sequentially updating both, as well as the instruments used, they need to offer a dramatic new paradigm, not pretend that some approaches can avoid the need to revise when adverse results are found.

3 Structure

To clarify the basis of the 'LSE' approach to determining sustainable empirical economic relations, I first consider two other key areas which Faust and Whiteman neglect, namely, structure and its determination. This will let us examine their claims about the identifiability of the resulting models.

The concept of structure has many meanings in econometrics. It began by denoting a set of exact relations (Frisch, 1934), then evolved to mean a system of invariant (autonomous) equations which characterized the behavior of economic agents (Haavelmo, 1944, and Wold and Juréen, 1953). Later, it became an entity (called a structural model) to be contrasted with a system having derived parameters (called a reduced form): see Koopmans (1950). Structure has also been employed as a synonym for the population value $\boldsymbol{\theta}_0$ of an unknown parameter $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ (see Bårdsen and Fisher, 1993, and Juselius, 1993). Another usage connotes 'being derived from inter-temporal optimization by economic agents', so such a structure aims to embody the 'deep parameters' of the relevant theory (see e.g., Kydland and Prescott, 1991).

Parts of all these ideas persist in general parlance. For example, consider (see Hendry, 1995b):

$$\mathsf{E}\left[y_t \mid \mathcal{I}_{t-1}\right] = \rho \mathsf{E}\left[z_t \mid \mathcal{I}_{t-1}\right] \text{ for } \rho \in \mathbb{R}. \tag{1}$$

In (1), y_t and z_t are the 'output' and 'input' variables for an agent's decision, given available information $\mathcal{I}_{t-1}$, where $\mathsf{E}[\cdot|\cdot]$ denotes the conditional expectations operator. This equation satisfies the first sense (i.e., is an exact relation), could satisfy the second and third depending on the state of nature (i.e., ρ might be invariant to changes in the generating process of z_t; and the equation could represent the structure from which mappings of y_t and z_t onto $\mathcal{I}_{t-1}$ are derived); when $\rho = \rho_0$, it satisfies the fourth; and if a rational-expectations optimization theory underlay the equation, it could match the last as well. Conversely, the same equation could fail on the last four senses if ρ was not invariant; the equation was derived by solving a more 'basic' set of relations which were the actual decision functions of agents; ρ did not have a fixed population value; and was not a deep parameter.

I construe structure as the set of invariant features that directly characterize the economic mechanism. This notion aims to capture the idea of permanence, within an economic 'framework' which is hidden from direct view and needs to be uncovered (see Hendry, 1995a). The parameters of an economic structure generally include those of agents' decision rules. Then, $\boldsymbol{\theta}$ defines a structure if it is invariant and directly characterizes the relations of the economy under analysis (i.e., is not derived from more basic parameters). Structure, therefore, entails a correspondence between model and reality that is not fully open to independent testing. Even so, a parameter can be structural only if it is constant (invariant to an extension of the sample period), remains the same under policy interventions or natural experiments (invariant to changes elsewhere in the economy), and is not altered by extensions of the information set (invariant to adding variables to an analysis). A 'parameter' that changed with policy interventions would be non-constant over time; and one that altered with model specification would not be a well-defined entity. Thus, all three conditions are required, but as these aspects are open to empirical scrutiny, necessary conditions for structure are testable, even if sufficient ones are not.

Succinctly, 'LSE' focuses on structure as invariance under extensions of the information set over time, across regimes, and for new sources of information. Whether the model corresponds to agent behavior is less easily tested, but indirect evidence can be obtained by testing theoretically-based restrictions. However, since structure is a relation between an empirical model and reality, theory *per se* cannot endow structure, although it may lead to a model that does. Conversely, aspects of structure are amenable to empirical discovery (see the examples in Hendry, 1995b), so as Frisch (1933) expressed it, the 'mutual penetration of quantitative economic theory and statistical observation' must have center stage.

3.1 Finding structure

Given this characterization of structure, how can it be discovered? As a correspondence between invariant features of reality and an empirical model, an investigator needs at least one of the following ingredients:

a) brilliance;
b) creative ability;
c) luck.

Empirical econometric modeling usually draws on all three.[1] Fortunately, it has some supportive companions: steadily improving economic theory, accumulating empirical evidence, and advances in econometric instruments. These have no precedence ordering when undertaking new empirical research: which it is best to exploit most in any historical period depends on how good that companion happens to be: contrast phlogiston theories of combustion with thermodynamics as a theory basis; and biology pre and post microscopes. In general, to discover, and to evaluate claimed discoveries, all three interact and mutually complement each other. They underlie the progressive accumulation of knowledge that is the hallmark of any science.

In this context, the 'LSE' econometric approach is 'theory neutral' in that one can apply it after any economic theory or postulated model starting point, whether it be RBC, VAR, or Keynesian macro-systems. Any viable empirical methodology must be neutral in that sense. For example, in an RBC approach, one can check theories of technology shocks being the source of non-stationarity (corresponding to common factors of unity in Sargan, 1980), or the constancy over time of the 'key moments' to be matched; in VARs, one can check the constancy of the 'parameters' of interest (e.g., impulse response patterns in subsamples); in Keynesian models, the dynamics and error specifications can also be tested. Thus, one can embed economic-theory models in general empirical representations to check if the former matches the salient features of the latter. Later, one can test the outcome against new evidence, perhaps simplifying the model to improve evaluation power. Obviously, many non-'LSE' approaches also embody tests of auxiliary hypotheses – see, *inter alia*, the Bayesian approach in Florens, Mouchart and Rolin (1990) – so it does not claim a monopoly.

If the postulated model is a truly brilliant characterization of the behavior of the relevant agents, it will pass all the specification and mis-specification tests, subject to the inevitable caveat about Type-I errors. But what if it fails at a significance level far beyond possible rationalization as a Type-I error? Either an anomaly must be tolerated or some form of revision undertaken (see §2 above). Unfortunately, one cannot uniquely infer the causes of failure from the observed outcome (a variant of the Duhem–Quine thesis: see e.g., Cross, 1982). Thus, one implication of the 'LSE' approach, on which Faust and Whiteman also seem to concur, is "do not patch": a specific-to-general approach with randomly-determined stopping rules does not work well in practice (see e.g. Anderson, 1971, and Mizon, 1995). Rather, rejection implies 'back to the drawing board' for all ingredients. Conversely, repeatedly postulating and rejecting a sequence of models ignoring previous reject information could be an endless and fruitless process. Thus, a framework for accumulating and consolidating knowledge is essential: minimal re-usage of available data is most desirable precisely when aggregate experimentation is expensive and difficult. A structured modeling approach helps do so (§6.1 discusses model simplification).

The next key issue is whether structure is identifiable, and to resolve that I first reconsider the general problem of identification.

[1] There is no uniquely best way to discover structure, and 'LSE' does not claim such, but seeks to place this aspect of the debate in the context of the research efficiency of alternative strategies.

4 Identification

Identification has three distinct attributes: 'uniqueness', 'correspondence to the desired entity', and 'satisfying the required interpretation (of a theory model)'. As an example, any regression of quantity on price delivers a unique function of the data second moments, so satisfies the first attribute. However, such a regression need not correspond to any underlying demand or supply behavior, and may be a composite of several relations as in textbook illustrations. Further, despite being a composite, the outcome may be incorrectly interpreted as a supply schedule due to a positive sign on price. In practice, the meaning of 'identified' usually entails all three notions, as in 'Have you identified the parameters of the money-demand function?' – i.e., is it unique, a demand schedule, and maps how agents behave.

The first sense of identification was used by the Cowles Commission (Koopmans, Rubin and Leipnik, 1950) when formalizing the conditions for the uniqueness of coefficients in simultaneous systems. This is the sense on which Faust and Whiteman focus in their discussion of rank and order conditions, checking if there are other (linear) combinations of the relations in the system that also satisfy the restrictions on the equation of interest. If so, then the claimed equation is not unique and hence not identified. Conditions for the correspondence of a model's parameters to those of the DGP are not easily specified, but usually one can directly check the interpretation of parameters in the light of a theory.

It is well known that the rank condition is not necessary for identification when other forms of restriction exist (see e.g., Fisher, 1966, and Hausman, 1983). Further, the rank condition is not sufficient for the two other attributes of identification just discussed. Perhaps the most baffling aspect of the Faust and Whiteman 'critique' is the conjunction of their claims that covariance restrictions are frowned upon in the 'LSE' approach, and that 'LSE-style' models are not identified by the rank condition. Many of the 'LSE' models which they seek to criticize are based on economic theories in which agents act contingently on the available information. Faust and Whiteman seem suspicious of such an approach, and that is their privilege: the proof of empirical puddings lies in their eating, not a priori views. But 'suspicion' does not justify their ignoring the fact – of which they are well aware – that such models are uniquely identified by conditioning. Given such identification, then other hypotheses become testable; the equations in question are not identified by that other information. In the context of systems modeling, full information maximum likelihood will fail unless the rank condition is satisfied, so 'prior' knowledge is certainly required, and §2 discussed that issue.[2]

For example, when 'LSE' single-equation models are identified (in the Cowles' sense) by conditioning, predicated on that, constancy and invariance can be tested, as can other auxiliary hypotheses if the model specification is not rejected. Genuine tests always require new evidence, which accrues steadily over time, and hence can be conducted systematically. For example, when a new structural break has occurred in one area of the economy, invariance of other relations can be tested. If the equations pass, then their evidential basis is strengthened; for those that fail, then as always in science, better ideas are needed (although large data revisions may sometimes reinstate models later).

Their confusion about 'LSE' views on identification seems to stem from not differentiating within-theory from between-theory discrimination, so I turn next to clarify that issue.

4.1 Insufficiency of over-identification to determine structural models

Despite the rigor of the Cowles developments on identification, uniqueness on the rank condition is insufficient to determine whether a given model actually corresponds to the underlying structure. For

[2] PcFiml was the first computer program to formally check for rank conditions *prior* to undertaking FIML estimation: see Hendry, Neale and Srba (1988).

example, all just-identified 'structural' equations are unique, are not rejectable against their 'reduced form', and have several representations which are observationally equivalent. It follows inexorably that finding an equation which is identified by the rank condition does not ensure that it corresponds to the underlying structural model, even when that model satisfies the theory restrictions. This problem is exacerbated when the reduced form does not characterize the data, so none of the observationally-equivalent representations is in fact structural.

Now I have reached the issue that concerns 'LSE' econometricians, which is not some unjustifiable attempt to 'redefine' the Cowles' results. Several models can be over-identified, satisfy the rank condition, and not fail over-identification tests empirically, even when such models conflict theoretically (see Hendry and Mizon, 1993). Thus, the rank condition is insufficient for the three attributes, although it is sufficient for uniqueness within theories. Readers may be forgiven for thinking that the exposition they are offered on the rank condition is complete, as Faust and Whiteman claim to consider all linear transforms $\mathbf{R}$ of $(\mathbf{\Gamma} : \mathbf{B})$ in their equation (3) to establish whether the uniquely admissible $\mathbf{R}$ is $\mathbf{R} = \mathbf{I}$. When $(\mathbf{\Gamma} : \mathbf{B})$ are unconstrained, $(\mathbf{R\Gamma} : \mathbf{RB})$ comprise all linear systems. Otherwise, admissible $\mathbf{R}$s are relative to the restrictions on the given $(\mathbf{\Gamma} : \mathbf{B})$ so no longer span all relevant linear models. The 'LSE' concern is that there can be other $(\mathbf{\Gamma}^* : \mathbf{B}^*)$ which have their own (but different) over-identifying restrictions such that only $\mathbf{R}^* = \mathbf{I}$ is admissible for them. This proliferation problem is exacerbated by modelers using different selections of the unmodeled variables $\mathbf{Z}_t$ and dynamic specifications, so that $\mathbf{X}_t$ vectors differ, combined with a failure to test the adequacy of the reduced form.

If the world were a weakly stationary process, and several such models were postulated, empirical evidence would not discriminate between them. Consistency over one time period would virtually ensure consistency over another. This is the stage onto which the regime-shift analysis enters: by inducing structural change in parts of the economic system, some non-structural representations will cease to be invariant. In terms of Faust and Whiteman's translation into 'extra instruments', by disturbing the status quo, the new variables reveal which (falsely-claimed) structures can be rejected. As an analogy, if the solar system were truly stationary, alternative models thereof would be hard to discriminate: but a massive object perturbing orbits would undoubtedly allow rejection of some models (perhaps causing mayhem to humanity en route – as some bad economic policies have done historically).

4.2 Identifying structure

By 1938, Frisch had evolved a theoretical viewpoint which he thought entailed that structure would be unidentifiable (see Frisch, 1938, and the commentary offered in Hendry and Morgan, 1995). As an example of unidentified structure, consider the system:

$$\begin{aligned} q_t - \alpha p_t - \lambda m_t &= \boldsymbol{\beta}_1' \mathbf{z}_t + \mathbf{v}_{1,t} && (2) \\ m_t - \phi p_t &= \boldsymbol{\beta}_2' \mathbf{z}_t + \mathbf{v}_{2,t} && (3) \\ p_t &= \boldsymbol{\beta}_3' \mathbf{z}_t + \mathbf{v}_{3,t}. && (4) \end{aligned}$$

Here the $\mathbf{z}_t$ are 'exogenous' in every sense, and the $\mathbf{v}_{i,t}$ are errors about which nothing is known. When there are no restrictions on $\boldsymbol{\beta}_1$, then (2) is not identifiable, since within this specification, the rank condition is definitive. However, consider a setting where no one knew that m_t mattered, so it failed to enter any analyses, delivering the non-structural system:

$$\begin{aligned} q_t - (\alpha - \lambda\phi)\, p_t &= \left(\boldsymbol{\beta}_1' + \lambda \boldsymbol{\beta}_2'\right) \mathbf{z}_t + \mathbf{v}_{1,t} + \lambda \mathbf{v}_{2,t} && (5) \\ p_t &= \boldsymbol{\beta}_3' \mathbf{z}_t + \mathbf{v}_{3,t}. && (6) \end{aligned}$$

It is easy to imagine situations where theories correctly specify sufficient elements of $\boldsymbol{\beta}_1' + \lambda\boldsymbol{\beta}_2'$ to be zero, such that (5) is identifiable on the rank condition in the bivariate process. Equation (5) may even be interpretable (e.g., as a demand equation when $\alpha - \lambda\phi < 0$), but it obviously does not correspond to the structure. As argued in Hendry and Mizon (1993), such 'spurious' structures can be detected using structural breaks induced by natural experiments and policy changes. Both (Elmer) Working and Frisch understood this problem, given their concept of identification: see Working (1927) and Frisch (1938). The breaks are not needed to 'uniquely identify' (5) in its bivariate system by the rank condition; they are needed to determine if it is a structure at all.

Surprisingly, Faust and Whiteman seem to take it for granted in their identification discussion that the exogeneity status of their $\mathbf{Z}_t$ variable is certain. I now consider that issue.

5 Weak exogeneity

Their discussion of this topic suggests a failure to understand either the logic of conditioning or the associated concept of weak exogeneity. Indeed, their discussion of weak exogeneity is so riddled with errors and incorrect assertions that this section of the 'translation' has transmogrified into a new chapter of fiction. One is tempted to adopt Keynes (1936) famous quote from Henrik Ibsen's *Wild Duck*: "The wild duck has dived down to the bottom ... and bitten fast hold of the weed and tangle... and it would need an extraordinarily clever dog to dive after and fish her up again." (p183, 1961 edition). However, the importance of the issue enforces a clarification.

Weak exogeneity concerns inference with no loss of information in conditional models. There are no generic implications for "estimation methods", or "estimation efficiency". A failure of weak exogeneity may be of little relevance in some settings because the resulting loss of information is trivial, but in others can induce inconsistency, as in least-squares estimation of a single equation with endogenous regressors. Examples spelling out these points are provided in Engle, Hendry and Richard (1983).

Their misunderstandings of weak exogeneity may stem from their confounding the existence of parametric dependence across conditional and marginal models, with the imposition (or not) of the resulting restrictions. A violation of weak exogeneity can go far beyond the 'estimation inefficiency' of not imposing valid restrictions. This is because two conditions are needed for the weak exogeneity of a conditioning variable $\mathbf{z}_t$ for a parameter of interest $\boldsymbol{\mu}$. The first is that $\boldsymbol{\mu}$ can be obtained from the conditional model alone; the second is that the parameters of the conditional and marginal models are variation free. The latter usually requires the absence of cross-equation restrictions, but the absence of cross-equation restrictions does not necessarily entail weak exogeneity. For example, the parameter of interest may not be derivable from the conditional model even when all restrictions are used. Ignoring the *existence* of some parameter dependencies can produce conditional models with non-constant coefficients which do not correspond to the parameters of interest in agent behavior: there are many illustrations of this in Hendry (1995a), from which the following example is adapted.

Consider the static bivariate normal distribution:

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} \sim \mathsf{IN}_2 \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \right]. \tag{7}$$

The five parameters of (7) are denoted by $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^5$ where $\boldsymbol{\theta} = (\mu_1 : \mu_2 : \sigma_{11} : \sigma_{12} : \sigma_{22})'$: all the elements $\boldsymbol{\theta}$ are assumed to be variation free, beyond $\boldsymbol{\Sigma} = (\sigma_{ij})$ being positive definite. Both variables are endogenous in the context of (7), but that does not preclude conditioning. Let the economic theory model entail a linear relation between the planned outcome and the expected value of z_t, so:

$$\mu_1 = \beta\mu_2, \tag{8}$$

where β is the parameter of interest and $\mu_2 = \mathsf{E}[z_t] \neq 0$. From (7) and (8):

$$\mathsf{E}\left[y_t \mid z_t\right] = \mu_1 + \sigma_{12}\sigma_{22}^{-1}\left(z_t - \mu_2\right) = \left(\beta - \delta\right)\mu_2 + \delta z_t, \tag{9}$$

where $\delta = \sigma_{12}\sigma_{22}^{-1}$. The conditional variance is $\omega^2 = \sigma_{11} - \delta\sigma_{21}$. Thus, the parameters of the conditional and marginal densities are:

$$\left(\left[\beta - \delta\right]\mu_2 : \delta : \omega^2\right)' \text{ and } \left(\mu_2 : \sigma_{22}\right)'. \tag{10}$$

This factorization involves five parameters, so there are no within, or across, equation restrictions, and no implied restrictions on (7). Nevertheless, z_t is not weakly exogenous for β in (9) unless $\beta = \delta$; indeed, β cannot be learned from that conditional model alone when $\beta \neq \delta$. However, β can be learned from the joint density (7) using (8) since $\beta = \mu_1/\mu_2$. In their 'translation' (and any original), least squares applied to (9) is inconsistent for β.

Under weak exogeneity:

$$\mathsf{E}\left[y_t \mid z_t\right] = \beta z_t, \tag{11}$$

requiring:

$$\sigma_{12} = \beta\sigma_{22}. \tag{12}$$

This population covariance restriction may, or may not, hold depending on agent behavior. If it does hold, there is no loss from conditional inference – and the resulting equation is uniquely 'identified'.

Next I consider non-constancy. If weak exogeneity does not hold at all points in time, and the parameters (μ_2, σ_{22}) change due to policy intervention, then so would the parameters in the conditional model (9). This example shows the importance of the violation of weak exogeneity, even when there are no actual 'cross-equation restrictions' on the parameters of the conditional model (although there are mappings that let β be disentangled from $\left(\beta - \delta\right)\mu_2$ using the parameters of the joint distribution (7)). Further, consider redefining $\left(\beta - \delta\right)\mu_2$ as a new parameter ψ, and claiming the parameters of interest are $\left(\psi : \delta : \omega^2\right)$, so that z_t becomes weakly exogenous for $\left(\psi : \delta : \omega^2\right)$ by construction. This would not prevent predictive failure resulting when (μ_2, σ_{22}) changed. This class of examples is the reason I focus on super exogeneity (i.e., weak exogeneity plus invariance) as more fundamental. Since these are issues at the level of the process, not empirical modeling, words such as 'valid' are well defined, *pace* Faust and Whiteman.

Let us turn now to "covariance restrictions". Written as a model, (11) becomes:

$$y_t = \beta z_t + v_t \text{ where } \mathsf{E}\left[v_t \mid z_t\right] = 0. \tag{13}$$

Consequently, writing (7) as:

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix},$$

then $\mathsf{E}\left[v_t\epsilon_{2,t}\right] = 0$: hence my surprise at their remarks about "frowning" on covariance restrictions, and their ignoring such restrictions when discussing identification. In their 'translation', equation (13) is identified by such a restriction here, as is the general conditional equation (9) even when the zero intercept is not imposed. The usual problem of covariance restrictions being invariant only for orthogonal rotations does not arise here, as the factorizations are always conditional/marginal. It is one thing to dislike this approach; but that hardly justifies a long discussion claiming a lack of identification that does not seriously address the approach used: this 'translation' is a highly expurgated rendering.

A further issue of substance that Faust and White misunderstand is the relation between weak exogeneity and simultaneity. The existence of simultaneity does *not* preclude weak exogeneity in a conditional model. Consider the following bivariate DGP for the vector $\mathbf{x}_t = (y_t : z_t)'$:

$$\begin{aligned} y_t &= \beta z_t + \epsilon_{1,t} \\ \Delta z_t &= \rho \left(y_{t-1} - \beta z_{t-1} \right) + \epsilon_{2,t} \end{aligned} \tag{14}$$

when $\Delta z_t = z_t - z_{t-1}$, and:

$$\left(\begin{array}{c} \epsilon_{1,t} \\ \epsilon_{2,t} \end{array} \right) \sim \mathsf{IN}_2 \left[\left(\begin{array}{c} 0 \\ 0 \end{array} \right), \left(\begin{array}{cc} \sigma_1^2 & \gamma\sigma_1\sigma_2 \\ \gamma\sigma_1\sigma_2 & \sigma_2^2 \end{array} \right) \right] = \mathsf{IN}_2 \left[\mathbf{0}, \boldsymbol{\Sigma} \right]. \tag{15}$$

This is an integrated-cointegrated process (see e.g., Engle and Granger, 1987, Johansen, 1988, Phillips, 1991, Banerjee, Dolado, Galbraith and Hendry, 1993, and Hamilton, 1994), with the dynamics of the second equation explicitly specified as in Hendry (1995c). The first equation is to be modeled, with the parameter β of interest. Let $\mathcal{I}_{t-1}$ denote available lagged information, then from (15) and (14), normalizing $\sigma_1 = \sigma_2 = 1$, the conditional expectation of y_t given $(z_t, \mathcal{I}_{t-1})$ is:

$$\mathsf{E}\left[y_t \mid z_t, \mathcal{I}_{t-1} \right] = \beta z_t + \gamma \Delta z_t - \gamma\rho \left(y_{t-1} - \beta z_{t-1} \right). \tag{16}$$

When $\rho = 0$:

$$\mathsf{E}\left[y_t \mid z_t, \mathcal{I}_{t-1} \right] = \beta z_t + \gamma \Delta z_t. \tag{17}$$

from which β can be learned without loss, and the parameters of the conditional and marginal processes are variation free. Thus, despite simultaneous shocks when $\gamma \neq 0$ (i.e., $\mathsf{E}[z_t \epsilon_{1,t}] \neq 0$), z_t is weakly exogenous for β.

Conversely, even if $\gamma = 0$, so:

$$\mathsf{E}\left[y_t \mid z_t, \mathcal{I}_{t-1} \right] = \beta z_t, \tag{18}$$

and hence the conditional equation is both correctly specified and coincides with the equation determining agent behavior, nevertheless when $\rho \neq 0$, z_t is not weakly exogenous for β. The consequence in this case is that inference, such as tests of hypotheses about β, can be badly distorted, for example, rejecting the correct hypothesis 50% of the time at the 5% level.

Thus, this section refutes the claim by Faust and Whiteman that the 'LSE' procedure is hanging by its own bootstraps, or is a smoke ring suspended in thin air. The associated models are identified; other restrictions are testable from such identified models; structural breaks play a key role in revealing non-structural representations; and weak exogeneity sustains inference with no loss of relevant information, a state of nature not precluded by simultaneity.

6 Theory of reduction

The theory of reduction is the centerpiece of Hendry (1995a). It investigates what happens statistically as information is eliminated from the vast complexity of reality to obtain empirical models. As such, the theory operates at the level of the population, as opposed to the sample. For example, it corresponds to a study of regression theory, when the data generation process is given; not to the application of least squares to a sample where the generating process is unknown.

This theory offers an explanation for the origin of all empirical models, including any that might be espoused by Faust and Whiteman, 'RBCers', or 'VAR' builders. Nowhere do they so much as note this essential point: rather, as remarked above, they conflate population and sample throughout their analysis.

One value-added aspect of the theory of reduction is that almost all the fundamental concepts of econometrics arise as corresponding to no loss of information from particular reductions. For example, Granger non-causality (Granger, 1969) is needed to justify marginalizing with respect to lagged information on variables to be excluded from the analysis. Thus, this theory delivers the comprehensive taxonomy of evaluation information, the null hypotheses of which are the most-frequently tested aspects of empirical econometrics. Since the theory of reduction is a logical analysis of population properties of procedures, the notion of 'validity' is well defined. Tests of the validity of reductions apply here, in exactly the same sense as one would test the validity of assuming serially uncorrelated errors in a Neyman–Pearson framework: the term validity qualifies the status of the hypothesis in the theory, not the outcome in practice since there is always a non-zero null rejection frequency of tests in finite samples.

Hendry (1995a) is clear on the distinction between different levels of econometric analysis, namely what applies when investigating the logical status of methods, as in probability theory; how these methods work when the model is correctly specified, as in estimation theory; and what can be said when the model does not coincide with the mechanism, but needs to be discovered from theory and evidence, as in modeling theory.

This sample/population confusion is disastrously inimical to any claimed 'translation': it is akin to ignoring the gender distinction when translating English into French. For example, general-to-specific modeling is one implementation of the theory of reduction in an empirical setting. It does not have the strong basis of the latter, and while it appears to perform well in practice (see Hoover and Perez, 1996), and avoids important research inefficiencies, it is not immutable.

6.1 General-to-specific modeling

The distinction between the initially-postulated general model and the final simplified model is also lost in their translation, adding to the confusion. The former is subject to mis-specification tests in most empirical approaches. If these tests reject, then there is no more point in continuing than there would be in an RBC approach where the 'key moments' did not match at all, or a VAR approach with substantive residual autocorrelation and non-constant impulse responses. Outright failure *ab initio* is bound to entail revision of model, data, or method in any approach that seeks to characterize empirical evidence. When used once on a virgin data set, with the outcome reported irrespective of accept/reject, such tests have a Neyman–Pearson justification. Otherwise, in *all* approaches they become indices of congruency, since the design of the empirical model is a function of the 'test' information: this problem is not unique to the 'LSE' approach as Faust and Whiteman seem to imply.

Some of the tests are for potential mis-specifications, usually directed to checking if the initial errors have 'well-behaved' properties. In part, this is to ensure that specification tests have their anticipated distributions, but a more important reason, which Faust and Whiteman do not mention, is that otherwise, a multitude of sins can be camouflaged. For example, functional-form mistakes will often manifest themselves as residual heteroskedasticity, so it is a *non sequitur* to interpret the latter as deriving from error heteroskedasticity. Similarly, dynamic mis-specification will result in serially-correlated errors, and if the latter are found, the former cannot be excluded. On the other hand, when the dynamics entail a serially-independent innovation, that does not rule out an autoregressive-error interpretation, providing the appropriate common factors exist in the dynamic representation (see Sargan, 1980, and Hendry and Mizon, 1978). As noted above, this idea has applications to testing RBC claims about the behavior of postulated technology shocks, so is far from specific to the 'LSE' approach (for once, this idea is explicitly associated with the LSE, albeit that Sargan conceived of it prior to going there! See Sargan, 1964). When the initial tests do not reject, later reuse checks that model simplification is acceptable, so they become diagnostic indices. Certainly, the vast gulf in our views about the role of testing is well exemplified by one of the best jokes in the paper, namely, their idea (under 'Dictate 4') that estimates of a claimed parameter of unity being 0.5 in one period and 2 in another are 'less supportive' of the theory than a constant value of unity. When the standard errors of the two estimated parameters exceed 10 (say), such a result would tell us nothing beyond the hopeless extent of the uncertainty; but when standard errors are less than 0.01, the associated theory would seem worthless, and may even be dangerous if the sample average is unity.

There are three main reasons for simplifying the initial model rather than stopping. First, to allow interpretations of the parameters in the light of theoretical analyses: see Hendry, Leamer and Poirier (1990). Secondly, a model that 'mops up' data variation by a profligate parameterization may be difficult to reject despite considerable mis-specification, and simplification can increase the power to detect this. Finally, and closely related, the Scylla and Charybdis of modern time-series econometrics are theory-dependence and sample-dependence: when the former is excessive, advances in theory can lead to the discarding of hard-won econometric evidence, and when the latter dominates, extensions to the sample can induce model change, again vitiating what was previously believed. The initial general model seeks to mitigate theory becoming a strait-jacket, rather than a productive companion; and simplification attempts to reduce excessive sample dependence.

The value of any empirical study depends almost entirely on the goodness of the model's specification as a representation of what is being modeled, rather than on the modeling approach. A bad start will not become a good end just because a clever simplification strategy is used, even if a good start may be wrecked by poor modeling. In their study of the usefulness of general-to-simple strategies when the generating process is a special case of the initial model, Hoover and Perez (1996) find a simplification procedure based on encompassing works well, and helps determine a choice of model that is generally close to the generating mechanism, despite a proliferation of competing 'explanatory' variables. Further, there are increasingly useful developments underpinning data-based modeling: see e.g., White (1997). The converse, of employing 'theory-based restrictions, even if not strictly correct', depends on how wrong the restrictions are for the purpose in hand. But only empirical evidence can tell us that, which takes us full circle to the need to test the assertions in question, and in turn entails checking if the basis on which they are tested is valid. I do not see in the Faust–Whiteman approach how they can tell what inferences are reliable and what unfounded, without relying on the belief that their theory models are correct before confronting data evidence. Such assumed omniscience has not proved to be a sound basis in any discipline historically, and bids fair to return us to scholasticism — a time when most effort was devoted to 'translations and critiques' rather than advances.

7 Policy analysis

For econometric analyses to deliver sustainable and useful policy advice requires:

(a) causality, so the instrument variables affect the target outcomes;
(b) constancy, so the policy claims are not wrecked by concomitant changes in agent behavior;
(c) invariance, so the policy itself does not alter behavior (i.e., the critique in Lucas, 1976);
(d) an accurate model, so that advice is not radically altered by extending the information set.

I am ‘deeply suspicious’ of any policy analyses based on theoretical models that lack strong empirical corroboration. There are four reasons for such suspicion. First, theory-based causal links may not hold in reality, so that the entailed policy conclusions are misleading. To make viable policy statements, the targets and instruments must be causally linked in practice: simply asserting they are in a theory model, and including them with significant coefficients in empirical models, is not adequate — see the critique in Granger and Deutsch (1992). Secondly, the assumptions in the theory about the constancy of the model’s parameters may not hold empirically, so the magnitudes of the policy claims are incorrect: remember the Faust–Whiteman example of treating as unity a parameter that is the average of 0.5 and 2. Next, the variables in the theory may not be a complete specification of the system, and unmodeled interdependencies may vitiate the policy claims. Only careful empirical evaluation can ascertain the validity or otherwise of such claims. Finally, there are invariably rival policy views, and when these are implemented in models, encompassing can be used to evaluate their mutual claims: see Mizon and Richard (1986) and Gourieroux and Monfort (1994). Absent such evaluation, or failure thereon, and the policy debate is reduced to assertion.

7.1 ‘LSE’ policy conclusions

As just emphasized, the causal issue must come first, and cannot be determined by theory alone. Faust and Whiteman – reread your claims about policy advice *when in fact there is no causal link*. Absent any actual link, and all the statements are unsustainable. The first objective of the ‘LSE’ approach has been to develop methods of ascertaining whether such links exist. There is no claim that it is easy to do so; merely that, without an empirical basis, policy assertions are a hazardous basis on which society might proceed, however much a subgroup may believe in their theoretical views. Perhaps the approach strives too much to avoid the syndrome of “fools rush in where angels fear to tread”, but the converse of “act in haste and repent at leisure” is not very appetizing.

It may surprise Faust and Whiteman to learn that the form of model which they so dislike, and claim to be useless for policy, is in fact used by most European policy agencies, such as Central Banks, Treasuries, etc. The address by Anders–Møeller Christiansen of the Danish Central Bank to the Copenhagen University City-of-Culture Conference documented that claim for European Central Banks’ studies of money demand. Moreover, from the mid-1970s onwards, many macro-econometric systems switched to developing equations using general-to-specific approaches embodying ‘error-correction’ mechanisms (see, among many others, H.M. Treasury, 1980). Indeed, as Faust is at the ‘Fed’, he is in an excellent position to compare their macro-econometric model today with that of the early 1980s: does he really claim that the model has not been substantively influenced by the ‘LSE’ approach?

Once causal links are (tentatively) established, various policy implications do follow. Certainly, the relevant equations may need to be embedded in the system-wide context for many policy purposes. Nevertheless, some implications were drawn in the money-demand examples Faust and Whiteman select. First, in Hendry (1985), from the worse empirical fit of tax-adjusted interest rates compared to unadjusted (consistent with tax evasion on income paid gross of tax), I concluded that when interest was forced by law to be paid net, own rates on money would jump close to competitive rates. As Figure 1 records from the time series on outside and own rates over 1980–89, this is precisely what occurred with the introduction of interest-bearing sight deposits, paid post-tax, in 1984. The resulting opportunity cost falls sharply, as the second panel shows.

Now Faust and Whiteman dismiss this work as “uncontroversial” and by innuendo, not as convincing as other evidence. The theoretical implications of a near-zero opportunity cost of holding money are spelt out in Sprenkle and Miller (1980), namely an almost unbounded precautionary demand. Thus,

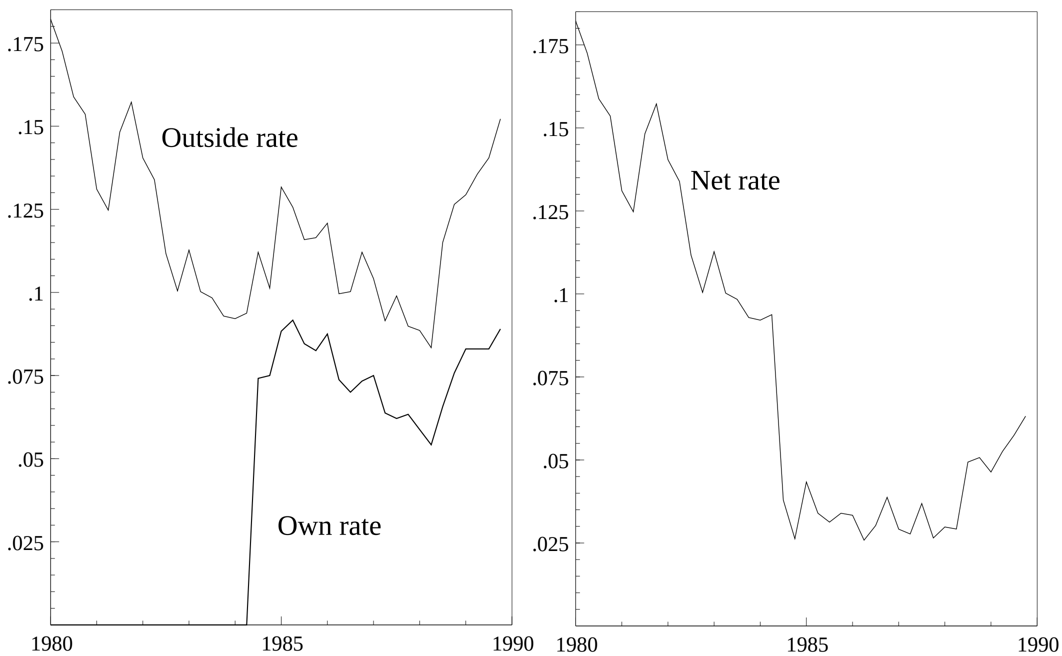

Figure 1 Competitive and own interest rates in the UK.

the consequent leap in money holdings can be seen after the event to be a portfolio adjustment, not a precursor to inflation: to see how "uncontroversial" that implication is, perhaps they should read the evidence summarized in Hendry (1981). Moreover, in the face of a monetary explosion, it is crucial to understand its causes for interest rate policy, and hence economic growth or recession: compare the views on an earlier "monetary explosion" in Friedman (1983). Our analysis correctly assigns the UK case to portfolio adjustment, and therefore shows no need for a rise in interest rates. Since M1 grew at 15–20% p.a., but only depended on the interest-rate differential, there was plenty of scope for serious policy errors in trying to 'control' it by raising the level of interest rates.

Their second and third implications concern the direction of causation between money and inflation, an issue which is central to the meaning of monetary policy. If money is endogenously determined by the private sector in response to the income and price levels and opportunity costs which agents face, it is unclear what attempts to change the money stock hope to achieve. This answers some of Faust and Whiteman's later questions, albeit not in a way they may like: absent causation of the form desired by a policy maker seeking to manipulate the economy, and the outcome will not be what is anticipated. Of course, such models cannot answer what will happen to unemployment when that variable was not in the model under analysis; but embedding the relevant equations in a system would allow quantitative scenario studies, with some expectation that the results would not be distorted by invariance failures.

Figure 2 shows US data on changes in the 1-month T-bill rate, highlighting the period during the New Operating Procedures. Since the interest-rate coefficient in the Baba, Hendry and Starr (1992) model estimated prior to 1979, is unaffected by such large changes, and induced no ARCH effects in the residuals, what explains its constancy if the response is not from the structure of agent behavior? It is difficult to understand such invariance when it is asserted not to reflect agents' responses. Later evidence in Hess, Jones and Porter (1997) shows the model did not remain constant through the 1990s, so it remains too simplistic to capture the relevant behavior, since collinearity, model mis-specification, and a lack of parsimony *per se* cannot explain predictive failure — breaks somewhere are needed: see Hendry and Doornik (1997). As an analogy, if a spacecraft is hit by a meteor, and knocked off course, previous forecasts from NASA would be falsified, and a modification required. This outcome is only

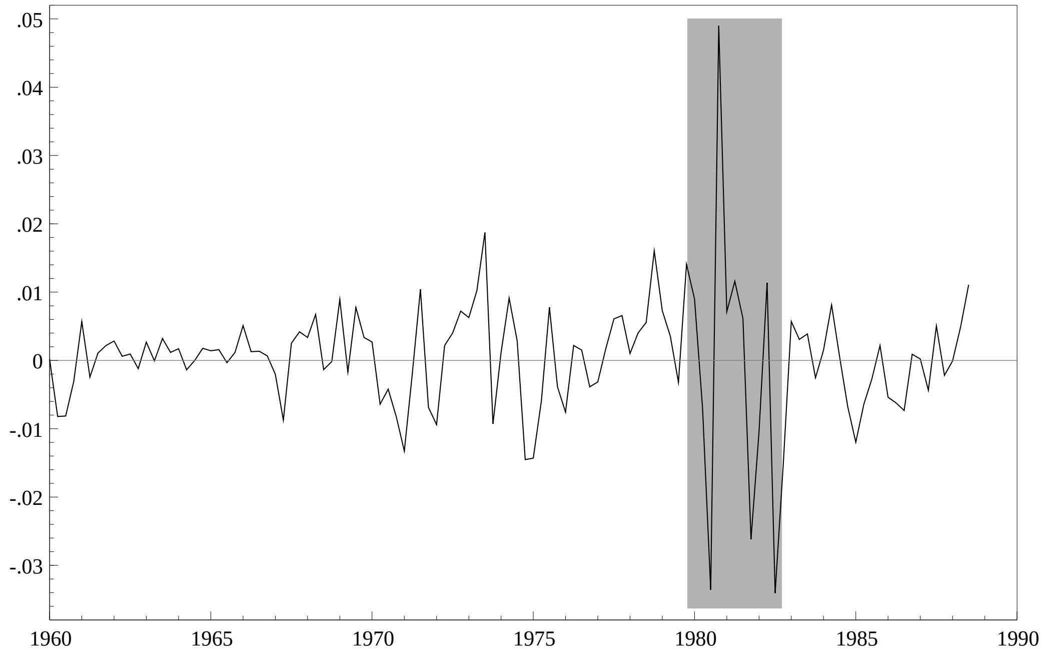

Figure 2 Time series of changes in US 1-month bill interest rate.

dependent on the forecast-period event, not on the quality of NASA's forecasting algorithm, although that could be criticized for failing to allow for accidents. Nevertheless, theories of gravitation would not be refuted by the forecast failure, nor would methods of modeling in engineering or astrophysics (except perhaps to robustify future spacecraft against such shocks if the costs do not prove prohibitive).

8 Rhetoric and substance

The multiple redundancies in their title alert the reader to the implicit views of Faust and Whiteman: I take their humor to be an attempt to enliven what can be a tedious topic, rather than the scoundrel's refuge of seeking to demean what they dislike by tricks of language when lacking substance. Nevertheless, their use of language to prejudice reader's attitudes is regrettable. For example, the claim that the 'LSE' approach is essentially about '*single*' equations (their first paragraph, emphasis added), in contrast to RBC, VAR, and large-scale econometric systems, is unreasonable. Prior to the mid-1980s, in order to understand the main issues in empirical modeling of economic time series, the 'LSE' approach sought to isolate problems in the simplest context of one equation. Once the key concepts, tools, and approaches were developed, the analysis was explicitly extended to systems (see e.g., Hendry, 1986, and Hendry *et al.*, 1988). Since Faust and Whiteman reference several system-based analyses, their focus on pre-system methods is not based on lack of knowledge.

When the 'LSE' approach uses a VAR as its initial model, but seeks to focus on a subset of equations deemed to be of interest, they call the remaining equations 'auxiliary, cursory'. This apparent criticism applies even when the whole system of which they are part has been rigorously evaluated: so do all VAR equations have that status?

Cointegration restrictions are called 'purely statistical': this again will not do. The theory of cointegration evolved from modeling long-run relations derived from economic theory: see for example, the optimization model of the behavior of UK Building Societies (akin to Savings and Loans Associations) in Hendry and Anderson (1977). Clive Granger documents its later evolution in his recent *Econometric Theory* Interview. Worse still are comments with implicit innuendos such as "What is remarkable about the insistence on weak exogeneity in this (the cointegration) context is that in general the estimation of the cointegrating relationship in isolation (that is, inefficiently) still results in a superconsistent estimate ...” as if that meant anything. Presumably the reader is invited to imagine that superconsistency is sufficient for “good” estimation in some sense. Yet Faust and Whiteman are well aware of the simulation results for superconsistent estimators in Banerjee, Dolado, Hendry and Smith (1986) showing small-sample biases in excess of 50% of the parameter to be estimated when weak exogeneity is violated.

8.1 Citations

Citation studies can be a useful bibliographic source, but need to be carefully undertaken as the choice of which papers to investigate is central to a successful appraisal. Unfortunately, Faust and Whiteman fail to do so, perhaps because of their tight budget constraint. Since ‘Hendry (1985)’ and ‘Hendry (1991)’ in Table 1 are not listed in their references, I am unsure what papers are intended. Using Sims (1986) to represent his work selects a paper that is not in a mainstream journal. To see the extent of the misrepresentation, contrast the citations for the first major ‘LSE’ paper applying the evolving modeling methodology, namely Davidson, Hendry, Srba and Yeo (1978), and the first by Sims (1980) emphasizing VAR modeling: there are gross citations of more than 400 and almost 800 respectively. “Some would find this fact to be of central relevance to the section”. I suspect these last two papers have also seen much more cross-referencing than those Faust and Whiteman select.

9 Conclusion

As a ‘translation’, the paper by Faust and Whiteman falls into the genre in which ‘the spirit is willing, but the flesh is weak’ in one language comes out as ‘the vodka is strong, but the meat is rotten’ in another. The difficulty is not so much that ‘translations’ are best done from a complete command of both languages, which few possess, but that errors in translated documents remain glaring to native-language speakers. Stephen Jay Gould (1991) provides several powerful examples of how a misleading view of a topic can emerge from secondary sources alone being consulted; interested readers would do well to read some of the originals. However, many of Faust and Whiteman’s substantive errors and claimed criticisms have been corrected above.

‘LSE’ models are identified in the Cowles’ sense of uniqueness; restrictions are tested given identification; structural breaks are used to reveal non-structural, but identified representations; and weak exogeneity sustains inference without loss of relevant information on the parameters of interest. The initial general model is subject to mis-specification testing, and simplifications are used to enhance interpretability and increase power to detect mis-specification, as well as reduce sample dependence. Such methods are applicable to almost all initial formulations, and examples were offered for its application to RBC and VAR approaches.

As a critique, the Faust and Whiteman paper lacks substance, is incorrect in important aspects, confounds different concepts and modes of analysis, and offers no viable or coherent alternative. Certainly, one can make telling points in a nihilistic assault, but the authors clearly wish to make policy pronouncements at the end of their analyses, so incoherence is a serious problem for them. They seem not to realize that the implications of the approach they seek to evaluate apply to other methods of empirical modeling in economics. Indeed, they have fallen for the famous Guinness advert – they have not tried it because they don’t like it. I can but reiterate their conclusion that: “Finding improved ways to exploit economic theory while thoroughly examining the empirical properties of the models would surely be beneficial”. However, they may find that I parse it differently.

References

Anderson, T. W. (1971). *The Statistical Analysis of Time Series.* New York: John Wiley & Sons.

Baba, Y., Hendry, D. F., and Starr, R. M. (1992). The demand for M1 in the U.S.A., 1960–1988. *Review of Economic Studies*, **59**, 25–61.

Banerjee, A., Dolado, J. J., Galbraith, J. W., and Hendry, D. F. (1993). *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.

Banerjee, A., Dolado, J. J., Hendry, D. F., and Smith, G. W. (1986). Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence. *Oxford Bulletin of Economics and Statistics*, **48**, 253–277.

Bårdsen, G., and Fisher, P. G. (1993). The importance of being structured. Discussion paper, Norwegian School of Economics, Bergen.

Cross, R. (1982). The Duhem-Quine thesis, Lakatos and the appraisal of theories in macro-economics. *Economic Journal*, **92**, 320–340.

Davidson, J. E. H., Hendry, D. F., Srba, F., and Yeo, J. S. (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *Economic Journal*, **88**, 661–692. Reprinted in Hendry, D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Engle, R. F., and Granger, C. W. J. (1987). Cointegration and error correction: Representation, estimation and testing. *Econometrica*, **55**, 251–276.

Engle, R. F., Hendry, D. F., and Richard, J.-F. (1983). Exogeneity. *Econometrica*, **51**, 277–304. Reprinted in Hendry, D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Faust, J., and Whiteman, C. H. (1997). General-to-specific procedures for fitting a data-admissible, theory-inspired, congruent, parsimonious, encompassing, weakly-exogenous, identified, structural model of the DGP: A translation and critique. mimeo, Federal Reserve Board, Washington. Forthcoming, Carnegie–Rochester Conference Series on Public Policy.

Fisher, F. M. (1966). *The Identification Problem in Econometrics*. New York: McGraw Hill.

Florens, J.-P., Mouchart, M., and Rolin, J.-M. (1990). *Elements of Bayesian Statistics.* New York: Marcel Dekker.

Friedman, M. (1983). A case of bad good news. *Newsweek*, **September 26**, 84.

Frisch, R. (1933). Editorial. *Econometrica*, **1**, 1–4.

Frisch, R. (1934). *Statistical Confluence Analysis by means of Complete Regression Systems*. Oslo: University Institute of Economics.

Frisch, R. (1938). Statistical versus theoretical relations in economic macrodynamics. Mimeograph dated 17 July 1938, League of Nations Memorandum. Reproduced by University of Oslo in 1948 with Tinbergen's comments. Contained in Memorandum 'Autonomy of Economic Relations', 6 November 1948, Oslo, Universitets Økonomiske Institutt. Reprinted in Hendry D. F. and Morgan M. S. (1995), *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press.

Gould, S. J. (1991). *Bully for Brontosaurus*. London: Penguin.

Gourieroux, C., and Monfort, A. (1994). Testing non-nested hypotheses. In Engle, R. F., and McFadden, D. L. (eds.), *Handbook of Econometrics*, pp. 2583–2637. Amsterdam: North–Holland.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, **37**, 424–438.

Granger, C. W. J., and Deutsch, M. (1992). Comments on the evaluation of policy models. *Journal of Policy Modeling*, **14**, 497–516.

Haavelmo, T. (1944). The probability approach in econometrics. *Econometrica*, **12**, 1–118. Supplement.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Hausman, J. A. (1983). Specification and estimation of simultaneous equation models. In Griliches, Z., and Intriligator, M. D. (eds.), *Handbook of Econometrics*, Vol. 1, pp. 391–448. Amsterdam: North-Holland.

Hendry, D. F. (1981). Econometric evidence in the appraisal of UK monetary policy. In *The Third Report of the Select Committee of the House of Commons on the Treasury and Civil Service*, Vol. 3, pp. 1–21: HMSO.

Hendry, D. F. (1985). Monetary economic myth and econometric reality. *Oxford Review of Economic Policy*, **1**, 72–84. Reprinted in Hendry, D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Hendry, D. F. (1986). Empirical modelling in dynamic econometrics: The new-construction sector. *Applied Mathematics and Computation*, **21**, 1–36.

Hendry, D. F. (1995a). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F. (1995b). Econometrics and business cycle empirics. *Economic Journal*, **105**, 1622–1636.

Hendry, D. F. (1995c). On the interactions of unit roots and exogeneity. *Econometric Reviews*, **14**, 383–419.

Hendry, D. F., and Anderson, G. J. (1977). Testing dynamic specification in small simultaneous systems: An application to a model of building society behaviour in the United Kingdom. In Intriligator, M. D. (ed.), *Frontiers in Quantitative Economics*, Vol. 3, pp. 361–383. Amsterdam: North Holland Publishing Company. Reprinted in Hendry, D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Hendry, D. F., and Doornik, J. A. (1997). The implications for econometric modelling of forecast failure. *Scottish Journal of Political Economy*. forthcoming, Centenary Issue.

Hendry, D. F., Leamer, E. E., and Poirier, D. J. (1990). A conversation on econometric methodology. *Econometric Theory*, **6**, 171–261.

Hendry, D. F., and Mizon, G. E. (1978). Serial correlation as a convenient simplification, not a nuisance: A comment on a study of the demand for money by the Bank of England. *Economic Journal*, **88**, 549–563. Reprinted in Hendry, D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Hendry, D. F., and Mizon, G. E. (1993). Evaluating dynamic econometric models by encompassing the VAR. In Phillips, P. C. B. (ed.), *Models, Methods and Applications of Econometrics*, pp. 272–300. Oxford: Basil Blackwell.

Hendry, D. F., and Morgan, M. S. (1995). *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press.

Hendry, D. F., Neale, A. J., and Srba, F. (1988). Econometric analysis of small linear systems using PC-FIML. *Journal of Econometrics*, **38**, 203–226.

Hess, G. D., Jones, C. S., and Porter, R. D. (1997). The predictive failure of the Baba, Hendry and Starr model of M1. forthcoming, Journal of Economics and Business.

Hoover, K. D., and Perez, S. J. (1996). Data mining reconsidered: Encompassing and the general-to-specific approach to specification search. Mimeo, Economics department, University of California, Davis.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231–254.

Juselius, K. (1993). VAR modelling and Haavelmo's probability approach to econometrics. *Empirical Economics*, **18**, 595–622.

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. London: Macmillan.

Koopmans, T. C. (ed.)(1950). *Statistical Inference in Dynamic Economic Models*. No. 10 in Cowles Commission Monograph. New York: John Wiley & Sons.

Koopmans, T. C., Rubin, H., and Leipnik, R. B. (1950). Measuring the equation systems of dynamic economics. in Koopmans (1950), Ch. 2.

Kydland, F. E., and Prescott, E. C. (1991). The econometrics of the general equilibrium approach to business cycles. *Scandanavian Journal of Economics*, **93**, 161–178.

Lucas, R. E. (1976). Econometric policy evaluation: A critique. In Brunner, K., and Meltzer, A. (eds.), *The Phillips Curve and Labor Markets*, Vol. 1 of *Carnegie-Rochester Conferences on Public Policy*, pp. 19–46. Amsterdam: North-Holland Publishing Company.

Mizon, G. E. (1995). A simple message for autocorrelation correctors: Don't. *Journal of Econometrics*, **69**, 267–288.

Mizon, G. E., and Richard, J.-F. (1986). The encompassing principle and its application to non-nested hypothesis tests. *Econometrica*, **54**, 657–678.

Pagan, A. R. (1987). Three econometric methodologies: A critical appraisal. *Journal of Economic Surveys*, **1**, 3–24. Reprinted in Granger, C. W. J. (ed.) (1990), *Modelling Economic Series*. Oxford: Clarendon Press.

Phillips, P. C. B. (1991). Optimal inference in cointegrated systems. *Econometrica*, **59**, 283–306.

Phillips, P. C. B. (1997). The ET interview: Professor Clive Granger. *Econometric Theory*, **13**, 253–303.

Sargan, J. D. (1964). Wages and prices in the United Kingdom: A study in econometric methodology (with discussion). In Hart, P. E., Mills, G., and Whitaker, J. K. (eds.), *Econometric Analysis for National Economic Planning*, Vol. 16 of *Colston Papers*, pp. 25–63. London: Butterworth Co. Reprinted as pp. 275–314 in Hendry D. F. and Wallis K. F. (eds.) (1984). *Econometrics and Quantitative Economics*. Oxford: Basil Blackwell, and as pp. 124–169 in Sargan J. D. (1988), *Contributions to Econometrics*, Vol. 1, Cambridge: Cambridge University Press.

Sargan, J. D. (1980). Some tests of dynamic specification for a single equation. *Econometrica*, **48**, 879–897. Reprinted as pp. 191–212 in Sargan J. D. (1988), *Contributions to Econometrics*, Vol. 1, Cambridge: Cambridge University Press.

Schumpeter, J. (1933). The common sense of econometrics. *Econometrica*, **1**, 5–12.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, **48**, 1–48. Reprinted in Granger, C. W. J. (ed.) (1990), *Modelling Economic Series*. Oxford: Clarendon Press.

Sims, C. A. (1986). Are forecasting models useable for policy analysis?. *Federal Reserve Bank of Minneapolis Quarterly Review*, **Winter**, 2–16.

Sprenkle, C., and Miller, M. (1980). The precautionary demand for narrow and broad money. *Economica*, **47**, 407–421.

Treasury, H. (1980). *Macroeconomic Model Technical Manual*. London: H.M. Treasury.

White, H. (1997). Data snooping with confidence. Mimeo, Economics Department, University of California at San Diego.

Wold, H. O. A., and Juréen, L. (1953). *Demand Analysis: A Study in Econometrics*, 2nd edn. New York: John Wiley.

Working, E. J. (1927). What do statistical demand curves show?. *Quarterly Journal of Economics*, **41**, 212–235.

Phillips (1997)