Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 576

March 1997

GENERAL-TO-SPECIFIC PROCEDURES FOR FITTING A DATA-ADMISSIBLE, THEORY-INSPIRED, CONGRUENT, PARSIMONIOUS, ENCOMPASSING, WEAKLY-EXOGENOUS, IDENTIFIED, STRUCTURAL MODEL TO THE DGP: A TRANSLATION AND CRITIQUE

Jon Faust and Charles H. Whiteman

General-to-Specific Procedures for Fitting a Data-Admissible, Theory-Inspired, Congruent, Parsimonious, Encompassing, Weakly-Exogenous, Identified, Structural Model to the DGP: A Translation and Critique

Jon Faust and Charles H. Whiteman*

Abstract: We characterize the LSE approach by its implications for reduced-form modeling and structural interpretations. Much of what has come to be associated with the LSE methodology involves the approach to fitting reduced forms, and can be thought of as a pragmatic solution to problems created by short samples plagued by serial correlation. The policy analysis one might be able to do with an "LSE model," on the other hand, hinges on structural identification arguments which do not meet the classic Cowles Commission standards, and is thus suspect.

Keywords: reduced form, identification, instruments.

* Faust: International Finance Division, Board of Governors of the Federal Reserve System; Whiteman: The University of Iowa. This paper was prepared for the Carnegie-Rochester Conference Series on Public Policy, fall 1996. We thank Dave DeJong, Neil Ericsson, David Hendry, Beth Ingram, Ed Leamer, Mico Loretan, Ben McCallum, and Christopher Sims for helpful comments. Whiteman gratefully acknowledges the support provided for this research by The College of Business Administration at The University of Iowa, the Federal Reserve Bank of Atlanta, and the NSF under grant SBR 9422873. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System. the Federal Reserve Bank of Atlanta, or of any other person associated with the Federal Reserve System.

Introduction

The breakdown of the 1960s consensus regarding the theoretical foundation of macroeconomics was striking, but even more impressive are the deep rifts that have emerged over the proper way to tease empirical facts from macroeconomic data. Three prominent (and perhaps dominant) approaches are the real business cycle (RBC) approach pioneered by Prescott (1986), the identified vector autoregression (VAR) approach of Sims (1986), and what we will call the London School of Economics (LSE) approach (Hendry, 1995). A brief look at a paper in each of these traditions reveals that there is almost nothing in common in their ultimate empirical products. An RBC paper might report a table of key moments of data implied by a theoretical model together with the asserted empirical analogs; the VAR paper would present impulse response functions from a just-identified vector autoregression; and the LSE paper would present a single parsimonious equation along with a host of specification test statistics. Yet while these three schools dominate academic work, updated versions of the large-scale econometric models that were born in the 1960s remain prominent in policy institutions (see, for example, the papers by Stockton, Reifschneider, and Wilcox; and Brayton, Levin, Tryon, and Williams in this volume).

Calling on an old parable, the proponents of the empirical approaches are like blind seers, sent to form a description (model?) of an elephant. The seer holding the trunk imagines a snake, the one holding a leg imagines a tree, and so on. The macroeconomic seers are the LSE'ers (pronounced "el-seers"), the RBC'ers ("arby-seers"), and VARs'ers (less alliteratively, "vahr-seers"); each uses a distinct approach to examining the economy and each ends up describing a very different beast. Ultimately, the seers are called upon to make suggestions about how the beast could or should move.

Our charge in this paper is to explain and evaluate the prescriptions which might be seen by LSE'ers. Interpreted broadly, this would be unmanageable given the breadth and depth of the fundamental econometric contributions by LSE economists and statisticians such as James Durbin, Andrew Harvey, David Hendry, A.W. Phillips, Dennis Sargan. Instead, we focus more narrowly on what Mizon (1995) and others have called the "LSE methodology," what Gilbert

(1986) has called "Professor Hendry's econometric method," and what is laid out most thoroughly in Hendry's *Dynamic Econometrics* (1995). Our choice of the LSE label provides us with a parallel acronym to set against RBC and VAR, but surely has elements of misnomer:[1] many of the distinguishing features of the approach were developed after Hendry left the LSE for the greener meadows of Oxford, and many at the LSE probably do not subscribe to key elements of the Hendry approach. But aside from misnomer, these issues will be of limited importance as we will not attempt to determine the pedigree of the ideas espoused, and will instead emphasize the elements of the approach that distinguish it from standard practice in other schools, not the common elements that may have historical origins at the LSE.

The primary goal of the paper is to illuminate the strengths and weaknesses of the LSE approach to policy analysis. We begin by recording what policymakers might want and reviewing the pitfalls the macroeconometrician faces in attempting to deliver policy analysis. We next provide a rough summary of the LSE method as a response to those pitfalls, highlighting distinguishing features by contrasting the LSE solution with the other approaches to empirical macroeconomics. In this section and throughout, work on monetary policy issues forms the basis for comparison among the econometric camps. This is both convenient and sensible, since each camp focused on such issues over the past 10 or 15 years: RBC'ers attempted to bound the share of the variance of output due to nominal disturbances, VARs'ers also attempted to measure this share as well as the dynamic response of the economy to a money supply disturbance. At the same time, LSE'ers developed and refined their techniques studying money demand and the related questions of the exogeneity of money and stability of money demand.

After the summary, we explore the LSE approach in detail, providing our assessment of its merits as we proceed. In doing so, we characterize the most important principles underlying it as pertaining either to reduced-form data-fitting issues or to identification and structural issues. On the former, the LSE school has laid out a pragmatic approach to dealing with the problems of short samples of serially correlated data, and has been resolute in the defense of the importance of constant parameters, well-behaved error terms, and providing improved explanations for existing findings. The approach to identifying structure, in contrast, does not arise from pragmatism. We argue that it strays from the narrow path laid out by the Cowles Commission and does not, therefore, aid in learning about economic structure. Finally, we offer some perspective on whether the LSE approach can answer important policy questions, and how the approach in fact seems to be used throughout the world.

Macroeconometric Policy Analysis

Macroeconomists are called upon to render opinions on many questions of interest to policymakers. The spectrum of possibilities includes historical questions (what *was*...?), measurement questions (what *is* ...?), simulation questions (what *will be if* ...?), and policy questions (what *should be* ...?) Examples of historical questions are: Has the Fed caused recessions? When have monetary shocks occurred and what were their effects? Measurement questions are of the form: What is the interest elasticity of money demand? Simulation questions are: What would happen if the Bundesbank engineered a long-lasting change in the growth rate of the German money supply? What would happen if the U.S. government budget deficit followed a prescribed path? Most of these questions are in turn directed in one way or another toward policy questions: What path for the Fed funds rate should the Fed pursue? Should the funds rate be changed at the next FOMC meeting?

Providing useful answers to these questions requires an econometric model which matches *the salient features* of the data. *Salience* is, of course, viewed quite differently by different econometric camps. Further, answering many of the important policy questions requires determining the effects of altering the path of one variable in a simultaneous system. Such analysis requires going beyond fitting the data to confront a host of familiar issues involved in structural inference--e.g., the selection of identifying restrictions and simultaneity bias in estimation.

Although fitting the data and identifying the economic structure need not be entirely distinct steps either in theory or in practice, for purposes of discussion we will lump all issues into one of these two categories. Roughly speaking, we will treat any issues of exogeneity, simultaneity bias, and identifying restrictions as identification issues; all other issues of model specification and estimation will be treated as reduced-form issues.

The LSE method was explicitly designed as a response to the unique problems of macroeconometrics. What are they? Two of the progenitors of the approach remind us:

> [A]n economy is an inherently dynamic, stochastic, multidimensional, interdependent, non-linear and evolving entity of considerable complexity, the overall functioning of which is only partly understood...[D]ata are scarce and of uncertain relevance, experimentation is uncontrolled and available theories are highly abstract and rarely uncontroversial. (Hendry and Richard, 1983, p. 111-112)

The first problem the authors isolate is that there are a great many macroeconomic variables that are potentially of interest. In the context of a monetary model, there are half a dozen or more monetary measures--reserves, the monetary base, M1, M2, transactions portions of M2, and various adjusted versions of these series; and there are more potentially relevant interest rate measures than one can easily count. It is not clear which of these series should be the focus of analysis. The data series are highly correlated with a rich lead-lag structure. Many special events--the oil shocks of the 1970's and the imposition of credit controls in 1980--plague the data. Finally, samples are short relative to length of the typical business cycle. Together, these problems present the econometrician with nasty problems of what variables and which lags to include in the model. As we shall see below, many of the most dramatic and contentious differences among macroeconometric approaches come from different perspectives on this variable selection problem.

Settling the question of which features of the data to fit leaves one with the identification problem--if interest rates, income, prices, and the quantity of money are simultaneously determined in the economy, then one cannot infer the answer to most policy questions from the reduced form. Researchers of the Cowles Commission (see, e.g. Koopmans, 1953) demonstrated that *a priori* identifying restrictions, if valid, can allow one to pick out the economic structure from the reduced form. The plausibility and usefulness of various approaches to identification has, of course, been at the center of major debates in the macroeconometric policy evaluation literature. Sims (1980) argued that the exclusion restrictions used to identify the large macroeconometric models of the 1960s and 1970s were incredible, and his argument emphasized the rich dynamic interdependence in macroeconomic data that was also emphasized by Hendry and Richard. In the wake of the Sims critique, several new approaches to identifying the money demand and or supply functions have been proposed and critiqued: Cooley and LeRoy (1985) echoed the Cowles Commission in arguing that arbitrary block recursivity assumptions were unlikely to reveal economic structure; the historical approach of Romer and Romer (1989) was critiqued by Hoover and Perez (1994); and the use of long-run restrictions (Blanchard and Quah, 1989; King, *et al*, 1991; Shapiro and Watson, 1988) was critiqued by Faust and Leeper (1996). The LSE school has proposed a new approach to identification, and as it is unseemly to have an approach to macroeconometric identification that has not yet been labeled *incredible*, we take up this challenge below.

This discussion of identification has largely presumed a fixed economic structure to identify. The monetary history since the 1960s has led to particular interest in dealing with structures in which both the demand and supply sides may change endogenously. The inflation of the 1970s may have sped the pace of financial innovation during that decade and lead to long-lasting effects on money demand. On the supply side fundamental changes include the switch from fixed to floating exchange rates in 1973 and the change in Federal Reserve operating procedures in 1979. Of course, through the mechanism described in the Lucas (1976) Critique, changes in the money supply process could in principle lead to changes in the money demand process.[2] How such changes are dealt with in econometric modeling is also a fundamental feature distinguishing major econometric camps.

The remainder of the paper explains and critiques the LSE response to the problems of policy analysis. In Section 3, we briefly sketch the approach, hoping to provide an aerial view of the forest before digging into the roots of LSE model fitting in section IV and of policy inferences from those models in section V.

A Sketch of the LSE Approach and its Competitors

At its core, the LSE approach to model estimation simply involves "... seeking models that account for previous findings and explain additional phenomena." (Hendry and Ericsson, 1991b, p.9.) While this goal is probably embraced by most empirical macroeconomists, the devil is in the details. "Accounting for previous findings" and "explaining additional phenomena" mean one thing to real business cycle modelers (Prescott, 1986), for example, and quite another to the LSE econometricians. Indeed, the LSE approach is intensively data-based, relying almost exclusively on a battery of tests to choose the final specification. By observing the outcome of the process, one can deduce that the goal is to formulate the most parsimonious model possible subject to the constraint that the model pass a host of classical tests. To quote Friedman and Schwartz (1991, p.39), the LSE approach is to "start with a collection of numerical data bearing on the question under study, subject them to sophisticated econometric techniques, place great reliance on tests of significance, and end with a single hypothesis (equation), however complex, supposedly 'encompassing' ... all subhypotheses."

These requirements sound as if they apply only to fitting the reduced form -- no mention was made of simultaneous equations bias or of *a priori* identifying restrictions. Not so. Under the LSE approach, one finds relations that are called *structural* through tests, especially tests of stability. In short, if an equation passes enough statistical tests, it simply must be a useful approximation to economic structure. Even more concisely, "... the three golden rules of econometrics are test, test, and test." (Hendry, 1995, p. 557).

Throughout the paper, we will use as examples the U.K. and U.S. money demand equations derived by Hendry and Ericsson (1991a) and Baba, Hendry, and Starr (1992). The Hendry-Ericsson updating of Hendry's (1988) money demand equation for the U.K. is [p.384]:

$$(HE3)\quad \Delta(m-p)_t = -\underset{[0.08]}{0.31}\,\Delta(m-p)_{t-1} - \underset{[0.012]}{0.102}\,(m-p-y)_{t-2}$$

$$-\underset{[0.24]}{0.80}\,\Delta p_t - \underset{[0.10]}{0.63}\,R3_t + \underset{[0.13]}{0.28}\,\Delta y_{t-1} + \underset{[0.006]}{0.022}$$

where heteroscedasticity-consistent standard errors (White, 1980) are in brackets, the data are quarterly from 1964(3)-1979(4), Δ denotes first difference, *m* is the log of M_1 (which includes interest-bearing "sight" accounts), *y* is the log of real total final expenditure in 1985 prices, *p* is the log of the associated deflator, and *R3* is the three-month local authority interest rate. Except for the cointegration term (*m-p-y*) included at lag two, (HE3) is relatively simple and conventional. But if the U.K. equation is well done, the Baba, Hendry, Starr (1992, BHS) U.S. equation may look a bit overcooked:

$$
\begin{aligned}
(BHS18)\ \ \Delta(m-p)_t = &-\underset{[0.097]}{0.334}\ \Delta_4(m-p)_{t-1} - \underset{[0.039]}{0.156}\ \Delta^2(m-p)_{t-4} - \underset{[0.015]}{0.249}\ (m-p-\tfrac{1}{2}y)_{t-2} \\
&- \underset{[0.046]}{0.33}\ \Delta\hat{p}_t - \underset{[0.132]}{1.097}\ \Delta_4 p_{t-1} + \underset{[0.079]}{0.859}\ V_t + \underset{[1.49]}{11.68}\,\Delta SV_{t-1} \\
&- \underset{[0.104]}{1.409}\ AS_t - \underset{[0.063]}{0.973}\ AR_{1t} - \underset{[0.049]}{0.255}\ \Delta R_{mat} + \underset{[0.055]}{0.435}\ R_{nsat} + \underset{[0.07]}{0.395}\,\Delta Ay_t \\
&+ \underset{[0.02]}{0.352} + \underset{[0.003]}{0.013}\ D_t
\end{aligned}
$$

where again heteroscedasticity-consistent standard errors are in brackets; m is the log of M_1; y is the log of real GNP using base year 1982; p is the log of the deflator; Δ^2 is the square of the first-difference operator; $\Delta\hat{p}_t = \Delta(1+\Delta)p_t$; $\Delta_4 x_t = 0.25(x_{t-1} - x_{t-5})$; V_t is a nine-quarter moving average of quarterly averages of twelve-month moving standard deviations of 20-year bond yields; $SV_t = \max(0,S_t)\times V_t$, where S_t is the spread between the 20-year Treasury bond yield and the coupon equivalent yield on a one-month Treasury bill; $AS_t = 0.5(S_t + S_{t-1})$; AR_{1t} is a two-quarter moving average of the one-month T-bill yield; R_{mat} is the maximum of a passbook savings rate, an ogive-weighted certificate of deposit rate, and an ogive-weighted money market mutual fund rate; R_{nsat} is the average of ogive-weighted NOW and SuperNow rates, Ay_t is a two-quarter moving average of y_t; and D_t is a credit control dummy which is -1 in 1980(2), +1 in 1980(3), and zero otherwise.

Both equations are subjected to a battery of tests for residual serial correlation, parameter non-constancy, heteroscedastic errors, non-normality, ARCH errors, nonzero mean, and excluded regressors. For example, BHS report 11 such tests; needless to say, all the tests are passed.

To researchers not from the LSE camp, (HE3) and, particularly, (BHS18) may seem unusual, and many may wonder about the genesis of these equations. In the beginning, presumably, there was theory. For example, Hendry and Ericsson (1991a) appeal to transactions and portfolio-balance motives for holding money, leading to

> Thus, the aggregate real quantity of money demanded ... is an increasing function of some measure of the volume of real transactions (*Y*).... Further, money demand declines as the opportunity costs of holding money increase, with the latter depending upon the returns to the alternative forms in which wealth might be held....Thus we have
>
> $M^d/P = h(Y,R)$,
>
> where *R* is a vector of interest rates on the alternatives to money, and $h(.,.)$ is increasing in *Y* and decreasing in the elements of *R*. Often, much more specific functional forms are adopted, e.g.,
>
> $$m^d - p = \delta y + \gamma R \qquad (1b)$$
>
> where (here and elsewhere) variables in lower case are in logarithms....
>
> Dynamic adjustment is characterized by a contingent planning model of the form:
>
> $$\Delta(m\text{-}p)_t = \mu_0(\mathrm{L})\, \Delta(m\text{-}p)_{t-1} + \mu_1(\mathrm{L})\, \Delta p_t + \mu_2(\mathrm{L})\, \Delta y_t + \mu_3(\mathrm{L})\, \Delta R_t \quad (2)$$
> $$+ \mu_4[(m^d - p)_{t-1} - (m - p)_{t-1})] + \varepsilon_t,$$
>
> where $\mu_i(\mathrm{L})$ (i= 0,...,3) are finite polynomials in the lag operator L, ε_t is the deviation of the outcome from the plan, and m^d_{t-1} denotes the long-run target value m^d in (1b).... Eq. (2) is an error-correction model, and so is a re-parameterization of an autoregressive-distributed lag model of [*m*,*p*,*y*,*R*] (i.e., in levels). It generalizes the conventional partial-adjustment model, allows separate rates of reaction to the different determinants of money demand .... (Hendry and Ericsson, 1991a, pp. 836-838)

The data are used to determine what cointegrating relationship can be found to fill the role of (HE1b), and which lags in the $\mu_i(\mathrm{L})$ are nonzero, if any. At this stage, a general distributed lag model in all the variables is estimated and simplified according to a mixture of creative art (Hendry, 1995) and rules that we review below. It is in this “simplification” that Hendry-Ericsson and Baba, Hendry, Starr “test, test, and test.”

After completing estimation, one can proceed to the policy conclusions that can be derived from a money demand equation. The LSE’ers take as an important implication of their work that only limited policy conclusions can come from a money demand equation (e.g., Hendry and Ericsson, 1991a). One negative conclusion that arises is that one cannot re-normalize these equations and use the result for the determination of prices. Some long-run conclusions also follow from the cointegration analysis. However, we share with the LSE’ers the view that one cannot answer many of the “What if ...” and “What should...” questions without a more complete model.

Before getting into the details of the LSE approach, it will first prove useful to put into context what has been wrought. Equations (HE3) and (BHS18) are very different from what would be produced by work in the identified-VAR tradition. The model specification principles in that tradition might be summarized as: 1) Start with a general time series model. 2) Stop. That is, estimate a general VAR in which each variable is generally regressed on k lags of every other. Do no further simplifications. After specifying and estimating the model, choose a particular identification, or representation, of the model. Report output not in the form of equation estimates, but rather in sets of “impulse responses” to shocks identified by using the weakest set of economic restrictions possible. These impulse responses or associated summary statistics provide a ready framework in which to answer various questions about the dynamic effects of shocks.

Equations (HE3) and (BHS18) also bear no relationship to what we might expect from RBC-style work on a monetary model, which would generally not be accompanied by any test statistics, but rather by population values of certain statistics implied by the model along with asserted sample analogs. The principles guiding model specification in that sort of work are: 1) Start with a highly stylized structural model. 2) Stop. That is, use economic theory to provide a fully specified structural model in which agents’ decision rules are solutions to well-specified optimization problems. Then “calibrate” (Prescott, 1986) and judge fit by the match between a small number of moments produced from the model and analogues computed from band-pass-filtered data.[3]

So how were (HE3) and (BHS18) obtained and why do they differ so much from what we would see in the work of RBC’ers and VARs’ers? In the next section, we lay out the principles underlying the formulation of these equations, and compare them to the “1) start with... 2) stop” principles used in VAR and RBC work.

The LSE Way

To perform applied policy analysis, one needs a model, and under the LSE approach, the model must pass tests that verify that it is *congruent* and that it *encompasses* all serious contenders. Understanding congruence and encompassing is the key to appreciating much that the LSE approach has to offer. In laying out his “Theory of Reduction,” Hendry (1995, p. 363 ff) provides a “Taxonomy of evaluation information” in which he expands congruence and encompassing into several more specific requirements. Adapting his list, we have eight dictates:

1. Verify that the model is *data admissible* and that the data accurately measure the quantity of interest.

2. Fail to reject the null hypothesis that the model *encompasses* each rival model.

3. Fail to reject the null hypotheses that the residuals are *well-behaved*--homoscedastic and innovations relative to available information.

4. Fail to reject the null hypothesis that the parameters of interest are *constant*.

5. Verify that the model is *theory consistent*.

6. Fail to reject the null hypothesis that conditioning variables are *weakly exogenous*.

7. Verify that the structure is *identifiable*.

8. Fail to reject the null hypothesis that the parameters of interest are *invariant* to changes in parameters of equations for other variables.

The first requirement is just good sense; we categorize the second through fourth requirements as dealing with the reduced form; the last four deal with structure. Initially, we describe the dictates individually, discussing any conflict that arises among the dictates after all are laid out.

Dictate 1: Data admissibility. This is an uncontroversial place to begin: surely we want the best available data. Data admissibility is defined to mean the model incorporates all positivity constraints, accounting identities, etc. that hold in the data. For example, nominal interest rates cannot be negative, and the modeling procedure should reflect this. Likewise, GDP should equal the sum of its components. Yet this can be taken too far. An equation in the level of interest rates with normal errors could in principle produce a negative value if an enormous negative innovation occurred, yet LSE’ers accept that the pragmatic advantages of this modeling assumption might be great enough that the modeler would tolerate a negligible probability of violating data admissibility.

The reduced form dictates

Dictate 2: Encompassing rival models. The encompassing dictate as applied by LSE'ers is a powerful model selection scheme and is the heart of the reduced form part of the method. Encompassing tests are designed to measure whether one model better explains the data than another. The tests were developed largely by LSE econometricians but often take the form of familiar tests of nested or non-nested hypotheses (see, e.g., Mizon, 1984). The encompassing logic goes as follows. Suppose we have two competing money demand models, $M_1(\theta)$ and $M_2(\varphi)$ with associated parameter vectors θ, φ, estimated by $\hat{\theta}, \hat{\varphi}$. The models may, but need not, be nested. Based on the assumption that $M_1(\hat{\theta})$ is true, one can generally--perhaps by specifying additional assumptions--derive the value, $\tilde{\varphi}$, one should estimate for the misspecified model. The test of whether M_1 encompasses M_2 is a test of whether the actual estimate, $\hat{\varphi}$, equals the derived $\tilde{\varphi}$. Thus, encompassing basically involves testing the overidentifying restrictions that the truth of M_1 places on the estimates of the misspecified model M_2. In the simplest case when M_1 simply is M_2 with zero restrictions applied to some subset of exogenous regressors, the F test of the exclusion restriction is an encompassing test. In more general cases, many complications have been elaborated.

The encompassing dictate becomes a powerful model selection scheme when one specifies the class of rivals the preferred model must encompass. Every linear autoregressive model is nested in an unrestricted VAR on the same variables, and to the LSE school, the unrestricted VAR should always be considered a rival to the preferred model. Thus the preferred model must encompass the VAR, implying that the additional variables in the VAR cannot help explain the residuals of the model (after proper allowance for degrees of freedom). As the VAR can be viewed as a reduced form, in standard Cowles Commission language, the encompassing requirement subsumes the requirement that one test any overidentifying restrictions of the current model.

Not only is the unrestricted VAR a rival to the preferred model, a large class of models nested in the VAR and preferred models is also considered in competition. From equations (HE3) and (BHS18) above, it is clear that this class includes models formed by imposing zero restrictions (Δp_t is excluded from HE3), by imposing any level and combination of differences (to get regressors like $\Delta(1+\Delta)p_t$), and imposing moving averages (as on income and volatility in BHS18). One assesses all combinations of all such restrictions and retains those that are not rejected in classical tests.

The twin requirements that the preferred model encompass (or not be parsimoniously encompassed by) each of these nested models and that it also encompass the VAR essentially require that one select the *most parsimonious* acceptable model that encompasses the VAR, where the set of acceptable models includes those defined by applying the transforms reviewed above to the VAR. It is not clear, however, what bounds the set of permissible transforms. While it is clear that one is not allowed to *impose* a coefficient at its estimated value, it does seem that one has great flexibility in designing regressors so that a model with few estimated coefficients will fit the data.

The encompassing dictate clearly plays a major role in the design of the two money demand equations. Since the final model must encompass the VAR in which it is nested, one begins with a VAR and attempts to improve on it. More specifically, the process begins with the levels of the variables and an analysis of whether there are any integrated variables and cointegrating relations among the variables. This step begins with a Johansen (1988)-style analysis of an unrestricted vector error correction model for the variables of interest (*m, p, y,* etc.). In both the U.S. and U.K. cases, one cointegrating vector is found (in both cases involving *m, p, y*). Next, since the focus is only on the money equation, one limits consideration to an autoregression of money on contemporaneous values and lags of all the variables, the dummies of interest, and the error correction term defined by the cointegrating vector. In the U.K. case, the route to a simple model involved imposing zero restrictions, differences, and differences of differences. In practice, one acquires a knack for spotting profitable opportunities for such restrictions--if two coefficients on adjacent lags are about equal in magnitude but opposite in sign, one replaces them with a single coefficient on a first difference.

In the U.S. case, we are less sure of how this process worked. In particular, it is difficult to discern how the original general model was specified and how a natural sequence of progressively simpler models led to the final equation. Some of the simplifications are straightforward: for example, the coefficients on R_{1t} and R_{1t-1} in BHS Table 1 are nearly equal in magnitude, so the two-quarter moving average is imposed, and AR_{1t} appears in place of a linear combination of R_{1t} and R_{1t-1} But how ΔSV_{t-1} came to be included in (BHS18) is more mysterious.[4] Hendry (1987), for example, would not view this as a criticism: model discovery is an art that cannot be codified. We take up this argument below.

Dictate 3: Well-behaved errors. The encompassing requirement ensures that the residuals cannot be explained by the variables in the unrestricted VAR. The third dictate further requires that the residuals pass a battery of tests against heteroscedasticity, ARCH (autoregressive conditional heteroscedasticity), nonzero mean, serial correlation, and non-normality.

There are three motivations for such tests. First, some can be viewed as tests of auxiliary assumptions needed for other tests to have desired properties. Second, violation of some of the tests suggests that an alternative modeling strategy may be preferred. If the residuals are far from normal, then the assumption that little is lost in using the asymptotic approximation to the distribution of the test statistics may be more questionable than usual. Similarly, under non-normality, an alternative to the ordinary least squares estimator or some alternative transform of the data might be in order.

The third motivation is that some of these tests may change how one interprets the economic questions of interest. For example, the presence of ARCH in the residuals of certain equations might lead one to abandon models of constant risk in the economy in favor of time-varying risk.

It is clear that the "well-behaved error" dictate is an important subtext in the tale of the two money demand equations: the relevant tests are satisfied. There may well have been other models which would pass the encompassing tests versus one of the models, but which did not pass the tests required in dictate 3. Such models are rejected.

Dictate 4: Constant parameters. The LSE school stands out among econometric approaches in its emphasis on stability testing. The workhorse software packages of the approach, PcGive and PcFiml, provide the capability to do hundreds of stability tests on estimated equations. For example, one can automatically do Chow (1960) tests for every feasible sample split, estimate all the parameters in the model using a rolling sample (fixed beginning point, progressively later endpoint) and view their stability, and do various one-step-ahead tests on rolling samples. Hendry and Ericsson and BHS report a number of these tests for their estimated equation. Once again, any simplification of the unrestricted VAR would be abandoned if the hypothesis of constancy of the model could be rejected for that simplification.

The stability tests are motivated by the same three factors motivating the residual diagnostics in the previous section: test properties and optimal inference may be different if parameters are not stable. More than with "well-behaved errors," however, it is clear that stability is itself often of fundamental importance for the economic questions being asked. If one is interested in whether the income elasticity of money demand is about one, a point estimate of one is less supportive of the theory if it comes from a sample in which the first half estimate is zero and a second half estimate is two. As stability is often central to the economic interpretation, testing of stability is central to the LSE method.

Interpreting the reduced form dictates

In the abstract, the goals underlying the reduced form dictates are laudable: seek stable models with well-behaved errors and which can be shown to be better than obvious alternatives. Indeed, some variation on this theme is part of each of the progressive modeling strategies we are discussing. Perhaps more than others, the LSE'ers have emphasized the application of such standards and are to be commended for defining their own notion of *better model* and developing the tools needed to implement that notion. In this section, we attempt to further elucidate the foundations of this notion by comparing it to the VAR and RBC approaches.

The LSE, VAR, and RBC approaches to reduced-form issues are sharply different on two essential questions: 1) What features of the available data should one attempt to fit? 2) How much theory-free modification of the empirical specification is allowed to improve parsimonious data fit? These two questions go hand in hand and arise naturally from the fact that the macroeconomic data have exceedingly rich and complex dynamics--far richer than we can account for with parsimonious *a priori* theory. Given this fact, answers to the two questions lay out clear courses for the modeler. If one wants to match a great deal of the variation in a macroeconomic dataset, then one must allow for a general time series specification guided only loosely by *a priori* theory. Alternatively, if one wants to start with tightly specified theory, one cannot hope to fit much of the variation in the data. This is the approach in the RBC school, which answers both questions “not much:” one starts with a few data series, pre-filters in some way, and then throws out all but a few moments of the filtered data; the goal is to fit only these few features with a tightly specified theoretical model, and little theory-free tweaking is allowed to improve the fit.

The VARs’ers, in contrast, answer the questions “as much as possible” and “not much:” to fit the data, one must start with a very flexible time series model, and if one is not going to allow much purely data-based modification of the model, then the final product cannot be parsimonious.[5] The LSE schools answers “as much as possible” and “a great deal:” one starts with a general time series model, then, applying purely data-based criteria according to the encompassing criterion, one moves toward the most parsimonious model that fits about as well as the general model.

Each of the three schools answers both questions in an extreme way. (The fourth possibility, “not much” together with “a great deal,” is lambasted in most introductory econometrics texts.) Can statistical theory tell us which approach is best? One of the simplest elements of this issue is the variable selection problem, on which there is a large and venerable statistical literature (e.g., Amemiya, 1980; Akaike, 1973; Schwarz, 1978; Stone, 1981; see also Chapter 21 of Judge, Griffiths, Hill, Lütkepohl, and Lee, 1985). As Hendry (1993a, p.442) emphasizes, very little can be said formally in cases as rich and complicated as those found in macroeconomics--where, for example, the sequence of hypotheses is neither nested nor naturally ordered. Thus, optimality theorems will not help us choose among these approaches.[6] Results from some simpler cases may suggest some limited guidance, however. Thus, if one is going to engage in a data-based model search, a general-to-specific search scheme, such as that suggested by the LSE, seems less likely to lead one astray than specific-to-general.

The clearest conclusion from the statistical model selection literature, however, is that any selection criterion reflects a preference ordering over relevant factors such as specification error and parsimony versus model complexity. One should not hope to find an unambiguously best criterion for all contexts. Different beliefs about complexity and preferences over these factors will lead to very different criteria, which select models with very different degrees of parsimony--that is, the loss function and prior matter a great deal. In simple contexts that can be analyzed analytically, the analogs of the unrestricted model chosen by the VAR camp and a highly parsimonious model based on testing criteria can both be shown to be admissible (Stone, 1981). In our view, claims that one approach is best could profitably be replaced by a discussion of the beliefs and preferences under which the approach is preferred.

Setting optimality aside, one still might wonder about the proper statistical interpretation of the tests in the long, dependent sequence proposed by the LSE. The LSE'ers have staked out a unique position in this regard:

> When an empirical model is selected to ensure that a test criterion is insignificant, such an outcome must occur--independently of the correctness of the solution. Within-sample tests reflect design adequacy and are *indices of congruency*, not model validity: *genuine* testing requires fresh evidence. (Hendry, 1993b, p 20, *emphasis added*)]

That LSE'ers are computing indices and not genuine tests is apparent in the practice of degrees-of-freedom reclamation common in this work. For example, if a test suggests simplification of a linear combination of two lags to a simple difference, the equation has one less right-hand-side variable, and LSE'ers use that additional degree of freedom in ensuing tests.[7] Thus, along the path to a model for which a certain vector of tests does not reject, the things that look like classical tests are merely design criteria--side constraints on a systematic program of data mining. Importantly, in pursuing the LSE approach, one brings to the project a set of tools that virtually guarantees that one can find a model satisfying the test criteria on any dataset, regardless of whether the underlying process satisfies the criteria (as noted by Hendry, 1993b, p.24). Put most sharply, the method dictates that one continue the search until the criteria are met with no stopping rule for deciding that the sought-after relation just is not in the data.

Of course, some check must be put on this process in order to avoid creating models that just happen to fit the current sample. The LSE school relies on the accumulation of new evidence, which provides the basis for *genuine* tests. This provides some insight as to the contexts in which one might prefer other approaches over the LSE approach: it may be true that in the long-run new evidence will weed out bad LSE equations, but policy analysts who must give day-to-day advice may lean toward Keynes's view of the long run. Such analysts might choose some other model selection approach that is not so reliant on the slow accumulation of new macroeconomic data, and in this regard, theory-based restrictions, even if not strictly correct, may provide a valuable tool.

At a more general level, the testing dictates of the LSE method could facetiously be re-cast as the dictate to attempt to answer the standard class of "What if...?" seminar and referee questions before finishing the paper: "What if you included variable q at lag p?" LSE'er's response: "done that, insignificant." "What if your errors have fat tails so that your test statistics won't behave properly?" LSE'er response: "tried that, test reveals no excess kurtosis." And so on. When one regards the final equation in an LSE econometric paper and is aware of all the steps that went into its formulation, one knows the answer to a huge number of such questions. Holding fixed the final model, knowing the answers to such questions can be a very useful thing. Below we discuss an example in the money demand context in which stability tests led to important modifications of LSE models that have not been pursued in VAR work.[8]

The clear virtue of the LSE in testing maintained assumptions becomes hazier when one does not hold the final model fixed. It is the modification of the model based on tests that reduces the *tests* to *indices of congruence.* We suspect that Prescott and Hendry would agree on the reasons why the tests are not genuine tests, but Prescott's RBC school holds the view that tests are neither genuine nor interesting. Indeed, while the LSE'ers have replaced the label (*test* with *index of congruence*), they have provided little guidance or theoretical foundation as to the proper interpretation of such indices.

To summarize our interpretation of the reduced-form dictates, practitioners in the three camps take extreme approaches to which features of the data to fit and whether to employ data-based methods for obtaining parsimony. The LSE and VAR camps worry that the RBC practice of ignoring most features of the data runs the risk of throwing out features that are central to resolving issues in question.[9] The LSE'ers join the VARs'ers in attempting to fit all the data, but the LSE'ers use data-based methods for obtaining parsimony. LSE'ers see danger in the lack of parsimony in VAR work, while VARs'ers question the validity of inferences from approaches that obtain parsimony in the data-based manner advocated by the LSE school. If statistical theory suggests anything, it suggests that the choice among these methods will to turn on prior beliefs and specifics of loss functions. Very little theoretical work has been done evaluating which of these positions is likely to be most useful in the context of policy analysis, or whether some middle ground might be more productive. We review some of the existing evidence on performance of the LSE method below.

The Structural Dictates

Dictate 5: Theory consistency. This dictate is uncontroversial viewed in isolation: make your model consistent with received theory. But received by whom? Certainly by the standards of modern monetary theory, (HE1b) does not constitute a model of money--there is no specification of preferences faced by agents, no specification of constraints, and no specification of special features of the environment which give rise to money. Moreover, the "contingency planning" interpretation of (HE2) lets the barbarians through the gates, since it expands the specification to a general time series model in the relevant (potentially integrated) variables.

But even if LSE'ers began from received theory, the theory consistency dictate inherently conflicts with the encompassing dictate. Starting from a general model, a nested model chosen strictly on the basis of the encompassing criterion must encompass the theory-constrained model.[10] Thus, if we follow the encompassing dictate, theory consistency is largely irrelevant. Alternatively, if we put weight both on theory consistency and encompassing, then we need some guidance as to how to trade off theory versus data consistency. We find no such guidance, and the evidence we have on this count indicates that LSE preferences are lexicographic over data and theory consistency: impose theory consistency only if it has no cost in terms of data consistency.[11]

This discussion is much simpler for the Bayesian: the posterior mean reported in any work need not (and generally will not) be at the peak of the likelihood, and the stronger one's prior in favor of some--perhaps theoretically motivated--point in the parameter space, the further the posterior mean can be from the peak likelihood. The Bayesian has a ready-made and coherent--though, still controversial--approach for combining prior commitment to theory with data, but there is no comparable general scheme for the LSE econometrician. The apparent LSE response is to place theory consistency below data consistency (encompassing) in a lexicographic preference ordering, and may lead one to wonder whether this approach "takes the econ out of econometrics." The clearest role for economic theory in LSE work seems to be in helping specify the variables to include in the general time series model that is the starting point of the model search.

Of course, the role of economic theory consistency is another topic on which the three camps take extreme positions. Each school uses theory to define the relevant variables upon which to focus. Theory has little other contribution in LSE work, and in VAR work theory is additionally used to generate a set of restrictions just sufficient to render the model economically identified. Of course, RBC'ers place theory consistency above all.

Dictate 6: Weak exogeneity. In traditional econometric language, independent variables are required to be exogenous in order to avoid simultaneity bias, and the more rigorously defined requirement of *weak exogeneity* still plays this role, but it is a far stronger requirement. Absence of simultaneity bias in (HE3) essentially requires a lack of correlation between the contemporaneous regressors (Δp_t and $R3_t$) and the error in the money demand equation. If this holds, OLS will, for example, give consistent estimates of the parameters.

Weak exogeneity, as defined by Engle, Hendry, and Richard (1983), is a necessary condition for estimates conditioned on Δp_t and $R3_t$ to be *efficient.* Thus, weak exogeneity could fail due to higher order dependence among Δp_t, $R3_t$, and the money demand error. Further, weak exogeneity requires an absence of cross-equation restrictions between the coefficients of the money demand equation and the equations for the independent variables. If such restrictions exist, a system method of estimation that jointly estimates the equations while imposing the restrictions will be more efficient. So long as one has no real commitments to economic theory, this latter restriction seems as if it would have little bite, but in the face of commitments to theory, the weak exogeneity dictate amounts to the rule that every theory restriction be exploited.

There is one purely statistical source of cross equation restrictions coming from cointegration, and this has been the source of considerable discussion among LSE econometricians (Johanasen, 1992). What is remarkable about the insistence on weak exogeneity in this context is that in general the estimation of the cointegrating relationship in isolation (that is, inefficiently) still results in a superconsistent estimate of the cointegration parameters.

Weak exoegeneity is an important concept elegantly defined by Engle, *et al.* When efficient estimation is the goal, it is an indispensable concept. However, in many contexts efficient estimation is not the overriding goal, and some other exogeneity concept will be appropriate. Thus, in our view, weak exogeneity should be placed alongside the related concepts giving conditions for conditional estimation to have other desirable properties, and the weak exogeneity dictate should be followed when appropriate.

Dictate 7: Identification. LSE'ers have pioneered a new approach to identification based on the existence of structural breaks in equations. This approach has been the source of many strong claims regarding endogeneity in the monetary sector and the empirical relevance of the Lucas Critique. In this section, we argue that the approach is flawed. The scheme has a natural interpretation in terms of the classical approach to identification laid out by the Cowles Commission, differing most substantially in the use of data-based zero restrictions and failure to check analogs of the classic rank and order conditions for identification. Although we came to the criticism in conversations with Ed Leamer, in the end, we attribute the ideas to Engle and Hendry (1993). We should be clear from the outset that some LSE work relies in part on standard, say, instrumental variables approaches to identification and avoiding simultaneity bias [e.g., Hendry and Ericsson, 1991b]. This section concerns only the innovative, breaks-based approach that has been the most prominent approach to identification in the LSE money demand work. We begin with a brief review of identification in simultaneous systems.

In a linear simultaneous equations system, there are many observationally equivalent ways to represent any reduced form. Identification is a matter of choosing a unique representation, and one usually hopes to pick a representation that has a natural interpretation in terms of economic behavior. The classic example is when we have two equations jointly determining the equilibrium price and quantity of, say, econometrics texts. One representation of the system will involve the supply equation and the demand equation, and every other observationally equivalent representation will involve two equations, each of which is a linear combination of the supply and demand equations. The assumptions allowing one to pick out a unique representation of the system are called identifying assumptions.

Suppose an LSE'er setting out to model behavior of the money market believed that the universe of economic time series relevant to this task were {(*m*-*p*), y, *R*, D_1, D_2, D_3}, and that D_1, D_2, D_3 were exogenous in every possible sense of the word. Letting Y_t denote the column vector of contemporaneous values of the three endogenous variables and Z_t the three exogenous variables, suppose that the true dynamic behavior of these variables can be written

$$\Gamma Y_t = A(L)Y_{t-1} + CZ_t + U_t \tag{1}$$

where the first equation is money demand, Γ denotes the square matrix of contemporaneous coefficients (normalized to have ones on the diagonal), $A(L)$ is a matrix polynomial in the lag operator L which gathers together the lag coefficients, and U_t denotes the vector of three error terms. Defining a new variable X_t to contain all relevant lags of Y_t and the vector Z_t, the system is seen to be in typical simultaneous equations form:

$$\Gamma Y_t + BX_t = U_t. \tag{2}$$

Suppose that someone correctly claims that the first equation of this system is money demand. Koopmans (1953) laid out the fundamental problem of identifying economic relations in noting that a "prankster" might deceitfully ask, "Why shouldn't we consider the first equation of

$$R\Gamma Y_t + RBX_t = Ru_t \tag{3}$$

to be the money demand equation, where R is any full rank matrix?" The two reduced forms are the same,

$$Y_t = \prod X_t + V_t, \tag{4}$$

where

$$\prod = -\Gamma^{-1}B; \;\; V_t = \Gamma^{-1}U_t, \tag{5}$$

and fit the data equally well. Since both systems have the same empirical implications, no fact about the data alone can answer the prankster. What is required is information from outside the system--*a priori* identifying information.

The simplest textbook (e.g., Koopmans, 1953; Theil, 1971) approach to identification is when we have *a priori* information about the placement of zeros in Γ and B. Such information is identifying if whenever $R\Gamma$ and RB satisfy the zero restrictions, the first row of R is the first row of the identity matrix. In this case, all transformations of the system that satisfy the *a priori* zero restrictions have the same coefficients for the first equation, and, the equation is identified. The well-known rank and order conditions for identifying the first equation are of this type. The order condition, which is necessary for identification in this way, requires that the number of zeros in the first row of B be greater than or equal to the number of nonzero elements in the first row of Γ less one. That is, the *number* of excluded exogenous variables must exceed the number of included "other" endogenous variables (the "less one" corresponds to the variable on which the equation is normalized). The instrumental variables language for the order condition is that we need an exogenous or predetermined variable excluded from the equation to serve as an instrument for each of the other endogenous variables. The rank condition, which is sufficient for identification, requires that if one makes a matrix out of the columns of Γ and B corresponding to zeros in their first rows, the rank of that matrix must be equal to the number of equations less one. That is, not only must the right number of exogenous variables be excluded from the equation of interest, but they must be included in a rich enough way elsewhere. If the rank condition is just met, the model is just identified and the restrictions yield no testable implications. If there are more than enough restrictions to meet the condition, then the overidentifying restrictions can be tested using standard methods.

The identified-VAR school follows this same logic in identifying its models. The school uses restrictions on Γ and (and less often, on B) and restrictions on the covariance matrix of U to achieve identification. Typically, the models are just identified. Of course, the RBC school starts with fully solved models, which are strongly overidentified under the usual assumptions concerning the source of stochastic error. The overidentifying restrictions would generally be rejected in formal tests, but such testing is not part of the scheme. But while the VAR and RBC approaches to identification might be deemed conventional, the LSE school seems to take a different path, arguing that the classical identification framework needs to be "reinterpreted in worlds of parameter change." (Hendry, 1987, p.40) As will become clear below, the details of this reinterpretation, especially regarding the need for *a priori* identifying restrictions, is difficult to discover. We begin by describing the simplest form of LSE breaks identification.

First, one refines the specification of the equation of interest until one fails to reject the hypothesis of parameter constancy using a battery of tests. The equation may involve contemporaneous, potentially endogenous variables, but in this step one ignores the possible simultaneity problem. Next, one estimates simple models--say, univariate autoregressions--for each of the endogenous variables, and tests the stability of those equations.[12] If those equations reject the hypothesis of stability, one concludes that simultaneity is not a problem in the equation of interest and that the equation must be one of the structural equations of the system. According to Hendry (1987, p.41), "Models that remain constant despite asserted regime shifts have a clear claim to the epithet structural." Since Hendry makes no mention of a priori restrictions, one might suppose that he is using the term *structural* in a different way from when we use it to mean that the equation deserves an economically meaningful title like "money demand" or "money supply." This is not the case, however, since in the LSE work on money one proceeds from the stability tests to applying one of those titles based on the signs of the coefficients: a negative interest rate coefficient, for example, justifies the name money demand.[13] As Baba, Hendry, and Starr summarize the argument,

> Moreover, when the resulting conditional demand model is constant, but it is known that the supply function shifted during the sample period ..., then the "classical" identification problem does not arise, since all linear combinations involving the shifting equations are automatically excluded. (Baba, Hendry, Star, 1992, p.33)

As stated, the argument begs the question by calling the equation of interest the "conditional *demand* model," and indicating *a priori* knowledge of a shift in *supply*. *Un*-begging the question, the statement becomes: "if one finds a stable equation with real balances on the left-hand side and finds unstable equations for contemporaneous regressors in that equation, then the classical identification problem does not arise." Why not? The prankster can take the equation of interest, stack it with the equations for the other endogenous variables into a system, premultiply by an arbitrary full rank matrix R, and ask why we should not call the first equation of the resulting system money demand. The answer might be that the resulting equation does not satisfy the *a priori* restrictions on the signs of the coefficients. If this is so for every full rank R (except those with first row equal to that of the identity matrix) then the model is identified by economic sign restrictions, and the discussion of breaks is a sideshow. In general, however, there will be a wide range of Rs giving rise to systems satisfying the sign restrictions.

A second argument, suggested by the passage from Baba, *et al*. is that the observationally equivalent models proposed by the prankster will involve unstable money demand functions--as the quote implies, a linear combination of a stable and unstable equation is unstable. However, no assumptions rule this out. Absent the *a priori* assumption that the money demand equation is stable, the stability tests provide no basis for answering the prankster.

The issue may be further clarified by reinterpreting breaks identification in terms of classical identification. For simplicity, we limit the discussion of instability to changes caused by shifts in the intercepts of the equations.[14] These intercept shifts can be modeled as the inclusion of a step dummy--an exogenous variable that is zero in some time periods and one in the remaining periods--in the equation. The test results showing that the equation of interest is stable while the others are not could be caused by a step dummy that is not in the first equation but is in the others. This, of course, is merely one, test-based, zero restriction on an exogenous variable in the equation of interest. If we had enough such *a priori* restrictions to satisfy the rank condition, then the system would be identified; otherwise it is not.

From the classical identification perspective, the breaks-test approach to identification uses data-based approaches to find zero restrictions on step dummies. There are two problems with this. First, since the zero restrictions are data-based they provide no basis for answering the prankster who claims that the equation is a linear combination of underlying structural equations. Second, even if one made an *a priori* assumption that the money demand equation were stable, the breaks tests do not reveal how many step variables are in the system nor how they enter the system; thus, there is no basis for the assertion that the order or rank conditions are satisfied.

This latter criticism is set aside by a second approach used by the LSE'ers. In this approach, the first step remains the same: build a constant equation of interest ignoring simultaneity. In the second step, one modifies the unstable equations for the endogenous regressors by adding constructed dummy variables until the equations pass stability tests. One then tests for the significance of these dummies in the equation of interest. The zero restrictions are once again data-based. In this case, however, if one were to treat the dummies as instruments, one could check the rank and order conditions on identification.

In laying out the formalities of identification-by-breaks, Engle and Hendry (1993) provide a little-recognized paragraph that seems to foreshadow our critique. The authors posit existence of a set of instruments (known to be exogenous) called Z and containing a subset z that is included in the equation of interest. The Zs may be shift dummies. The authors argue

> In the common case where one is unwilling or unable to specify the entire set Z, it may still be possible to perform the test. Partition Z into (z, Z_1, Z_2) where Z_1 and Z_2 are excluded from (5) [the equation of interest] on *a priori* grounds (rather than merely because of insignificant coefficients in pre-tests) and Z_1 is observed. Z_1 might be dummies for shifts in regimes which, under superexogeneity, ought not to enter (5). A test of superexogeneity would then be whether Z_1 enters (5). (Engle and Hendry, 1993, p.130)

The authors do not elaborate on what is different about the uncommon case, in which one knows the entire set Z, but we suspect the difference flows from an implicit assumption that z is a proper subset of Z. This once again implies that the modeler knows that there are valid instruments excluded from the equation of interest. We call this paragraph *little recognized* because in none of the LSE work we reviewed did we see any discussion of whether the common case applies or of *a priori* assumptions about which equations contain the breaks.

We conclude, as Engle-Hendry (1983) seem to, that identification based on breaks will resolve the identification problem posed by Koopman's prankster only if the breaks are viewed as any other exogenous variable and restrictions on those variables meet the classical conditions for identification. Identification based on exogenous regime changes is potentially a very important idea, as such breaks may add considerably to the list of potential instruments. The breaks will be useful, however, only in cases where one can make an *a priori* case that the breaks affected one equation and not others.

One might suppose that those in the LSE school would be open to the idea of assuming *a priori* that a certain break occurs in one equation but not others. However, such an assumption would stand in the way of another project regarding invariance.

Dictate 8: Invariance. Parameter invariance is defined by the LSE school to mean that a parameter remains the same under some class of interventions on coefficients in another equation. If a parameter is invariant in this way, then the Lucas (1976) Critique does not apply to that parameter under that class of interventions.

LSE econometricians have been in the forefront of proposing tests of invariance and in claiming that the Lucas Critique appears to be of little relevance empirically (e.g., Favero and Hendry, 1992; Ericsson and Irons, 1995; see also Leeper, 1995, for an additional critique of these tests, and Neftci and Sargent, 1978, for an early example of such work). The breaks tests just discussed form the basis of this argument. If money demand is stable and the other equations of the model are not, then the Lucas Critique must not have been operative.

Given our discussion of identification above, we find a circularity in the LSE critique of the Lucas Critique: the breaks tests only properly identify money demand if they are accompanied by the *a priori* assumption that the break enters other equations, but not money demand. But this is equivalent to the assumption that the Lucas Critique does not apply. Without this assumption, there is no reason to call the equation money demand, and, hence, no way to test if the Lucas Critique applies.

The circularity has another implication. The LSE school finds a stable equation which they call money demand and finds that all the other equations are unstable. If identification is wrong, then the true money demand equation is a linear combination of all the equations of the system, and will, in general be unstable. In other words, suspicion about the identification supports the view that the empirical evidence is consistent with the Lucas Critique.

Some inessential tenets

We have come to the end of our account of the LSE dictates without stating a number of claims common in LSE'er accounts. For example:

1) Weak exogeneity is necessary for valid conditional inference (Hendry and Ericsson, 1991b; Mizon, 1995; Hendry 1995).

2) The final model after the data-based reduction involves no loss of information relative to the original data.

> Encompassing ensures that there is no loss of information in the reduction as a whole... (Hendry, 1995, p.365; see also Mizon, 1995).

3) The data-based path taken to select the final model is irrelevant in assessing the credibility of the model.

> Indeed, it seems a crucial difference from other viewpoints (e.g., Leamer 1983) that the credibility of the model is not dependent on its mode of discovery but on how well it survives later evaluation of all of its properties and implications... (Hendry, 1987, p.37)

If the claims were accorded a natural interpretation, they would be exceedingly important, and we would be remiss in not giving them prominence. Everyone is interested in valid inference, and inference without loss of information sounds wonderful. Fears about data mining seem to be set aside in the view that the credibility of a model is independent of its mode of discovery.

The straightforward interpretation of these claims is not warranted, however, and the claims seem to be shorthand for less important claims. For example, as we have discussed above, the first claim seems to be LSE shorthand for the established claim that weak exogeneity is necessary for efficient conditional inference. Weak exogeneity is often not necessary to attain other desirable properties.

The second claim states that the final model, after data-based simplification, involves no loss of information from the original unrestricted starting point of the specification search. At times, claim 2 seems to be interpreted as applying only when employing the notion of encompassing in population, rather than in finite samples, or, equivalently, as holding conditional on all the encompassing tests getting the right answer.[15] Such claims are correct, but are nearly vacuous--so long as the variable selection criterion gets the right answer, there is no loss of information in following it. Most importantly, the claim has no finite-sample implications: the encompassing tests have type II error rates, and imposing the null may involve loss of information.

The third claim is that the credibility of estimated models is independent of the way in which they are discovered. Once again, there is a closely related claim that will generally be correct: the mode of model selection cannot affect the intrinsic validity of the final chosen model (Hendry, 1987). If intrinsic validity is taken to involve some metric on the distance between the model estimates and some true model, then such a distance between truth and model will not depend on the way the model was found. Of course, we can never evaluate intrinsic validity. In practice, we evaluate the credibility of models using probabilistic inferences about how close the model is to the truth. In both a Bayesian and classical framework, such inferences depend on how the reported statistics were generated.

For example, suppose a researcher has time series on three variables: x_1, x_2, and y. It is known that y depends on one and only one of the x's and, in particular, that $y = a + b\, x_j + \varepsilon$ where ε is independent of the x's, and j is either 1 or 2. The researcher reports the ordinary least squares estimate of b and the associated t statistic for the regression of y on x_1. The t statistic is 2.1. The credibility of the claim that it is x_1 that belongs in the regression clearly depends on how this model was obtained. Consider two possibilities: 1) A coin flip determined which single regression was run and reported. 2) Both regressions were run and the one with the larger t statistic was reported. Under neither approach is the reported information a sufficient statistic for the data under the problem at hand, and different information is contained in the statistics in the two cases. In the first case, we know a single t statistic, whereas in the second case we know the t statistic and the fact that the other t statistic is smaller. From both a Bayesian and classical perspective, the additional information in the second case is important in assessing the credibility of the claim that it is x_1 that is important.

LSE-Style Policy Analysis in Practice

Our review of the principles underlying the LSE approach showed that the reduced-form aspects are quite different from the two competitors we consider, but that reason alone was unlikely to resolve which is most useful. We found stronger reasons to be suspicious of identification claims in LSE-style work. But perhaps the most important test of an approach to macroeconomic policy analysis is not whether one can criticize its theoretical foundations: given the problems faced in policy analysis, any approach that yields answers will be subject to criticisms. It is more important to ask whether the approach has led to important results. Are models estimated under this approach used to answer important policy questions? Have the answers been persuasive to the general economic community? Have the answers proven reliable? Would those answers have been likely to go unnoticed under other approaches?

A natural starting point for an examination of these questions is with a review of the policy results that the LSE'ers have drawn from their work. Next we turn to a brief review of the reduced-form performance of the approach. Finally, we attempt to draw some conclusions about whether LSE results have proven persuasive to the general econometric community.

Policy-related claims in LSE applied econometrics.

The money demand example that we have followed throughout probably provides the clearest example of policy analysis in the LSE approach, and the best summary of those results comes in a section labeled "Policy Implications" that follows an extended discussion of LSE work on UK money demand in Hendry (1995, p. 618). Hendry's discussion draws on much of the money demand work already discussed, but especially, Hendry and Mizon (1993) and Hendry and Doornik (1994). Four implications are listed: First, the opportunity cost of holding money is a spread between the own rate and the rate on substitutes, such as short term interest-bearing notes, and is not simply the rate on some 3-month instrument. Thus, one cannot control the stock of real money simply by controlling the level of short-rates. Second, the money demand equation cannot be re-normalized on the price level to give a stable equation for price determination. Any analysis that does this will be misguided. Third, since the exogeneity tests suggest that money historically came last in a recursive ordering of the economy, stochastic shocks to money demand contemporaneously affect real balances but not interest rates and prices in the economy. Thus, prices were not historically determined by the intersection of a shifting money demand function with an exogenous money supply curve. Fourth, *excess demand* for output is simultaneously determined with inflation and is affected by interest rates. This result comes from placing a structural interpretation on a cointegrating vector, and is said to imply that raising the interest rate to control inflation will operate by lowering output.

The most obvious point to make about these implications is that they do not directly answer any of what are surely the most important policy questions listed in the first section: What effects has policy had historically? What would be the effect of different policy options today? The four implications have a bearing on how these questions should be answered, but they do not singly or jointly provide the material with which to answer them. Thus, we have few clear examples of how such questions should be answered in the LSE approach, and we suppose that answering such questions has simply not been the focus of LSE work. One explanation for absence of more policy claims is that the attention in the LSE school until recently has been on single-equation methods, and the LSE practitioners are clear on the point that "policy implications do not follow directly from a constant money-demand equation..." (Hendry and Ericsson, 1991a, p 861). This is not a criticism: there is a long tradition of single-equation, money demand work, and we infer that there is a strong demand for such results.

The limited policy implications that are listed by Hendry do provide some interesting insights, however. The first claim--when money pays interest the opportunity cost of holding money is an interest rate spread--is uncontroversial. It is part of econometric work in many traditions and microeconomic theory and evidence probably are much more convincing on this point than any aggregate equation. One might be tempted to conclude that there is nothing unique that the LSE method contributed to this conclusion. It is true, however, that most VAR work on the effects of money shocks includes only a short-term interest rate (or perhaps short-term and long-term rates) but not any own rate on assets in the chosen monetary aggregate. Thus, the VARs, which are intended to sort out money demand from money supply, do not contain a measure of the opportunity cost of holding money. Further, the dictates of the LSE method forced LSE'ers to face this issue. In particular, as the own rate on money began to vary in the U.K. in the 1980s, re-estimates of money demand equations appeared unstable, and this instability disappeared with the inclusion of an opportunity cost measure (see, e.g., Hendry and Ericsson, 1991a). In the LSE tradition, one is forced to attempt to resolve this instability; in the VAR tradition, it is often ignored. In fairness to the VAR work, when LSE'ers estimate a system for money, output, prices, and interest rates, they include only the opportunity cost of money, but not the level of interest rates, which may be the relevant measure in the output and price equations (Hendry and Mizon, 1993; Hendry and Doornik, 1994; Hendry, 1995).

The interpretation of the second policy implication hinges crucially on the identification scheme. If the LSE approach has succeeded in identifying a structural equation, the implication follows.

The remaining two implications have a more direct bearing on the important policy questions. Given the LSE emphasis on single-equation, conditional models, it is important to note that neither of the final two implications flows from single-equation work. The third implication rests on the exogeneity tests, which are based on careful estimation of one equation of the system and auxiliary, cursory, estimation of equations for the rest of the system. The fourth implication regards the role of a cointegrating vector. Estimating and identifying multiple cointegrating vectors requires analysis of the entire system.

The third policy implication is essentially a re-statement of the fact that the exogeneity tests reveal real balances to be last in a recursive ordering of (the studied block of) the economy. While we are deeply suspicious of the identification scheme on which this conclusion rests, the conclusion would have important implications for the conduct of policy analysis. It also stands in direct conflict with the emphasis on the endogeneity of money in the VAR literature. As we note below, there has been little attempt to resolve this conflict by VARs'ers or by LSE'ers.

The final policy implication comes from estimation of a *system* involving money, and since we have had limited discussion of system methods in the LSE approach, these system estimates warrant some attention. The four-equation system on U.K. data involving interest rates, money, output and prices of Hendry (1995), Hendry and Mizon (1993) and Hendry and Doornik (1994) provides a good example. Initially, a cointegration analysis is performed to allow the authors to transform the variables into a set of four empirically I(0) variables--the change in real balances, the change in income, the change in inflation, and the change in the opportunity cost of holding money--along with two cointegrating vectors and some dummy variables.

Assessing various identifying assumptions for systems such as this has been the focus of much of the identified VAR literature. To identify the system, Hendry presents a list of zero restrictions on coefficients on contemporaneous and lagged variables in this system. The identifying restrictions are "are based on the accumulated evidence, the earlier theoretical analysis, and the data evidence in Table 16.7." (Hendry, 1995, p.600) In the various accounts of this system, we found no *a priori* argument in favor of any of the restrictions. While great emphasis is placed on satisfying a test of the overidentifying restrictions specified, without *a priori* restrictions it is not clear that such a test confers any legitimacy on the structural interpretation.

Published work in the LSE approach has given us only limited examples of policy analysis to discuss. This is, in part, because the most important and interesting policy questions involve more than a single equation, and much of the focus in the LSE approach has been on single-equation work. While we question the identification approach in the more recent system work, the results do have potentially important implications.

Reduced-form results.

We divided our analysis into reduced form and structural issues and have been critical of the LSE approach to identifying a structure. One can, however, strip the estimated equations of their structural interpretation and ask whether they have accurately characterized the data that arrived after their estimation. Has the LSE approach succeeded in capturing regularities in the data? While this alone is not sufficient for policy analysis, some form of stability is surely necessary for policy analysis, and finding stable macroeconometric relations is itself no mean feat.

Tracking the performance of earlier estimated equations is a notable strength of the LSE school. For example, the history of the Davidson, Hendry, Srba, Yeo (1978) consumption function--perhaps the first and most famous work in what we are calling the LSE school--is detailed in a number of articles collected in Hendry (1993a). The fact that the method is called progressive implies that LSE'ers expect to find problems with earlier work, and, thus, the relevant criterion is not whether the estimated equations required any alteration. A more generous criterion is to assess whether problems in the estimated equations stemmed from flaws in the methodology used to formulate them. The distinct tales of the two money demand equations provide an interesting contrast.

Equation (HE3) for U.K. money demand was first estimated in Hendry (1979) on data from 1963(3) to 1977(4). The estimates we reported were by Hendry and Ericsson for data through 1979(4) and involved little change from the original. Upon re-estimating with revised data and extending the sample to 1985(2), the estimates are essentially unchanged, but forecasts from the estimated model show a massive structural break. The equation drastically underpredicts the unprecedented growth of real balances in the U.K. from the end of the estimation sample through 1989(2). During this period, interest-bearing sight deposits were introduced in the U.K., thus generating variance in the own rate on money that was absent in the original sample period. Estimating on the new data through 1985(2), but including a learning-adjusted own rate on sight deposits, *Ro*, eliminates this instability while leaving the other coefficients of the model largely unchanged (within about 1 standard error of the original values from either equation):

$$\begin{aligned}\Delta(m-p)_t = \underset{[.005]}{.02} + \underset{[0.12]}{0.25}\,\Delta y_{t-1} - \underset{[0.15]}{0.70}\,\Delta p_t - \underset{[0.07]}{0.30}\,\Delta(m-p)_{t-1} \\ - \underset{[0.07]}{0.63}\,R_t - \underset{[0.33]}{0.094}(m-p-y)_{t-2} + \underset{[0.33]}{0.74}\,Ro_t \end{aligned} \qquad (6)$$

The opposite signs and nearly equal magnitudes on the interest rate terms are consistent with the spread between the two rates being the important quantity, as simple theory would predict, and Hendry and Ericsson's final equation reflects this restriction.

The results for this breakdown of an LSE equation represent an impressive victory for the approach. The original equation is both parsimonious and conventional in specification. The modification in response to the payment of interest is a natural one (except perhaps for the learning adjustment of the rate),[16] and the basic features of the equation stood up to the addition of 12 additional years of data.

A less supportive result emerges for the U.S. M1 money demand equation. The reported equation was estimated using data for 1960(3) through 1988(3).[17] Of course, there have been many valiant attempts to fit a stable money demand function over this period, and the wreckage from these attempts is well-documented (Goldfeld, 1976; Judd and Scadding, 1992). BHS succeed in this venture, but not with nearly as parsimonious and conventional an equation as was found for the U.K. The equation, recall, involves many more parameters than the U.K. equation and involves several variables that were constructed based on the data in the sample. The equation is stable over the sample period, which includes the “missing money” period in the 1970s and the great velocity decline in the early 1980s. Finding such a stable equation and a corresponding economic rationalization is a heroic achievement. The question is whether the achievement is testament to the ability of the method to uncover important economic regularities. The alternative, of course, is that the equation is testament to the ability of talented and imaginative practitioners to generate a relation that passes stability tests, regardless of the data.

Hess, Jones, and Porter (1997) claim that the latter interpretation has some validity. They report that the equation’s prediction is outside two-standard error bands for the first quarter after the sample and remains so for much of the period 1988(4) to 1993(4). The equation fails many standard stability tests when extended to 1993(4). Further, Hess, *et al.* attempt to trace the problems to the specially-designed volatility and adjusted rate variables, showing, for example, that the forecasts and errors from equations without the designer variables help predict the forecast errors of the Baba, *et al.*, equation (but not *vice versa*).

The U.S. money demand equation required a great deal more *tailoring* to satisfy the LSE design criteria than did the U.K. equation. It also appears to have had more trouble accounting for new data. While these two examples do not represent a complete review of all the work done in the LSE tradition, the work on money demand has been a primary focus of work in the school during the period of when many elements of the method have been codified. Moreover, the two examples surely highlight the optimistic and pessimistic perspectives on the approach. We are not aware of a body of evidence clearly supporting either perspective.

Inferring the impact of LSE policy conclusions.

In this section, we attempt to asses the impact that the LSE approach has had on the profession’s views regarding empirical regularities. It is clear that the LSE approach is used extensively throughout the world. For example, important LSE works on money demand in the U.S. and U.K. (Hendry [1985], Hendry [1991], Hendry and Ericsson [1992] and Baba, *et al.* [1992]) have jointly received over 80 unique citations, a large number of which appear to be further examples of econometrics in the LSE tradition. The citations include studies of money demand in Argentina, Bolivia, Canada, China, Greece, Switzerland, the U.K., and the U.S. While it unclear whether the popularity of the approach stems from its ability to generate useful policy implications, it is clear that the LSE approach to studying money demand is widely used.

Gross publication counts reveal little regarding the *relative* merits of the widely-used econometric approaches, however. A measure of the relative rates of growth of publications by the three major camps might be more informative, but such rates of growth would be difficult and costly to measure. We can shed some light on whether the results of one approach are drawing the attention and perhaps converts from other approaches, however, by examining cross citations. In particular, much of the identified VAR work on the importance of money in the economy probably cites the seminal works of Bernanke (1986), Blanchard and Watson (1986), and Sims (1986). Similarly, much of the LSE-style work on money demand probably cites at least one of the four LSE papers listed above. If subsequent research is influenced by one or both of these literatures, or if some VAR workers are converted to LSE-style work (or vice versa), one might expect to find papers citing both seminal LSE and VAR work.

We examined the citations to the 4 LSE money demand papers and the three seminal identified VAR papers (Table 1).[18] We would have liked to include seminal RBC papers in this examination, but this proved too costly.[19] Of all the papers, Sims (1986) has received the most citations (77). This work has also had a worldwide impact as indicated by cites in works on Australia, Chile, Canada, Germany, Japan, New Zealand, Pakistan. Both the LSE and the VAR papers have received over 80 citations, and several papers cite more than one LSE paper or more than one VAR paper.

	Hen85	Hen91	H/E91	Baba92	Sims86	B/W86	Ber86
Hen85	30						
Hen81	5/48	35					
H/E91	4/66	3/77	46				
Baba92	0/30	5/47	2/55	18			
Sims86	0/163	0/88	0/108	0/59	77		
B/W86	0/50	0/46	0/58	0/28	7/89	20	
Ber86	0/36	0/38	0/50	0/19	3/79	0/26	6

Notes: The diagonal elements are gross citations of the paper. The off-diagonal elements are N/D where N is the number of papers citing both the row and column paper, and D is the total number of unique papers citing either the row or column paper since the first year in which both papers were cited. Hen1 is Hendry (1985); Hen91, Hendry, 1991; H/E91 Hendry and Ericsson, 1991b; Baba92, Baba, et al. (1992), Sims86, Sims (1986); B/W86, Blanchard and Watson, 1986; Ber86, Bernanke (1986).

Table 1: A cross tabulation of citations to LSE and VAR work

The point we wish to emphasize is that there are no papers that cite both a seminal VAR paper and an LSE paper. One explanation for this would be that, despite the similar-sounding topics, the papers actually have little overlap in content. In fact, however, it is not difficult to find good reasons for cross fertilization. The above discussion of the differing treatment of the opportunity cost of holding money and the conflicting results on simultaneity provide important examples.

The view that the literatures have had little effect on each other is strongly supported by reading typical works that have cited the papers we examine. Both approaches have evolved: newer LSE work begins with recently developed cointegration analysis and newer VAR work evaluates the effect of different information sets and different identifying assumptions on various puzzles in earlier work. None of the innovations or extensions we noticed in either literature appear to be motivated by results in the other literature, however. None of the innovations reconcile the fundamental inconsistencies noted above.

Thus, our limited look at the academic research on empirical monetary economics reveals that the LSE approach is widely practiced, but reveals very little evidence that the LSE method has been successful in generating results that have persuaded the practitioners outside the method. Of course, the VAR approach and (we conjecture) the RBC approach have had similarly little effect on LSE practitioners, and this may simply reflect an insularity in applied macroeconometrics. If one believes, however, that approaches that are successful in generating useful results tend to gain converts, then the cross-citation results may be evidence for a view with which we have some sympathy: none of the approaches has yet clearly distinguished itself in revealing the empirical truths of monetary economics.

Conclusion

The LSE school has created an impressive body of work describing and justifying an approach to econometrics, building the required tools, and applying those tools. The approach has been widely used, and achieved some impressive results along with some failures. Any approach to applied econometrics will be subject to criticism, and we have attempted to present the major criticisms to the LSE approach. For example, the LSE school has relied heavily on a new approach to identification based on the view that it is possible to discriminate between, say, supply and demand functions using stability tests and without using *a priori* restrictions from economic theory. On this count, we applied the logic of the Cowles Commission and agree with its negative conclusion:

> Statistical inference unsupported by economic theory applies to whatever statistical regularities and stable relationships can be discerned in the data. Such purely empirical relationships when discernible are likely to be due to the presence and persistence of the underlying structural relationships, and (if so) could be deduced from a knowledge of the latter. However, the direction of this deduction cannot be reversed--from the empirical to the structural relationships--except possibly with the help of a theory which specifies the form of the structural relationships, the variables which enter into each, and any further details supported by prior observation or deduction therefrom. (Koopmans, 1953, p.28)

To someone attempting to choose among econometric approaches, however, the particularities of these criticisms may be rather less important than the dramatic differences among the tools, procedures, and results in the camps. In concluding his discussion of the LSE, VAR, and Leamer approaches, Pagan (1987) worried that the divergence of macroeconometric methodology since the 1960s might be interpreted as tacit admission of the failure of econometrics. Pagan hoped to see more integration of the approaches as evidence that the divergence was merely a widespread search for better methods, leading to renewed consensus on some basic issues. Perhaps the nine years since he wrote are insufficient to expect much progress in this regard. The LSE school and the RBC and VAR camps with which we have contrasted it remain at extremes on the basic issues of what features of the data to fit, of the advisable amount of testing-based modification of empirical models, and on the role of economic theory in restricting the model search.

We too believe in greater integration of the approaches. For example, on reduced-form issues, VAR and RBC work would benefit from a more thorough examination of the ways in which the models do and do not fit the data. The battery of tests prescribed in the LSE approach is one way to do this. Yet the LSE approach may push testing too far, as data-based modifications of the model are pursued to the point that tests are reduced to indices of congruence with no clear interpretation. The RBC and VAR schools are pursuing (very different) ways to exploit economic theory as a constraint on an otherwise purely data-based exercise. Finding improved ways to exploit economic theory while thoroughly examining the empirical properties of the models would surely be beneficial.

References

Akaike, H. (1973) “Information Theory and the Extension of the Maximum Likelihood Principle,” in B.N. Petrov and F. Csaki, eds., *2nd International Symposium on Information Theory*, Akailseoniai-Kiudo, Budapest, pp. 267-281.

Amemiya, T. (1980) “Selection of Regressors,” *International Economic Review*, 21, 331-354.

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.

Bernanke, B. (1986) "Alternative Explanations of the Money Income Correlation," in K. Brunner and A.Meltzer, eds., *Carnegie-Rochester Conference on Public Policy, Real Business Cycles, Real Exchange Rates and Actual Policies*, 25, 49-100.

Blanchard, O., and Quah, D. (1989) "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review* 79, 655-673.

Blanchard, O. and Watson, M. (1986) "Are All Business Cycles Alike?" in R.J. Gordon (ed.), *The American Business Cycle: Continuity and Change.* Chicago: University of Chicago Press.

Braun, R.A. (1994) "Tax Disturbances and Real Economic Activity in the Postwar United States," *Journal of Monetary Economics* 33, 441-462.

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

Christiano, L.J., and M.S. Eichenbaum, M.S. (1992) "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *American Economic Review* 82, 430-450.

Cogley, T., Nason, J. (1995) "Output Dynamics in Real-Business-Cycle Models", *American Economic Review*, 85, 492-511.

Cooley, T.F., and LeRoy, S.F. (1985) "Atheoretical Macroeconometrics. A Critique" *Journal of Monetary Economics*, 16, 283-308.

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

Engle, R.F. (1982) "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50, 987-1007.

Engle, R.F., and Hendry, D.F. (1993) "Testing Super Exogeneity and Invariance in Regression Models," *Journal of Econometrics* 56, 119-139.

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

Ericsson, N., and Irons, J. (1995) "The Lucas Critique in Practice: Theory without Measurement," in Kevin Hoover, ed., *Macroeconometrics: Developments, Tensions, and Prospects*. Boston: Kluwer Academic, 263-312.

Favero, C., and Hendry, D. (1992) "Testing the Lucas Critique: A Review," *Econometric Reviews*, 11, 265-306.

Faust, J., and Leeper, E.M. (1996) "When do Long-run Identifying Restrictions give Reliable Results?," *Journal of Business Economics and Statistics*, forthcoming.

Faust, J., and Whiteman, C. (1995) "Commentary on Chapter 4," in Kevin Hoover, ed., *Macroeconometrics: Developments, Tensions, and Prospects*. Boston: Kluwer Academic, 171-180. [A comment on "The LSE Approach, A Progressive Approach to Econometric Modeling" by Grayham Mizon.]

Florens, J.-P. and Richard, J.-F. (1985) "Encompassing," C.O.R.E Discussion Paper, University of Louviain-la-Neuve.

Friedman, M. and Schwartz, A.J. (1991) "Alternative Approaches to Analyzing Economic Data," *American Economic Review* 81, 39-49.

Gilbert, C.L. (1986) "Professor Hendry's Econometric Methodology," Oxford Bulletin of Economics and Statistics 48, 283-307.

Godfrey, L.G. (1978) "Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables," Econometrica 46 (November), 1293-1301.

Goldfeld, S.M. (1976) "The Case of the Missing Money," *Brookings Papers on Economic Activity* 3, 683-730.

Goodfriend, M., and McCallum, B. (1992) "Demand for Money: Theoretical Studies," in P. Newman, M. Lilgate, and J. Eatwell, eds, *The New Palgrave Dictionary of Money and Finance,* 1992, pp. 611-617. New York: The Stockton Press.

Hendry, D.F. (1979) "Predictive Failure and Econometric Modelling in Macro-Economics: The Transactions Demand for Money," in P. Ormerod, ed., *Economic Modelling*, pp. 217-242. (Reprinted in Hendry, D.F., *Econometrics: Alchemy or Science?* Oxford: Blackwell.)

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

_______, (1987) "Econometric Methodology: A Personal Perspective," in T.F. Bewley, ed., *Advances in Econometrics*. Cambridge: Cambridge University Press.

_______, (1988) "The Encompassing Implications of Feedback Versus Feedforward Mechanisms in Econometrics," *Oxford Economic Papers,* 40, 132-149.

_______, (1993a), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

_______, (1993b) "The Roles of Economic Theory and Econometrics in Time-Series Economics," Invited address, European Econometric Society, Stockholm.

_______, (1995), *Dynamic Econometrics*. New York, Oxford University Press.

Hendry, D.F., and Doornik, J.A. (1994) "Modelling Linear Dynamic Econometric Systems," *Scottish Journal of Political Economy* 41, 1-33.

Hendry, D.F., and Ericsson, N.R. (1991a) "Modeling the Demand for Narrow Money in the United Kingdom and the United States," *European Economic Review* 35, 833-886.

______, (1991b) "An econometric analysis of U.K. money demand in monetary trends in the United States and the United Kingdom by Milton Friedman and Anna J. Schwartz," *American Economic Review* 81, 8-38.

Hendry, D.F., and Mizon, G.E. (1993) "Evaluating Dynamic Econometric Models by Encompassing the VAR," in P.C.B. Phillips, ed., *Models, Methods and Applications of Econometrics: Essays in Honor of A.R. Bergstrom*, 272-300, Cambridge, Basil Blackwell.

Hess, G.D.; Jones, C.S.; and Porter, R.D. (1997) "The Predictive Failure of the Baba, Hendry and Starr Model of M1," *Journal of Economics and Business*, forthcoming.

Hoover, K. and Perez, S. (l994) "Post Hoc Ergo Propter Hoc Once More: An Evaluation of `Does Monetary Policy Matter?' in the Spirit of James Tobin." *Journal of Monetary Economics*, 34, 47-73.

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

_______ (1992) "Testing Weak Exogeneity and the Order of Cointegration in U.K. Money Demand Data," *Journal of Policy Modeling* 14, 313-334.

Judd, J. and Scadding, J. (1982) "The Search for a Stable Money Demand Function: A Survey of the Post-1973 Literature," *Journal of Economic Literature* 20, 993-1023.

Judge, G.G.; Griffiths, W.E.; Hill, R.C.; Lütkepohl, H.; and Lee, T.-C. (1985), *The Theory and Practice of Econometrics*. New York: Wiley.

King, R., Plosser, C., Rebelo, S.(1988) "Production, growth and business cycles," *Journal of Monetary Economics*, 21:195-232.

King, R., Plosser, C., Stock, J., and Watson, M. (1991) "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81, 819--840.

Koopmans, T.C. (1953) "Identification Problems in Economic Model Construction," Chapter 2 in W.C. Hood and T.C. Koopmans, eds., *Studies in Econometric Method*. New York: Wiley.

Kydland, F. and E. Prescott. (1982) "Time to build and aggregate fluctuations," *Econometrica*, 50:1345-1370.

Leamer, E. (1983) "Let's take the con out of econometrics," *American Economic Review* 73, 31-43.

Leeper, E.M. (1995) "Commentary," on "The Lucas Critique in Practice: Theory Without Measurement," in K.D. Hoover, ed., *Macroeconometrics: Developments, Tensions and Prospects*, 313-324. Boston, Kluwer.

Leeper, E.M., and Sims, C.A. (1994) "Toward a Modern Macro Model Usable for Policy Analysis," *NBER Macroeconomics Annual 1994*, 81-117.

Leeper, E.M., Sims, C.A., and Zha, T. (1996) "What Does Monetary Policy Do?," *Brookings Papers on Economic Activity*, forthcoming.

Lucas, R.E., Jr. (1976) "Econometric Policy Evaluation: A Critique," in K. Brunner, and A. Meltzer, eds, *The Phillips Curve and Labor Markets*, Vol. 1 *of Carnegie-Rochester Conference Series on Public Policy*, pp. 19-46. Amsterdam: North-Holland.

McGrattan, E. (1994) "The Macroeconomic Effects of Distortionary Taxation," *Journal of Monetary Economics* 33, 573-601.

Mizon, G.E. (1984) "The Encompassing Approach in Econometrics," in D.F. Hendry and K.F. Wallis eds., *Econometrics and Quantitative Economics*, 135-172, Oxford, Basil Blackwell.

_______, (1995) "Progressive Modeling of Macroeconomic Time Series: The LSE Methodology," in K.D. Hoover, ed., *Macroeconometrics: Developments, Tensions and Prospects*, 107-170. Boston, Kluwer.

Neftci, S. and Sargent, T. (1978) "A little bit of Evidence on the Natural Rate Hypothesis from the U.S." *Journal of Monetary Economics* 4:315-319.

Oliner, S., Rudebusch, G., and Sichel, D. (1996) "The Lucas Critique Revisited: Assessing the Stability of Empirical Euler Equations for Investment," *Journal of Econometrics* 70:291-316.

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

Prescott, E. (1986) "Theory Ahead of Business Cycle Measurement," in K. Brunner and A. Meltzer, eds., *Real Buysiness Cycles, Real Exchange Rates and Actual Policies*, Vol. 25 of *Carnegie-Rochester Conference Series on Public Policy*, pp. 11-44. Amsterdam: North-Holland.

Romer, C. and Romer, D. (l989). "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz," in Blanchard, Olivier J. and Stanley Fischer, eds., *NBER Macroeconomics Annual l989*, 121-70, Cambridge, MIT.

Schwarz, G. (1978) "Estimating the Dimension of a Model," *Annals of Statistics* 6:461-464.

Shapiro, M., and Watson, M. (1988) “Sources of Business Cycle Fluctuations,” in S. Fisher, ed., *NBER Macroeconomics Annual 1988*, 111-148, Cambridge, MA.: MIT Press.

Sims, C.A. (1980) “Macroeconomics and Reality,” *Econometrica* 48, 1-48.

Sims, C.A. (1986) “Are Forecasting Models Useable for Policy Analysis?” Federal Reserve Bank of Minneapolis *Quarterly Review* (Winter):2-16.

Sims, C.A., and Zha, T. (1996) “Bayesian Methods for Dynamic Multivariate Models,” Federal Reserve Bank of Atlanta manuscript.

Stone, C.J. (1981) “Admissible Selection of an Accurate and Parsimonious Normal Linear Regression Model,” *Annals of Statistics* 9:475-485.

Watson, M. (1993) “Measures of Fit for Calibrated Models,” *Journal of Political Economy* 101:1011-41.

White, H. (1980) “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity,” *Econometrica* 48 (May):817-838.

White, H. (1990) “A Consistent Model Selection Procedure Based on *m*-testing,” in Granger, C.W.J., ed., *Modelling Economic Series*,369-383, Oxford: Clarendon Press.

Endnotes

[1] The VAR and RBC labels are also objectionable, since, for example, RBC-style tools are used to study monetary economies.

[2] As noted by Goodfriend and McCallum (1992), the route through which the Lucas Critique is usually operative involves costly adjustment considerations typically absent in models of money demand. More generally, changes in the nature of supply processes could affect the form of demand functions, and any econometric procedure must confront this possibility. Below, we shall see that there is reason to believe the LSE approach has not done so satisfactorily whether the context is money demand or a much broader one.

[3] Others have chosen to estimate the models and judge the fit by likelihood methods (Eichenbaum and Christiano, 1992; Braun, 1994; Leeper and Sims, 1994; McGrattan, 1994).

[4] Recall that $SV_t = \max(0,S_t)\times V_t$, where S_t is the spread between the 20-year Treasury bond yield and the coupon equivalent yield on a one-month Treasury bill; and V_t is a nine-quarter moving average of quarterly averages of twelve-month moving standard deviations of 20-year bond yields.

[5] Forecasters have long used a Bayesian prior that reduces the effective number of parameters in the model shrinking coefficients toward zero. Until the recent work of Sims and Zha (1996) and Leeper, Sims, and Zha (1996), this approach had not been used in identified VAR work.

[6] White's (1990) theorem on consistency of testing-based model-selection is sometimes raised as a justification for the LSE model selection criterion. In our view, this result--and any consistency result--is weak support, since it says nothing about finite-sample model selection. Moreover, it relies on strong assumptions about the stability of the underlying economic structure with which many researchers might not agree.

[7] In some papers, LSE'ers seem to make ad hoc adjustments to the critical values in these tests. For example, each *t*-statistic on a retained regressor is greater than 3 in BHS. This approach has not been codified, and would be difficult to codify, especially when one considers applying the approach across separate papers.

[8] The failure to test stability when it is central to the theory has also been criticized in the literature attempting to estimate "deep parameters" using generalized method of moments (Oliner, et al., 1995).

[9] Evidence that the study of business cycles in the RBC literature has missed essential business cycle features is found in Watson (1993) and Cogley and Nason (1995).

[10] The encompassing model might coincide with the theory constrained model, of course, but a model always encompasses itself.

[11] In drawing this conclusion, we relied not on whether LSE papers have an economic theory section; rather, we attempted to judge whether theory ultimately was a constraint on the final model reported.

---

[12] These equations may or may not involve contemporaneous endogenous variables.

[13] Another interpretation of this procedure is that at the first stage one imposes the just-identifying assumption that money demand is last in a recursive ordering of the economy (so simultaneity is not an issue). We do not favor this interpretation because this assumption is never stated directly, no economic basis for it is given, and it involves a covariance matrix restriction, which is frowned upon in the LSE school. Under this interpretation, the only question is whether the breaks are overidentifying; the remainder of our argument is essentially unaltered.

[14] Engle and Hendry (1993) lay out this system of testing allowing for much more general kinds of breaks. We limit discussion to intercept shifts for simplicity. Our argument can be formalized for the general types of changes in slope coefficients and variances.

[15] Hendry, 1995, p. 367: "In the process of reduction to a claimed sufficient representation, information can be discarded conditional on the validity of the reduction."

[16] The raw rate on sight deposits is multiplied by a two-parameter weighting function that rises smoothly with time after 1984(3). See Hendry and Ericsson (1991a).

[17] Many features of the model were formulated in an earlier version of the work using data through 1985.

[18] The citations search was performed through a Social Science Citation index search. We thank Cathy Tunis, a research librarian at the Federal Reserve Board, for much assistance.

[19] Our access to the Social Science Citation Index is through an on-line service that charges $1.40 per downloaded citation. Works such as Kydland and Prescott (1982) and King, Plosser, and Rebelo (1988) have received more citations than the chosen VAR and LSE papers combined and including these important RBC papers would have pushed the direct costs of the search well into four figures. Some would find this fact to be of central relevance to the section. At the very least, LSE'ers and VARs'ers who were not convinced by our argument against the relevance of gross citation counts may be more persuaded now.