Some comments on the role of econometrics in economic theory

Martin Eichenbaum

In this article, I offer some comments on the role of econometrics in macroeconomics.[1] These reflect a specific perspective: The role of econometrics ought to be the advancement of empirically plausible economic theory. This is a natural perspective for any economist to take, but it is one that is particularly compelling for a macroeconomist. Lucas' (1976) critique of econometric policy evaluation aside, it seems obvious that most policy questions cannot be fruitfully addressed using traditional quasi-reduced form econometric models. In the end, there are no alternatives to the use of fully specified general equilibrium models for addressing many of the problems that interest macroeconomists.

The real issue is: Different fully specified general equilibrium models can generate very different answers to the same question. Indeed it is possible to work backwards from *any answer* to *some model*. So given a particular question, which model should a macroeconomist use? Developing the tools to answer this question is the key challenge facing econometricians. Because all models are wrong along some dimension, the classic Haavelmo (1944) program of testing whether models are "true" will not be useful in meeting this challenge.[2] We do not need high-powered econometrics to tell us that models are false. We know that. What we need to know are the dimensions along which a given model does well and the dimensions along which it does poorly. In Leamer's (1978) terminology, we need a workable version of "specimetrics" that is applicable to dynamic general equilibrium models.[3] Developing the diagnostic tools for this specimetrics program ought to be the primary occupation of econometricians, not developing ever-increasingly sophisticated tools for implementing the Haavelmo program.

The need for progress on this front is pressing. General equilibrium business cycle analysts have begun to move beyond their initial practice of assessing models on a small set of moments without a formal statistical methodology.[4] Real business cycle (RBC) theory is evolving to accommodate a wide variety of impulses to the business cycle, including shocks to fiscal and monetary policy. But the process is in its infancy. The ultimate success of the enterprise will depend on the willingness of econometricians to devote more energy to the development of diagnostic tools for structural models and less to the development of estimators for the parameters of reduced form systems and increasingly powerful tests of null hypotheses, such as 'The model is a literal description of the data-generating mechanism'.

What is at stake for econometricians in all this? Why should they care about the needs of macroeconomists? Because, as social scientists,

Martin Eichenbaum is a professor of economics at Northwestern University and a senior consultant to the Federal Reserve Bank of Chicago. The author is grateful to Craig Burnside, Larry Christiano, John Cochrane, Ian Domowitz, Jonas Fisher, Lars Hansen, Joel Mokyr, and Tom Sargent for their comments.

their product has to meet a market test. There is no point in producing elegant merchandise that is buried in the inventory of advanced econometrics textbooks. Unfortunately, this happens all too often. To many young macroeconomists, econometrics seems irrelevant.[5] To remedy the situation, econometricians need to write instruction manuals for their products in a language that their customers understand. The language of economists centers on objects like agents' criterion functions, information sets, and constraints.[6] Consequently, econometricians need to focus their efforts on developing tools to obtain information about those objects. To the extent that they concentrate on analyzing the parameters of reduced form representations of the data or devising tests of whether specific structural models are false, their output is likely to be ignored by most of their (macro) colleagues.

This is not to suggest that there is no room for specialization in research or that econometricians should not engage in basic research and development. No one knows in advance which tools will be valuable in applied research. Still, the paradigm within which econometricians operate affects the types of tools they are likely to develop. The fact is that economists need to work with false structural models. It follows that econometricians need to abandon the Haavelmo paradigm and adopt one that more closely captures the ongoing dialogue between theory and data. [7]

Building confidence in models

Focusing on the task of evaluating the effects of alternative policy rules is one way to make concrete the ongoing interaction between theory and data that marks actual practice in macroeconomics. With data drawn from otherwise identical economies operating under different policy rules, we could easily dispense with economic theory. Such data are not available. And real world experimentation is not an option. We can perform experiments only in structural models. Indeed, Lucas (1980) argues that one of *the* critical functions of theoretical economics is to provide fully articulated economic systems. These systems can serve as laboratories in which policies that would be prohibitively expensive to experiment with in actual economies can be tested.

This sounds fine in principle. But *which* fully articulated economic system should we use? Lucas suggests that we

> test models as useful imitations of reality by subjecting them to shocks for which we are fairly certain how actual economies, or parts of economies, would react. The more dimensions on which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions. ("Methods and problems in business cycle theory," *Journal of Money, Credit and Banking*.)

The problem with this advice is that Lucas doesn't specify what "more" and "mimics" mean or how we are supposed to figure out the way an actual economy responds to an actual shock. But absent specificity, we are left wondering just how to build trust in the answers that particular models give us. In the remainder of this article, I discuss two strategies. One strategy uses exactly identified vector autoregressions (VARs) to derive the answers that actual economies give to a simple question and then to see if structural models reproduce that answer.[8] The specific simple question that VARs can sometimes answer is: How does the economy respond to an exogenous shock in agents' environments? A different strategy, the one most RBC analysts have pursued, is to focus on a model's ability to account for selected moments of the data, like variances and covariances, that they believe are useful for diagnostic purposes.

Identifying the effects of actual shocks to actual economies

Without observable exogenous variables, it is not easy to determine the answers that real economies give to even simple questions. Limited progress can be made by combining historical and institutional knowledge with exactly identified VARs to isolate empirical measures of shocks to the economy. Reduced-form VAR-based exercises cannot provide answers to hard questions like 'How would the economy react to a systematic change in the Federal Reserve's monetary policy rule?' That's because they are not well suited to investigating the effects of systematic changes in agents' constraint sets. But they can, in principle, answer simpler questions like 'What is the effect of an exogenous shock to the money supply?'

To the extent that complete behavioral models can reproduce answers that exactly identified VARs provide, we can have greater confidence in the behavioral models' answers to harder policy questions. Suppose, for example, that we want to use a particular structural model to assess the impact of a systematic change in the monetary authority's policy rule. A *minimal* condition we might impose is that the model be consistent, qualitatively and quantitatively, with the way short-term interest rates respond to shocks in the money supply.

To the extent that the answers from VAR-based exercises are robust to different identifying assumptions, they are useful as diagnostic devices. For example, different economic models make sharply different predictions about the impact of a shock to monetary policy. Both simple monetized RBC models and simple Keynesian models imply that interest rates ought to rise after an expansionary shock to the money supply. Limited participation models embodying strong liquidity effects imply that interest rates ought to fall.[9] Bernanke and Mihov (1995) and Pagan and Robertson (1995) review recent VAR-based research on what actually happens to interest rates after a shock to monetary policy. The striking aspect of these papers is how robust inference is across a broad array of restrictions: expansionary shocks to monetary policy drive short-term interest rates down, not up. This finding casts strong doubt on the usefulness of simple monetized RBC and Keynesian models for addressing a host of monetary policy issues.

Often, historical and institutional information can be very useful in sorting out the plausibility of different identifying schemes. Just because this information is not easily summarized in standard macro time series does not mean it should be ignored. Consider the task of obtaining a 'reasonable' measure of shocks to monetary policy. We know that broad monetary aggregates like M1 or M2 are not controlled by the Federal Reserve on a quarterly basis. So it makes no sense to identify unanticipated movements in M1 or M2 with shocks to monetary policy. Similarly, based on our knowledge of U.S. institutions, we may have very definite views about the effects of monetary policy on certain variables. For example, a contractionary monetary policy shock is clearly associated with a decrease in total government securities held by the Federal Reserve. A measure of monetary policy shocks that did not have this property would (and should) be dismissed as having incredible implications.

Does this mean that we should only use VARs to generate results that are consistent with what we already think we know? Of course not. In practice we build confidence in candidate shock measures by examining their effect on the variables that we have the strongest views about. In effect we 'test' the restrictions underlying our shock measures via sign and shape restrictions on the dynamic response functions of different variables to the shocks. When enough of these 'tests' have been passed, we have enough confidence to use the shock measure to obtain answers to questions we don't already know the answers to.[10] To my knowledge, econometricians have not yet provided a formal Bayesian interpretation for this procedure. Such a framework would be extremely valuable to practitioners.

How well does a model mimic a data moment?

Another strategy for building confidence in models is to see whether they account for prespecified moments of the data that are of particular interest to economic model builders. This strategy is the one pursued by most RBC analysts. In so doing, they have made little use of formal econometric methods, either when model parameters are selected, or when the model is compared to the data. Instead a variety of informal techniques, often referred to as calibration, are used.

A key defect of calibration techniques is that they do not quantify the sampling uncertainty inherent in comparisons of models and data. Calibration rhetoric aside, model parameter values are not known. They have to be estimated. As a result, a model's predictions are random variables. Moreover, the data moments that we are trying to account for are not known. They too have to be estimated. Without some way of quantifying sampling uncertainty in these objects, it is simply impossible to say whether the moments of a fully calibrated model are "close" to the analog moments of the data-generating process. In the end, there is no way to escape the need for formal econometric methodology.

Do the shortcomings of calibration techniques affect inferences about substantive claims being made in the literature? Absolutely.

The claim that technology shocks account for a given percent, say λ, of the variance of output amounts to the claim that a calibrated model generates a value of λ equal to

$$\hat{\sigma}^2_{yn}(\hat{\Psi}_1)/\hat{\sigma}^2_{yd}.$$

Here the numerator denotes the variance of model output, calculated under the assumption that the vector of model structural parameters, Ψ_1, equals $\hat{\Psi}_1$ while the denominator denotes an estimate of the variance of actual output. The claim that technology shocks account for most of the fluctuations in postwar U.S. output corresponds to the claim that λ is close to one.[11]

In reality, Ψ_1 and the actual variance of output, σ^2_{yd}, have to be estimated. Consequently, λ is a random variable. Eichenbaum (1991) investigated the extent of the sampling uncertainty associated with estimates of λ. My conclusion was that the extent of this uncertainty is enormous.[12] The percentage of aggregate fluctuations that technology shocks actually account for could be 70 percent as Kydland and Prescott (1989) claim but it could also be 5 percent or 200 percent. Under these circumstances, it is very hard to attach any importance to the point estimates of λ pervading the literature.

There are a variety of ways to allow for sampling uncertainty in analyses of general equilibrium business cycle models. The most obvious is to use maximum likelihood methods.[13] A shortcoming of these methods is that the estimation criterion weights different moments of the data, exclusively according to how much information the data contain about those moments. At a purely statistical level, this is very sensible. But as decisionmakers we may disagree with that ranking. We may insist on allocating more weight to some moments than others, either at the estimation or at the diagnostic stage. Different approaches for doing this have been pursued in the literature.

Christiano and Eichenbaum (1992) use a variant of Hansen's (1982) generalized method of moments (GMM) approach to estimate and assess business cycle models using prespecified first and second moments of the data. Ingram and Lee (1991) discuss an approach for estimating parameter values that minimizes the second-moment differential of the actual data and the artificial data generated by the model. Diebold, Ohanian, and Berkowitz (1994) propose frequency domain analogs, in which the analyst specifies the frequencies of the data to be used at the estimation and diagnostic stages of the analysis. King and Watson (1995) pursue an approach similar in spirit to those mentioned above but geared more toward assessing the relative adequacy of competing models with respect to prespecified features of the data.

These approaches share two key features. First, the analyst has the option of using different features of the data for estimation and diagnostic purposes. Second, standard econometric methodology is used to provide information about the extent of uncertainty regarding differences between the model and the data, at least as these reflect sampling error. In principle, the first key feature differentiates these approaches from maximum likelihood approaches. In practice, it is easy to overstate the importance of this difference. In actual applications, we have to specify which variables' likelihood surface we are trying to match. So there is nothing particularly general or comprehensive about maximum likelihood methods in particular applications, relative to the approaches discussed above.

Still, the more moments an approach uses to diagnose the empirical performance of a model, the more general that approach is. An important shortcoming of many RBC studies (including some that I have conducted) is that they focus on a *very* small subset of moments. Some of the most interesting diagnostic work being done on general equilibrium business cycle models involves confronting them with carefully chosen but ever-expanding lists of moments. The evolution of RBC models beyond their humble beginnings parallels the wider range of phenomena that they are now being confronted with.

To illustrate this point, I now consider some of the strengths and weaknesses of a simple, prototypical RBC model. Using the approach discussed in Christiano and Eichenbaum (1992), I show that the model does very well with respect to the standard small list of moments initially used to judge RBC models. I then use this approach to display a point made by Watson (1993): Standard RBC models badly miss capturing the basic spectral shape of real macroeconomic variables, particularly real output. This reflects the virtual absence of any propagation mechanisms in these models. Model diagnostic approaches that focus on a small set of moments like the variance of output and employment mask this first-order failure.

A simple RBC model

Consider the following simple RBC model. The model economy is populated by an infinitely lived representative consumer who maximizes the criterion function

$$1)\quad E_0 \sum_{t=0}^{\infty} \beta^t\,[\ln(C_t) - \theta N_t].$$

Here $0 < \beta < 1$, $\theta > 0$, C_t denotes time t consumption, N_t denotes time t hours of work, and E_0 denotes expectations conditioned on the time 0 information set.

Time t output, Y_t, is produced via the Cobb-Douglas production function

$$2)\quad Y_t = K_t^{1-\alpha}(N_t X_t)^{\alpha},$$

where the parameter α is between 0 and 1, K_t denotes the beginning of time t capital stock, and X_t represents the time t level of technology. The stock of capital evolves according to

$$3)\quad K_{t+1} = (1 - \delta)K_t + I_t.$$

Here I_t denotes time t gross investment and $0 < \delta < 1$. The level of technology, X_t, evolves according to

$$4)\quad X_t = X_{t-1} \exp(\gamma + \upsilon_t),$$

where $\gamma > 0$, υ_t is a serially uncorrelated process with mean 0 and standard deviation σ_υ. Notice that unlike the class of models examined in Eichenbaum (1991), the level of technology is modeled here as a difference stationary stochastic process. The aggregate resource constraint is given by

$$5)\quad C_t + I_t + G_t \leq Y_t.$$

Here G_t denotes the time t level of government consumption which evolves according to

$$6)\quad G_t = X_t g_t^*.$$

The variable g_t^* is the stationary component of government consumption and $g_t = 1\text{n}(g_t^*)$ evolves according to

$$7)\quad g_t = g_0 + g_1 t + \rho g_{t-1} + \varepsilon_t,$$

where g_0 and g_1 are constants, t denotes time, $|\rho| < 1$, and ε_t is a mean zero shock to g_t that is serially uncorrelated and has standard deviation σ_ε. The variable ρ controls the persistence of g_t. The larger ρ is, the longer lasting is the effect of a shock to ε_t on g_t.

In the presence of complete markets, the competitive equilibrium of this economy corresponds to the solution of the social planning problem: Maximize equation 1 subject to equations 2 to 7 by choice of contingency plans for time t consumption, hours of work, and the time t+1 stock of capital as a function of the planner's time t information set. This information set is assumed to include all model variables dated time t and earlier.

Burnside and Eichenbaum (1994) estimate the parameters of this model using the GMM procedure described in Christiano and Eichenbaum (1992). To describe this procedure, let Ψ_1 denote the vector of model structural parameters. The unconditional moment restrictions underlying Burnside and Eichenbaum's estimator of Ψ_1 can be summarized as:

$$8)\quad E\,[u_{1t}(\Psi_1^0)] = 0,$$

where Ψ_1^0 is the true value of Ψ_1 and $u_{1t}(\bullet)$ is a vector-valued function that depends on the data as well as Ψ_1^0. In Burnside and Eichenbaum's (1994) analysis, the dimension of $u_{1t}(\bullet)$ is the same as that of Ψ_1^0. Because of this, the moment restrictions in equation 8 fall into two categories. The first category consists of conditions that require the model to match the sample analogs of various moments of the data, like the capital to output ratio, and average hours worked. The second category consists of conditions that lead to estimating parameters like those governing the behavior of government purchases, ρ, g_0, and g_1, via least squares, and parameters like the standard deviations of the shock to technology and government purchases, as the sample averages of the sums of squared fitted residuals.

Two features of equation 8 are worth noting. First, there is no reason to view this equation as holding only under the hypothesis that the model is "true". Instead equation 8 can be viewed as summarizing the rule by which Burnside and I chose model parameter values as functions of unknown moments of the data-generating process. Second, our model is one of balanced growth. This, in conjunction with our specification of the technology process, X_t, as a difference stationary process, implies a variety of cointegrating relationships among the variables in the model.[14] We exploit these relationships to ensure that the moments entering equation 8 pertain to stationary stochastic processes.

The salient features of the parameter estimates reported in Burnside and Eichenbaum (1994) is their similarity to the values employed in existing RBC studies. So what differentiates the estimation methodology is not the resulting point estimates, but that the approach allows one to translate sampling uncertainty about the functions of the data that define the parameter estimator into sampling uncertainty regarding point estimates.

The procedure used to assess the empirical plausibility of the model can be described as follows. Let Ψ_2 denote a vector of diagnostic moments that are to be estimated in ways not involving the model. The elements of Ψ_2 typically include objects like the standard deviations of different variables, as well as various autocorrelation and cross-correlation coefficients. The unconditional moment restrictions used to define the GMM estimator of Ψ_2 can be summarized as:

9) $E\,[u_{2t}(\Psi_2^0)] = 0.$

Here Ψ_2^0 denotes the true value of Ψ_2. The vector $u_{2t}(\bullet)$ has the same dimension as Ψ_2^0. It is useful to summarize equations 8 and 9 as

10) $E\,[u_t(\Psi^0)] = 0 \quad t = 1, \ldots, T.$

Here Ψ^0 is the true value of $(\Psi_1' \Psi_2')'$ and u_t is a vector valued function of dimension equal to the dimension of Ψ^0. As long as the dimension of $u_t(\bullet)$ is greater than or equal to the dimension of Ψ^0, equation 10 can be exploited to consistently estimate Ψ^0 via Hansen's (1982) GMM procedure.

Suppose we wish to assess the empirical plausibility of the model's implications for a $q \times 1$ subset of Ψ_2. We denote this subset by ω. Let $\Phi(\Psi)$ denote the value of ω implied by the model, given the structural parameters Ψ_1. Here Φ denotes the (nonlinear) mapping between the model's structural parameters and the relevant population moments. Denote the nonparametric estimate of ω obtained without imposing restrictions from the model by $\Gamma(\Psi)$. Then hypotheses of the form

11) $H_0 : F(\Psi^0) = \Phi(\Psi^0) - \Gamma(\Psi^0) = 0$

can be tested using a simple Wald test.

Early RBC studies often stressed the ability of the standard model to account for the volatility of output and the relative volatility of various economy-wide aggregates. To examine this claim, it is useful to focus for now on the standard deviation of output, the standard deviation of consumption, investment, and hours worked relative to output, and the standard deviation of hours worked relative to average productivity.[15] Column 1 of table 1 lists different moments of the data. Column 2 reports nonmodel-based point estimates of these moments, obtained using aggregate time-series data covering the period 1955:Q3–84:Q4. Column 3 contains the values of these moments implied by the model, evaluated at Ψ_1.

TABLE 1

Data and model moments

(Relative volatility tests[a])

Moment	U.S. data	Model
σ_y	0.0192	0.0183
	(0.0021)	(0.0019)
		[0.712]
σ_c/σ_y	0.437	0.453
	(0.034)	(0.005)
		[0.633]
σ_i/σ_y	2.224	2.224
	(0.079)	(0.069)
		[0.999]
σ_h/σ_y	0.859	0.757
	(0.080)	(0.050)
		[0.999]
σ_h/σ_{apl}	1.221	1.171
	(0.132)	(0.032)
		[0.729]

[a]The statistic σ_i is the standard deviation of the Hodrick-Prescott filtered variable *i, i = y* (output), *c* (consumption), *i* (investment), *h* (hours worked), and *apl* (average productivity of labor).

Notes: Numbers in parentheses denote the standard error of the corresponding point estimate. Numbers in brackets denote the probability values of the Wald statistics for testing the hypothesis that the model and nonmodel-based numbers are the same in population.

Source: This table is taken from Burnside and Eichenbaum (1994).

Numbers in parentheses are the standard errors of the corresponding point estimates. Numbers in brackets are the probability values of Wald statistics for testing whether the model and data moments are the same in population. The key thing to notice is how well the model performs on these dimensions of the data. In no case can we reject the individual hypotheses that were investigated, at a conventional significance level.

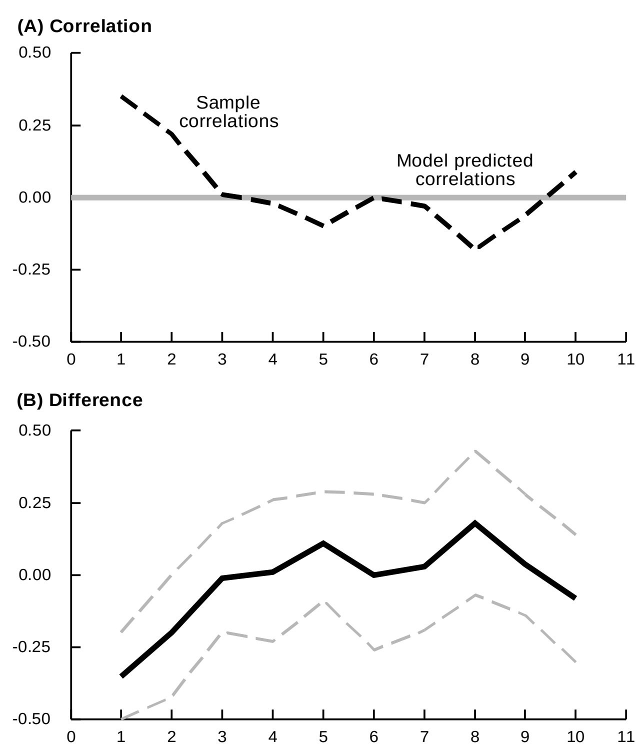
Once we move beyond the small list of moments stressed in early RBC studies, the model does not perform nearly as well. As I mentioned above, Watson (1993) shows that the model fails to capture the typical spectral shape of growth rates for various macro variables. For example, the model predicts that the spectrum of output growth is flat, with relatively little power at cyclical frequencies. This prediction is inconsistent with the facts. A slightly different way to see this empirical shortcoming is to proceed as in Cogley and Nason (1993) and focus on the autocorrelation function of output growth. Panel A of figure 1 reports nonmodel-based estimates of the autocorrelation function of $\Delta \ln (Y_t)$, as well as those implied by the model. These are depicted by the solid and dotted lines, respectively. The actual growth rate of U.S. output is positively autocorrelated: specifically the first two autocorrelation coefficients are positive and significant.[16] The model implies that all the autocorrelations are negative, but small. In fact they are so close to zero that the solid line depicting them is visually indistinguishable from the horizontal axis of the figure. Panel B displays the difference between the model and nonmodel-based estimates of the autocorrelation coefficients, as well as a two-standard error band around the differences. We can easily reject the hypothesis that these differences reflect sampling error.

Various authors have interpreted this empirical shortcoming as reflecting the weakness of the propagation mechanisms embedded within standard RBC models. Basically what you put in (in the form of exogenous shocks) is what you get out. Because of this, simple RBC models cannot simultaneously account for the time-series properties of the growth rate of output and the growth rate of the Solow residual, the empirical measure of technology shocks used in first generation RBC models.

How have macroeconomists responded to this failing? They have *not* responded as Haavelmo (1944) anticipated. Instead they have tried to learn from the data and modify the models. The modifications include allowing for imperfect competition and internal increasing returns to scale, external increasing returns to scale, factor hoarding, multiple sectors with nontrivial input-output linkages, and monetary frictions.[17] Evidently when econometricians convey their results in language that is interpretable to theorists, theorists do respond. Progress is being made. Granted, the econometric tools described here fall far short of even approximating the dynamic version

FIGURE 1

Autocorrelations of output growth



Note: In the Difference panel the dashed lines represent a two-standard error band around the difference.

of Leamer-style specimetrics discussed in the introduction. Still, they have proved to be useful in practice.

Conclusion

I would like to conclude with some comments about the classic Haavelmo program for testing economic models. I did not discuss this program at length for a simple reason: It is irrelevant to the inductive process by which theory actually evolves. In his seminal 1944 monograph, Haavelmo conceded that his program contributes nothing to the construction of economic models. The key issue he chose to emphasize was

> the problem of splitting on the basis of data, all *a priori* theories about certain variables into two groups, one containing the admissible theories, the other containing those that must be rejected. ("The probability approach in econometrics," *Econometrica*)

In reality, economic hypotheses and models are generated by the ongoing interaction of researchers with nonexperimental data. The Haavelmo program conceives of economic theorists, unsullied by data, working in splendid isolation, and "somehow" generating hypotheses. Only when these hypotheses appear, does the econometrician enter. Armed with an array of tools he goes about his grim task: testing and rejecting models. This task complete, the econometrician returns to the laboratory to generate ever-increasingly powerful tools for rejecting models. The theorist, no doubt stunned and disappointed to find that his model is false, returns to his office and continues his search for the "true" model.

I cannot imagine a paradigm more at variance with the way actual empirical research occurs. Theories don't come from a dark closet inhabited by theorists. They emerge from an ongoing dialogue with nonexperimental data or, in Leamer's (1978) terminology, from ongoing specification searches. To the extent that the Haavelmo program is taken seriously by anyone, it halts the inductive process by which actual progress in economics occurs.

The fact is that when Haavelmo attacked a real empirical problem, the determinants of investment, he quickly jettisoned his methodological program. Lacking the tools to create a stochastic model of investment, Haavelmo (1960) still found it useful to interact with the data using a "false" deterministic model. Fortunately, economic theory has progressed to the point where we do not need to confine ourselves to deterministic models. Still we will always have to make simplifying assumptions. In his empirical work, Haavelmo (1960) tried to help us decide which simplifying assumptions lead us astray. That is the program econometricians need to follow, not the utopian program that was designed in isolation from actual empirical practice. That road, with its focus on testing whether models are true, means abandoning econometrics' role in the inductive process. The results would be tragic, for both theory and econometrics.

NOTES

[1]This article is based on a paper that appeared in the November 1995 issue of *Economic Journal*.

[2]See Conclusion for further discussion of the Haavelmo program.

[3]By specimetrics, Leamer (1978, p. v) means: ". . . the process by which a researcher is led to choose one specification of the model rather than another; furthermore, it attempts to identify the inferences that may be properly drawn from a data set when the data-generating mechanism is ambiguous."

[4]Moments refer to certain characteristics of the data-generating process, such as a mean or variance. Moments are classified according to their order. An example of a first-order moment would be the expected value of output. An example of a second-order moment would be the variance of output.

[5]Some of the rhetoric in the early RBC literature almost suggests that econometricians and quantitative business cycle theorists are natural enemies. This view is by no means unique to RBC analysts. See for example Keynes' (1939) review of Tinbergen's (1939) report to the League of Nations and Summers' (1991) critique of econometrics.

[6]Econometricians have many customers, such as government officials and private businesses, for whom the language of economic theory may not be very useful.

[7]If these comments sound critical of econometricians who ignore economic theory, I have been as critical, if not more so, of business cycle theorists who ignore econometrics. See Eichenbaum (1991) for a discussion of the sensitivity of inference in the RBC literature to accounting for sampling uncertainty in the parameter estimates of structural models.

[8]A finite-ordered vector autoregressive representation for a set of variables Z_t expresses the time t value of each variable in Z_t as a function of a finite number of lags of all the variables in Z_t plus a white noise error term. The error term is often interpreted as a linear combination of the basic shocks affecting the economy. These shocks include unanticipated changes in monetary and fiscal policy. Exactly identified VARs make just enough assumptions to allow the analyst to measure the shocks from the error terms in the VAR. These assumptions are referred to as identifying assumptions.

[9]The key feature of limited participation models is the assumption that households do not immediately adjust their portfolios after an open market operation. Consequently, open market operations affect the bank-reserves portion of the monetary base. It is this effect that generates declines in interest rates following contractionary open market operations in the model. See King and Watson (1995) and Christiano and Eichenbaum (1995), as well as the references therein.

[10]See for example Christiano, Eichenbaum, and Evans (1996), who use this strategy to study the response of the borrowing and lending activities of different sectors of the economy to a shock in monetary policy.

[11]See for example Kydland and Prescott (1989).

[12]This conclusion depends on the nature of the estimators of Ψ_1 and σ^2_{yd} implicit in early RBC studies and the hypothesis that the level of technology is a trend-stationary process. The latter is an important maintained assumption of early RBC studies.

[13]See for example Leeper and Sims (1994) and the references therein.

[14]With the exception of hours worked, all model variables inherit a stochastic trend from the technology process, X_t.

[15]Here all moments refer to moments of time series that have been processed using the stationary inducing filter discussed in Hodrick and Prescott (1980).

[16]See Burnside and Eichenbaum (1994).

[17]This is a good example of Leamer's (1978) observation that a critical feature of many real learning exercises is the search for *new* hypotheses that explain the given data.

REFERENCES

Bernanke, B., and I. Mihov, "Measuring monetary policy," Princeton University, unpublished paper, 1995.

Burnside, C., and M. Eichenbaum, "Factor hoarding and the propagation of business cycle shocks," National Bureau of Economic Research, working paper, No. 4675, 1994.

Christiano, L. J., and M. Eichenbaum, "Current real business cycle theories and aggregate labor market fluctuations," *American Economic Review*, Vol. 82, June 1992, pp. 430–450.

_______________, "Liquidity effects, monetary policy, and the business cycle," *Journal of Money, Credit and Banking*, Vol. 27, No. 4, November 1995.

Christiano, L. J., M. Eichenbaum, and C. Evans, "What happens after a monetary shock? Evidence from the flow of funds," *Review of Economics and Statistics*, 1996 forthcoming.

Cogley, T., and J. Nason, "Do real business cycle models pass the Nelson-Plosser test?" University of British Columbia, unpublished paper, 1993.

Diebold, F. X., L. Ohanian, and J. Berkowitz, "Dynamic equilibrium economies: A framework for comparing models and data," University of Pennsylvania, unpublished paper, 1994.

Eichenbaum, M., "Real business cycle theory: Wisdom or whimsy?" *Journal of Economic Dynamics and Control*, Vol. 5, July/October 1991, pp. 607–626.

Haavelmo, T., "The probability approach in econometrics," *Econometrica*, Vol. 12, supplement, 1944, pp. 1–118.

_______________, *A Study in the Theory of Investment*, Chicago: University of Chicago Press, 1960.

Hansen, L. P., "Large sample properties of generalized method of moments estimators," *Econometrica*, Vol. 50, July 1982, pp. 1029–1054.

Hodrick, R. J., and E. C. Prescott, "Post-war business cycles: An empirical investigation," Carnegie-Mellon University, unpublished paper, 1980.

Ingram, B., and B. S. Lee, "Simulation estimation of time-series models," *Journal of Econometrics*, Vol. 47, February/March 1991, pp. 197–206.

Keynes, J.M., "Comment," *Economic Journal*, Vol. 44, September 1939, pp. 558–568.

King, R. G., and M. Watson, "Money, prices, interest rates, and the business cycle," University of Virginia, unpublished paper, 1995.

Kydland, F. E., and E. C. Prescott, "Hours and employment variation in business cycle theory," Institute for Empirical Economics, Federal Reserve Bank of Minneapolis, discussion paper, No. 17, 1989.

Leamer, E. E., *Specification Searches. Ad Hoc Inferences with Nonexperimental Data*, New York: John Wiley and Sons, 1978.

Leeper, E., and C. Sims, "Toward a modern macroeconomic model usable for policy analysis," in *NBER Macroeconomics Annual 1994*, Stanley Fischer and Julio J. Rotemberg (eds.), 1994, pp. 81–117.

Lucas, R. E., "Econometric policy evaluation: A critique," in *The Phillips and Labor Markets*, Karl Brunner and Alan H. Meltzer (eds.), Carnegie Rochester Conference Series on Public Policy, Vol. 1, Amsterdam: North-Holland Publishing Company, 1976, pp. 19–46.

Lucas, R. E., Jr., "Methods and problems in business cycle theory," *Journal of Money, Credit and Banking*, Vol. 12, November 1980, pp. 696–715.

Pagan, A., and John Robertson, "Resolving the liquidity effect," *Federal Reserve Bank of St. Louis Review*, Vol. 77, No. 3, May/June 1995, pp. 33–54.

Summers, L. H., "The scientific illusion in empirical macroeconomics," *Scandinavian Journal of Economics*, Vol. 93, June 1991, pp. 129–48.

Tinbergen, J., *A Method and Its Application to Investment Activity*, Statistical Testing of Business Cycle Theories, Geneva: League of Nations, 1939.

Watson, M., "Measures of fit for calibrated models," *Journal of Political Economy*, Vol. 101, December 1993, pp. 1011–1041.