Facts and Artifacts: Calibration and the Empirical Assessment of Real-Business-Cycle Models

Kevin D. Hoover

Oxford Economic Papers, New Series, Vol. 47, No. 1. (Jan., 1995), pp. 24-44.

Stable URL:
http://links.jstor.org/sici?sici=0030-7653%28199501%292%3A47%3A1%3C24%3AFAACAT%3E2.0.CO%3B2-5

Oxford Economic Papers is currently published by Oxford University Press.

Oxford Economic Papers 47 (1995), 24–44

FACTS AND ARTIFACTS: CALIBRATION AND THE EMPIRICAL ASSESSMENT OF REAL-BUSINESS-CYCLE MODELS

By KEVIN D. HOOVER

Department of Economics, University of California, Davis, California 95616-8578, USA

1. Whither quantitative macroeconomics?

THE RELATIONSHIP between theory and data has been, from the beginning, a central concern of the new-classical macroeconomics. This much is evident in the title of Robert E. Lucas's and Thomas J. Sargent's landmark edited volume, *Rational Expectations and Econometric Practice* (1981). With the advent of real-business-cycle models, many new classical economists have turned to calibration methods. The new classical macroeconomics is now divided between calibrators and estimators. But the debate is not a parochial one, raising, as it does, issues about the relationships of models to reality and the nature of econometrics that should be important to every school of macroeconomic thought, indeed to all applied economics. The stake in this debate is the future direction of quantitative macroeconomics. It is, therefore, critical to understand the root issues.

Lucas begins the second chapter of his *Models of Business Cycles* with the remark:

> Discussions of economic policy, if they are to be productive in any practical sense, necessarily involve *quantitative* assessments of the way proposed policies are likely to affect resource allocation and individual welfare. (Lucas 1987, p. 6; emphasis added)

This might appear to be a clarion call for econometric estimation. But appearances are deceiving. After mentioning Sumru Altug's (1989) estimation and rejection of the validity of a variant of Finn E. Kydland and Edward C. Prescott's (1982) real-business-cycle model (a model which takes up a large portion of his book), Lucas writes:

> ... the interesting question is surely not whether [the real-business-cycle model] can be accepted as 'true' when nested within some broader class of models. Of course the model is not 'true': this much is evident from the axioms on which it is constructed. We know from the onset in an enterprise like this (I would say, in *any* effort in positive economics) that what will emerge—at *best*—is a workable approximation that is useful in answering a limited set of questions. (Lucas 1987, p. 45)

Lucas abandons not only truth but also the hitherto accepted standards of empirical economics. Models that clearly do not fit the data, he argues, may nonetheless be calibrated to provide useful quantitative guides to policy.

Calibration techniques are commonly applied to so-called 'computable general-equilibrium' models. They were imported into macroeconomics as a

means of quantifying real-business-cycle models, but now have a wide range of applications. Some issues raised by calibration are common to all computable general-equilibrium models; the concern of this paper, however, is with real-business-cycle models and related macroeconomic applications; and, as will appear presently, these raise special issues. A model is calibrated when its parameters are quantified from casual empiricism or unrelated econometric studies or are chosen to guarantee that the model precisely mimics some particular feature of the historical data. For example, in Kydland and Prescott (1982), the coefficient of relative risk aversion is justified on the basis of microeconomic studies, while the free parameters of the model are set to force the model to match the variance of GNP without any attempt to find the value of empirical analogues to them.

Allan W. Gregory and Gregor W. Smith (1991, p. 3) conclude that calibration '... is beginning to predominate in the quantitative application of macroeconomic models'. While indicative of the importance of the calibration methodology, Gregory and Smith's conclusion is too strong. Aside from the new classical school, few macroeconomists are staunch advocates of calibration. Within the new classical school, opinion remains divided. Even with reference to real-business-cycle models, some practitioners have insisted that calibration is at best a first step, which must be followed '... by setting down a metric (e.g. one induced by a likelihood function) and estimating parameters by finding values that make the metric attain a minimum' (Gary Hansen and Sargent 1988, p. 293).[1]

Sargent advocates estimation or what Kydland and Prescott (1991) call the 'system-of-equations approach'. Estimation has been the standard approach in macroeconometrics for over 40 years. Sargent and like-minded new classical economists modify the standard approach only in their insistence that the restrictions implied by dynamic-optimization models be integrated into the estimations. The standard of empirical assessment is the usual one: how well does the model fit the data statistically? Lucas and Kydland and Prescott reject statistical goodness of fit as a relevant standard of assessment. The issue at hand might then be summarized: who is right—Lucas and Kydland and Prescott, or Sargent?

The answer to this question is not transparent. Estimation is the *status quo*. And, although enthusiastic advocates of calibration already announce its triumph, its methodological foundations remain largely unarticulated. An uncharitable interpretation of the calibration methodology might be that the advocates of real-business-cycle models are so enamored of their creations that they would prefer to abandon commonly accepted, neutral standards of empirical evaluation (i.e. econometric hypothesis testing) to preserve their

[1] Despite the joint authorship of the last quotation, I regard Sargent and not Hansen as the preeminent proponent of the necessity of estimation, because I recall him forcefully insisting on it in his role as discussant of a paper by Thomas F. Cooley and Hansen (1989) at the Federal Reserve Bank of San Francisco's Fall Acacemic Conference; see also Manuelli and Sargent (1988, pp. 531–4).

models. This would be an *ad hoc* defensive move typical of a degenerating research program.

This interpretation is not only uncharitable, it is wrong. Presently, we shall see that Herbert Simon's (1969) *Sciences of the Artifical* provides the materials from which to construct a methodological foundation for calibration, and that calibration is compatible with a well-established approach to econometrics that is nonetheless very different from the Cowles Commission emphasis on the estimation of systems of equations. Before addressing these issues, however, it will be useful to describe the calibration methodology and its place in the history and practice of econometrics in more detail.

2. The calibration methodology

2.1. *The paradigm case*

Kydland and Prescott (1982) is the paradigm new-classical equilibrium, real-business-cycle model. It is neoclassical optimal-growth model with stochastic shocks to technology which cause the equilibrium growth path to fluctuate about its steady state.[2] Concrete functional forms are chosen to capture some general features of business cycles. Production is governed by a constant-elasticity-of-substitution production function in which inventories, fixed capital, and labor combine to generate a single homogeneous output that may either be consumed or reinvested. Fixed capital requires a finite time to be built before it becomes a useful input. The constant-relative-risk-aversion utility function is rigged to possess a high degree of intertemporal substitutability of leisure. Shocks to technology are serially correlated. Together the structure of the serial correlation of the technology shocks and the degree of intertemporal substitution in consumption and leisure choices govern the manner in which shocks are propagagated through time and the speed of convergence back towards the steady state.

Once the model is specified, the next step is to parameterize its concrete functional forms. Most of the parameters of the model are chosen from values culled from other applied econometric literatures or from general facts about national-income accounting. For example, Thomas Mayer (1960) estimated the average time to construct complete facilities to be 21 months; Robert E. Hall (1977) estimated the average time from start of projects to beginning of production to be two years. Citing these papers, but noting that consumer durable goods take considerably less time to produce, Kydland and Prescott (1982, p. 1361) assume that the paramters governing capital formation are set to imply steady construction over four quarters.[3] The values for depreciation

[2] For general descriptions of the details and varieties of real-business-cycle models see Lucas (1987), Kevin D. Hoover (1988, ch. 3), and Bennett T. McCallum (1989). Steven M. Sheffrin (1989, ch. 3, especially pp. 80, 81), gives a step-by-step recipe for constructing a prototypical real-business-cycle model.

[3] Mayer's estimates were for complete projects only, so that new equipment installed in old plants, which must have a much shorter time-to-build than 21 months, was not counted.

rates and the capital/inventory ratio are set to rough averages from the national-income accounts. Ready estimates from similar sources were not available for the remaining six paramters of the model, which include parameters governing intertemporal substitution of leisure and the shocks to technology. These were chosen by searching over possible parameter values for a combination that best reproduced certain key variances and covariances of the data. In particular, the technology shock variance was chosen in order to exactly match the variance of output in the postwar US economy.

To test the model's performance, Kydland and Prescott generate a large number of realizations of the technology shocks for 118 periods corresponding to their postwar data. They then compute the variances and covariances implied by the model for a number of important variables: output, consumption, investment, inventories, the capital stock, hours worked, productivity, and the real rate of interest.[4] These are then compared with the corresponding variances and covariances of the actual US data.[5]

Kydland and Prescott offer no formal measure of the success of their model. They do note that hours are insufficiently variable with respect to the variability of productivity to correspond accurately to the data, but otherwise they are pleased with the model's ability to mimic the second moments of the data.

Real-business-cycle models, treated in the manner of Kydland and Prescott, are a species of the genus computable (or applied) general-equilibrium models. The accepted standards for implementing computable general-equilibrium models, as codified, for example, in Ahsan Mansur and John Whalley (1984), do not appear to have been adopted in the real-business-cycle literature. For example, while some practitioners of computable general-equilibrium models engage in extensive searches of the literature in order to get some measure of the central tendency of assumed elasticities, Kydland and Prescott's (1982)

[4] In fact, it is not clear in Kydland and Prescott (1982) that these are calculated from the cross-section of a set of realizations or from a single time-series realization. In a subsequent paper that extends their results, Kydland and Prescott (1988, p. 353) are quite precise about using a cross-section of many realizations. Because they are interested only in the moments of the variables and not in particular time-series, Kydland and Prescott initialize variables to their steady-state values or, equivalently in the context of detrended data, to zero. In order to generate a time path that can be compared to the history of a particular series, it is necessary, as in Hansen and Prescott (1993), to initialize at some actual historical benchmark.

[5] Despite my referring to Kydland and Prescott's model as a growth model, the model for which they calculate the variances and covariances does not possess an exogenous source of trend growth. Thus, to make comparisons, Kydland and Prescott (1982, p. 1362) detrend the actual data using the Hodrick-Prescott filter. The particular choice of filter is not defended in any detail. Prescott (1983, p. 6) simply asserts that it produces 'about the right degree of smoothness, when fit to the logarithm of the real GNP series' without any indication by what standard rightness is to be judged. Kydland and Prescott (1990, p. 9) claim that it generates a trend close to the trend that students of business cycles would draw by hand through a plot of actual GNP. Although the Hodrick-Prescott filter is almost universally adopted in comparing real-business-cycle models to actual data, Fabio Canova (1991b) shows that the use of Hodrick-Prescott filters with different choices for the values of a key parameter or of several entirely different alternative filters radically alters the cyclical characteristics of economic data (also see Timothy Cogley and James Nason, 1993).

choice of parameterization appears almost casual. Similarly, although Kydland and Prescott report some checks on robustness, these appear to be perfunctory.[6]

In the context of computable general-equilibrium models, calibration is preferred in those cases in which, because of extensive disaggregation, the number of parameters is too large relative to the available data set to permit econometric estimation.[7] Since typical real-business-cycle models are one-good, one-agent models, there is no difficulty in estimating them using standard methods such as maximum likelihood or generalized method of moments. Indeed, since the practitioners are often interested principally in matching selected second moments, method-of-moments estimators can concentrate on the moments of interest to the exclusion of others (see Watson 1993, p. 1037).

As noted earlier, Altug (1989) estimates and rejects a close relative of Kydland and Prescott's model using maximum likelihood. The central problem of this paper can be restated: is there a case for ignoring Altug's rejection of the Kydland and Prescott model? The case must be something other than the standard one of too many parameters found in the literature on computable general-equilibrium models.

2.2. *Calibration as estimation*

Various authors have attempted to tame calibration and return it to the traditional econometric fold. Manuelli and Sargent (1988), Gregory and Smith (1990a), Canova (1991a), and Bansal *et al.* (1991) interpret calibration as a form of 'estimation by simulation'. In such a procedure, parameters are chosen, and the relevant features of the simulated output of the calibrated model are compared to the analogous features of the actual data. Such a procedure differs from standard estimation methods principally in that it allows the investigator to expand or restrict the range of features considered to be relevant.

Lucas's argument, however, is that any form of estimation is irrelevant. In their early writings, Kydland and Prescott were not in fact as explicit as Lucas about the irrelevance of estimation. They merely argued that it would be premature to apply techniques to their model, such as those developed by Lars Peter Hansen and Sargent (1980), to account for the systemic effects of rational expectations (Kydland and Prescott, 1982, p. 1369). Prescott (1983, pp. 8–11)

[6] Canova (1991a) suggests a formal methodology and provides an example in which sensitivity analysis is conducted with respect to distributions for parameter values constructed from the different values reported in unrelated studies or from *a priori* information on the practically or theoretically admissible range of parameter values.

[7] Lawrence J. Lau (1984) notices that any model that can be calibrated can also be estimated. He uses 'calibration', however, in a narrow sense. A model is calibrated when its parameters are chosen to reproduce the data of a benchmark period. Thus, parameterization on the basis of unrelated econometric studies does not count as calibration for him. Lau's usage is diametrically opposed to that of Gregory and Smith (1991) for whom calibration is only the assignment of parameter values from unrelated sources. We use 'calibration' in both Lau's and Gregory and Smith's senses. Lau and, similarly, James MacKinnon (1984) make strong pleas for estimation instead of, or in addition to, calibration, and for subjecting computable general-equilibrium models to statistical specification tests.

was more pointed: real-business-cycle models are tightly parameterized. They will almost inevitably be rejected against a weakly restricted alternative hypothesis, but such alternative hypotheses arise from the introduction of arbitrary stochastic processes and, so, are not suitable benchmarks for economic inference. 'A model may mimic the cycle well but not perfectly' (Prescott 1983, p. 10). Similarly, Kydland and Prescott (1991, p. 174) write:

> Unlike the system-of-equations approach, the model economy which better fits the data is not [necessarily?] the one used. Rather, currently established theory dictates which one is used.

The dominance of theory in the choice of models lies at the heart of the difference between estimators and calibrators. To throw the difference into high relief, one can think of estimators pursuing a competitive strategy and calibrators pursuing an adaptive strategy. Under the competitive strategy, theory proposes, estimation and testing disposes. In fine, alternative theories compete with one another for the support of the data. The adaptive strategy begins with an unrealistic model, in the sense of one that is an idealized and simplified product of the core theory. It sees how much mileage it can get out of that model. Only then does it add any complicating and more realistic feature. Unlike the competitive strategy, the aim is never to test and possibly reject the core theory, but to construct models that reproduce the economy more and more closely within the strict limits of the basic theory.

The distinction between the competitive and adaptive strategies is sharply drawn and somewhat stylized, but focuses nonetheless on a genuine difference. On the one hand, the competitive strategy is the received view of econometricians, taught in an idealized form in most econometric textbooks, even if more honored in the breach than the observance by applied economists. The competitive strategy is explicit in Gregory and Smith's (1990b) 'Calibration as Testing'. Even if in practice no theory is ever decisively rejected through a test based on an econometic estimation, the theory is nonetheless regarded as at risk and contingent—even at its core. On the other hand, the real-business-cycle modeller typically does not regard the core theory at risk in principle. Like the estimators, the calibrators wish to have a close fit between their quantified models and the actual data—at least in selected dimensions. But the failure to obtain a close fit (statistical rejection) does not provide grounds for rejecting the fundamental underlying theory. Adaptation in the face of recalcitrant data is adaptation of peripheral assumptions, not of the core. Thus, the inability of Kydland and Prescott's (1982) original real-business-cycle model to match the data prompted more complicated versions of essentially the same model that included, for example, heterogeneous labor (Kydland 1984), a banking sector (Robert G. King and Charles I. Plosser 1984), indivisible labor (Gary Hansen 1985), separate scales for straight-time and overtime work (Gary Hansen and Sargent 1988), and variable capital intensity (Kydland and Prescott 1988).

One consequence of these strategies is that esimators possess a common ground, the performance of each theoretically-based specification against actual

data, on which to judge the performance of competing models. For the calibrators, however, data help discriminate only between different adaptations of the common core. The core theory itself is not questioned, so that, unintentially perhaps, the core theory becomes effectively a Lakatosian hard-core (Lakatos 1970, 1978; Blaug 1992, ch. 2). Calibration does not provide a method that could in principle decide between fundamentally different business-cycle models (e.g. real-business-cycle models or Keynesian business-cycle models) on the basis of empirical evidence derived from the calibration exercise itself.[8] Critics of real-business-cycle models who attempt such comparisons fall back either on attacking the discriminating power of calibration methods (e.g. Hartley *et al.* 1993) or on adaptations of standard econometric techniques (e.g. Canova *et al.* 1993). Kydland and Prescott are explicit in rejecting these applications of estimation techniques as missing the point of the calibration method. The aim of this paper is partly to explicate and appraise their view.[9]

2.3. *The mantel of Frisch*

Calibrators radically reject the system-of-equations approach. But Kydland and Prescott, at least, do not reject econometrics. Rather, they argue that econometrics is not coextensive with estimation; calibration is econometrics. Kydland and Prescott (1991, pp. 161, 162) point out that for Ragnar Frisch, Irving Fisher, and Joseph Shumpeter, the founders of the Econometric Society, 'econometrics' was the unification of statistics, economic theory, and mathematics. Unacknowledged by Kydland and Prescott, Mary Morgan's (1990) brilliant history of econometrics supports and elaborates their point. According to Morgan, even before the term 'econometrics' had wide currency, the econometric ideal had been to weld mathematical, deductive economics to statistical, empirical economics to provide a substitute for the experimental methods of the natural sciences appropriate to the study of society. This ideal collapsed with the rise of the system-of-equations approach in the wake of the Cowles Commission.

Kydland and Prescott point to Frisch's (1933) article, 'Propagation Problems and Impulse Response Problems in Dynamic Economics' as a precursor to both their own real-business-cycle model and to calibration methods. Frisch argues that quantitative analysis requires complete models: i.e. general-equilibrium models in a broad sense. He considers a sequence of models, starting with a very simple one, and then adding complications. He models the time-to-build feature of capital formation. He distinguishes between the impulses that start business cycles and the dynamic mechansisms that amplify and propagate them. He quantifies his models using calibration techniques. And,

[8] This is not to say that there could not be some other basis for some decision.

[9] Hoover 1994a (as well as work in progress) outlines a possible method of using econometric techniques in a way that respects the idealized nature of the core models without giving up the possibility of empirical discrimination.

precisely like Kydland and Prescott (1982), Frisch marvels at how well a very simple model can capture the features of actual data.

Although Kydland and Prescott are correct to see the affinities between Frisch's work and their own, they ignore the very real differences between Frisch and themselves. Frisch's approach is wholly macroeconomic. Frisch writes:

> In order to attack these problems on a macro-dynamic basis so as to explain the movement of the system taken in its entirety, we must deliberately disregard a considerable amount of the details of the picture. We may perhaps start by throwing all kinds of production into one variable, all consumption into another, and so on, imagining that the notions 'production', 'consumption', and on, can be measured by some sort of total indices. (1933, p. 173)

While his flight to aggregates parallels the practice of the new classical real-business-cycle model, Frisch does not suggest that this is a way station on the road to microfoundations. His article does not hint at the desirability of microfoundations, even of the pseudo-microfoundations of the representative-agent model: there is not an optimization problem to be found. Frisch appears to use calibration mainly for purposes of illustration, and not to advocate it as a preferred technique. He writes:

> At present I am guessing very roughly at these parameters, but I believe that it will be possible by appropriate statistical methods to obtain more exact information about them. I think, indeed, that the statistical determination of such structural parameters will be one of the main objectives of the economic cycle analysis of the future. (1933, p. 185)

Precisely which statistical methods are appropriate appears to be an open question.[10]

More generally, although Frisch stresses the importance of theory, there is no hint that his interpretation is limited to 'maximizing behavior subject to constraints' (Kydland and Prescott 1991, p. 164). Frisch does not define 'theory' in 'Propagation Problems . . .', but the examples he produces of theories are not of an obviously different character from the structures employed by Jan Tinbergen, Lawrence Klein, James Dusenberry, and the other 'Keynesian' macromodelers who are the special bugbears of the advocates of new-classical, real-business-cycle models.

Schumpeter (co-founder with Frisch of the Econometric Society) provides typically prolix discussions of the meaning of 'economic theory' in his magisteral *History of Economic Analysis* (1954). For Schumpeter (1954, pp. 14,

[10] Frisch's own shifting views illustrate how open a question this was for him. By 1936, he had backtracked on the desirability of estimating calibrated models. In 1938, he argued that structural estimation was impossible because of pervasive multicollinearity. In its place he proposed estimating unrestricted reduced forms (see Morgan, 1990, p. 97; also see Aldrich 1989, section 2). In this he comes much closer to Christopher Sims (1980) program of vector autoregressions without 'incredible' identifying restrictions. I am grateful to an anonymous referee for reminding me of this point. Hoover (1992) identifies Sims's program as one of three responses to the Lucas (1976) critique of policy invariance. Calibration and the systems-of-equations approach each possess a analogous response.

15), theories are, on the one hand, 'synonomous with Explanatory Hypotheses', and on the other hand, 'the sum total of the gadgets', such as '"marginal rate of substitution", "marginal productivity", "multiplier", "accelerator"', including 'stragegically useful assumptions', 'by which results may be extracted from the hypothesis'. Schumpeter concludes: 'In Mrs. Robinson's unsurpassingly felicitous phrase, economic theory is a box of tools'. Later Schumpeter defends the theoretical credentials of Wesley C. Mitchell, the subject of Tjalling Koopman's (1947) famous attack on 'Measurement without Theory':

> ... in intention as well as in fact, he was laying the foundations for a 'theory', a business cycle theory as well as a general theory of the economic process, but for a different one. (1954, p. 1166)

Kydland and Prescott (1991, p. 164) argue that the system-of-equations approach flourished in the 1950s only because economists lacked the tools to construct stochastic computable general-equilibrium models. They proclaim the death of the system-of-equations approach:

> The principal reason for the abandonment of the system-of-equations approach, however, was the advances in neoclassical theory that permitted the application of the paradigm in dynamic stochastic settings. Once the neoclassical tools needed for modeling business cycle fluctuations existed, their application to this problem and their ultimate domination over any other method was inevitable. (1991, p. 167)

This is an excessively triumphalist and whiggish history of the development of econometric thought. First, the work of Frisch and others in the 1930s provides no support for Kydland and Prescott's narrowing of the meaning of 'theory' to support such tendentious statements as: 'To summarize the Frisch view, then, econometrics is quantitative *neoclassical* theory with a basis in facts' (Kydland and Prescott 1991, p. 162; emphasis added). (A model is 'neoclassical' for Kydland and Prescott (1991, p. 164) when it is constructed from '... agents maximizing subject to constraints and market clearing'.) Second, the declaration of the death of the system-of-equations approach is premature and greatly exaggerated. Allegiance to the system-of-equations approach motivates the many efforts to interpret calibration as a form of estimation. Third, the calibration methodology is not logically connected to Kydland and Prescott's preferred theoretical framework. The example of Frisch shows that calibration can be applied to models that are not stochastic dynamic optimal-growth models. The example of Lars Peter Hansen and Sargent (1980) shows that, even those who prefer such models, can use them as the source of identification for systems of equations—refining rather than supplanting the traditional econometrics of estimation.

3. The quantification of theory

Although Kydland and Prescott overstate the degree to which Frisch and the econometrics of the 1930s foreshadowed their work, they are nonetheless correct to note many affinities. But such affinities, even if they were more complete than they turn out to be, do not amount to an argument favoring calibration

over estimation. At most, they are an illicit appeal to authority. To date, no compelling defence of the calibration methodology has been offered. An interpretation of the point of calibration and an assessment of its merits can be constructed, however, from hints provided in Lucas's methodological writings of the 1970s and early 1980s.

3.1. *Models*

'Model' is a ubiquitous term in economics, and a term with a variety of meanings. One commonly speaks of an econometric model. Here one means the concrete specification of functional forms for estimation. I call these observational models. The second main class of models are evaluative or interpretive models. An obvious subclass of interpretive/evaluative models are toy models.

A toy model exists merely to illustrate or to check the coherence of principles or their interaction. An example of a toy model is the overlapping-generations model with money in its simplest incarnations. No one would think of drawing quantitative conclusions about the working of the economy from it. Instead one wants to show that models constructed on its principles reproduce certain known qualitative features of the economy and suggest other qualitative features that may not have been known or sufficiently appreciated (cf. Diamond 1984, p. 47), Were one so rash as to estimate such a model, it would surely be rejected, but that would be no reason to abandon it as a testbed for general principles.

Is there another subclass of interpretive/evaluative models, one that involves quantification? Lucas seems to think so:

> One of the functions of theoretical economics is to provide fully articulated, artificial economic systems that can serve as laboratories in which policies that would be prohibitively expensive to experiment within actual economies can be tested out at much lower cost. (Lucas 1980, p. 271)

Let us call such models benchmark models. Benchmark models must be abstract enough and precise enough to permit incontrovertible answers to the questions put to them. Therefore,

> . . . insistence on the 'realism' of an economic model subverts its potential usefulness in thinking about reality. Any model that is well enough articulated to give clear answers to the questions we put to it will necessarily be artificial, abstract, patently unreal. (Lucas 1980, p. 271)

On the other hand, only models that mimic reality in important respects will be useful in analyzing actual economies.

> The more dimensions in which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions. This is the sense in which more 'realism' in a model is clearly preferred to less. (Lucas 1980, p. 272)

Later in the same essay, Lucas emphasizes the quantitative nature of such model building:

> Our task ... is to write a FORTRAN program that will accept specific economic policy rules as 'input' and will generate as 'output' statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies. (p. 288)

For Lucas, Kydland and Prescott's model is precisely such a program.[11]

One might interpret Lucas's remarks as making a superficial contribution to the debate over Milton Friedman's 'Methodology of Positive Economics' (1953): must the assumptions on which a theory is constructed be true or realistic or is it enough that the theory predicts 'as if' they were true? But this would be a mistake. Lucas is making a point about the architecture of models and not about the foundations of secure prediction. Lucas refers to a model as fully 'realistic' when it fully accounts for all the factors that determine the variables of interest. Lucas makes two points. Uncontroversially, he argues that toy models convey deep understanding of economic principles. More interestingly, he argues that benchmark models have an advantage over estimation. This is controversial because estimators believe that fully articulated specifications are required for accurate quantification. This is expressed in their concern for specification error, omitted variable bias, and so forth. Their view is widely shared. The point is not that estimated models are necessarily more realistic in Lucas's sense than calibrated models, nor that estimation is the only or even the most reliable way to quantify a model or its components.[12] Rather it is that any method of quantification that does not aspire to full articulation is likely to mislead. Lucas denies this, and the interesting issues are how to appraise his position, and, if his position is sustainable, how to appraise quantified benchmark models themselves.

To make this clear, consider Lucas's (1987, pp. 20–31) cost-benefit analysis of the policies to raise GNP growth and to damp the business cycle. Lucas's model considers a single representative consumer with a constant-relative-risk-aversion utility function facing an exogenous consumption stream. The model is calibrated by picking reasonable values for the mean and variance of consumption, the subjective rate of discount, and the constant coefficient of relative risk aversion. Lucas then calculates how much extra consumption consumers would require to compensate them in terms of utility for a cut in the growth of consumption and how much consumption they would be willing to give up to secure smoother consumption streams. Although the answers that Lucas seeks are quantitative, the model is not used to make predictions that might be subjected to statistical tests. Indeed, it is a striking illustration of why calibration should not be interpreted as estimation by simulation. Lucas's model is used to set upper bounds to the benefits that might conceivably be gained in the real world. Its parameters must reflect some truth about the world

[11] Kydland and Prescott do not say, however, whether it is actually written in FORTRAN.

[12] For example, to clarify a point raised by an anonymous referee, if the central bank had direct knowledge of it money supply function, that would be better than estimating it.

if it is to be useful, but they could not be easily directly estimated. In that sense, the model is unrealistic.[13]

3.2. *Artifacts*

In a footnote, Lucas (1980, p. 272, fn. 1) cites Simon's *Sciences of the Artificial* (1969) as an 'immediate ancestor' of his 'condensed' account. To uncover a more fully articulated argument for Lucas's approach to modelling, it is worth following up the reference.

For Simon, human artifacts, among which he must count economic models,

> can be thought of as a meeting point—an 'interface'...—between an 'inner' environment, the substance and organization of the artifact itself, and an 'outer' environment, the surroundings in which it operates. (Simon 1969, pp. 6, 7)

An artifact is useful, it achieves its goals, if its inner environment is appropriate to its outer environment.

The distinction between the outer and inner environments is important because there is some degree of independence between them. Clocks tell time for the outer environment. Although they may indicate the time in precisely the same way, say with identical hands on identical faces, the mechanisms of different clocks, their inner environments, may be constructed very differently. For determining when to leave to catch a plane, such differences are irrelevant. Equally, the inner environments may be isolated from all but a few key features of the outer environment. Only light entering through the lens for the short time that its shutter is open impinges on the inner environment of the camera. The remaining light is screened out by the opaque body of the camera, which is an essential part of its design.

Simon factors adaptive systems into goals, outer environments and inner environments. The relative independence of the outer and inner environments means that

> [w]e might hope to characterize the main properties of the system and its behavior without elaborating the detail of *either* the outer or the inner environments. We might look toward a science of the artificial that would depend on the relative simplicity of the interface as its primary source of abstraction and generality. (Simon 1969, p. 9)

Simon's views reinforce Lucas's discussion of models. A model is useful only if it foregoes descriptive realism and selects limited features of reality to reproduce. The assumption upon which the model is based do not matter, so long as the model succeeds in reproducing the selected features. Friedman's 'as if' methodology appears vindicated.

[13] Of course, Lucas's approach might be accepted in principle and still rejected in detail. For example, McCallum (1986, pp. 411, 412) objects to the characterization of consumption as fluctuating symmetrically about trend that is implicit in Lucas's use of a mean/variance model. If the fluctuations of consumption are better described as varying degrees of shortfall relative to the trend of potential maximum consumption, then the benefits of consumption smoothing will be considerably higher than Lucas's finds.

But this is to move too fast. The inner environment is only relatively independent of the outer environment. Adaptation has its limits.

> In a benign environment we would learn from the motor only what it had been called upon to do; in a taxing environment we would learn something about its internal structure—specifically, about those aspects of the internal structure that were chiefly instrumental in limiting performance. (Simon 1969, p. 13)[14]

This is a more general statement of principles underlying Lucas's (1976) critique of macroeconometric models. A benign outer environment for econometric models is one in which policy does not change. Changes of policy produce structural breaks in estimated equations: disintegration of the inner environment of the models. Economic models must be constructed like a ship's chronometer, insulated from the outer environment so that '... it reacts to the pitching of the ship only in the negative sense of maintaining an invariant relation of the hands on its dial to real time, independently of the ship's motions' (Simon 1969, p. 9). Insulation in economic models is achieved by specifying functions whose parameters are invariant to policy. The independence of the inner and outer environments is not something which is true of arbitrary models; rather it must be built into models. While it may be enough in hostile environments for models to reproduce key features of the outer environment 'as if' reality was described by their inner environments, it is not enough if they can do this only in benign environments. Thus, for Lucas, the 'as if' methodology interpreted as an excuse for complacency with respect to modeling assumptions must be rejected. Simon's notion of the artifact helps justify Lucas's both rejecting realism in the sense of full articulation and at the same time, insisting that only through carefully constructing the model from invariants—tastes and technology, in Lucas's usual phrase—can the model secure the benefits of a useful abstraction and generality.

Recognizing that a model must be constructed from invariants does not itself tell us how to quantify it. The emphasis on a maintained theory or inner environment presents a generic risk for quantified idealized models (see Section 2.2 above). The risk is particularly severe for the calibration methodology with its adaptive strategy. Gregory and Smith (1991, p. 30) observe that '[s]etting parameter values (i.e. calibrating), simulating a model and comparing properties of simulations to those of data often suggests fruitful modifications of the model'. Generally, such modifications leave the essential core theory intact and attempt to better account for the divergences from the ideal, to better account for the fudge factors need to link the output of the model to the phenomenal laws. The risk, then, is that the core of the model becomes completely insulated from empirical confirmation or disconfirmation—even in the weakest senses of

[14] Haavelmo (1944, p. 28) makes a similar point in his well-known example of the failure of autonomy: the relationship between the speed of a car and the amount of throttle may be well-defined under uniform conditions, but would break down immediately the car was placed in a different setting. To understand how the car will perform on the track as well as on the road requires us to repair to the deeper principles of its operation.

those terms. Kydland and Prescott (1991, p. 171) explicitly deny that the confidence in the answers a model gives to policy questions can '. . . be resolved by computing some measure of how well the model economy mimics historical data'. Rather, '[t]he degree of confidence in the answer depends on the confidence that is placed in the economic theory being used'. Kydland and Prescott do not explain what alternative sources there might be to justify our confidence in theory; the adaptive strategy of the calibration approach almost guarantees that empirical evidence will not be among those sources.

3.3. *Quantification without history*

Calibrators of real-business-cycle models typically concentrate on matching selected second moments of variables rather than, say, matching the actual historical evolution of the modeled variables. Why? Lucas (1977, p. 218) observes that 'business cycles are all alike', not in exact detail but qualitatively. An informative test of a model's ability to capture business-cycle behavior is not, therefore, its fit to some historical time series, which is but one of many possible realizations, but rather its ability to characterize the distribution from which that realization was drawn. Lucas (1977, pp. 219, 234) advocates the test of Irma Adelman and Frank L. Adelman (1959). The Adelmans asked the question, could one distinguish data generated by simulating a model (in their case, the Klein-Goldberger macroeconometric model) from actual data describing the economy, in the absence of knowledge of which was which? The Adelmans' test compares the distribution of outcomes of the model to the actual economy. Once a close relation is established, to experiment with alternative policy rules is an easy next step. Even though government is not modelled in Kydland and Prescott's initial models, policy analysis is their ultimate goal (Kydland and Prescott, 1982, p. 1369). Concentration on the second moments of variables can be seen as the practical implementation of the Adelmans' standard: one eschews the particular realization in favor of a more general characterization of the distribution of possible outcomes.[15]

One reason, therefore, not to apply a neutral statistical test for the match between model and reality is that it is along only selected dimensions that one cares about the model's performance at all. This is completely consistent with Simon's account of artifacts. New classical economics has traditionally been skeptical about discretionary economic policies. New classical economists are, therefore, more concerned to evaluate the operating characteristics of policy rules. For this, the fit of the model to a particular historical realization is largely irrelevant, unless it assures it will also characterize the future distribution of

[15] The Adelmans themselves examine the time-series properties of a single draw, rather than the characteristics of repeated draws. This probably reflects, in part, the computational expense of simulating a large macroeconometric model with the technology of 1959. It also reflects the application of Burns and Mitchell's techniques for characterizing the repetitive features of business cycles through averaging over historical cycles all normalized to a notional cycle length. King and Plosser (1989) attempt to apply precisely Burns and Michell's techniques to outcomes generated by a real-business-cycle model.

outcomes. The implicit claim of most econometrics is that it does assure a good characterization. Probably most econometricians would reject calibration methods as coming nowhere close to providing such assurance. Substantial work remains to be done in establishing objective, comparative standards for judging competing models.

4. Aggregation and general equilibrium[16]

Whether calibrated or estimated, real-business-cycle models are idealizations along many dimensions. The most important dimension of idealization is the the models deal in aggregates while the economy is composed of individuals. After all, the distinction between microeconomics and macroeconomics is the distinction between the individual actor and the economy as a whole. All new classical economists believe that one understands macroeconomic behavior only as an outcome of individual rationality. Lucas (1987, p. 57) comes close to adopting the *Verstehen* approach of the Austrians.[17] The difficulty with this approach is that there are millions of people in the economy and it is not practical—nor is it ever likely to become practical—to model the behavior of each of them.[18] Universally, new classical economists adopt representative-agent models, in which one agent or a few types of agents, stand in for the behavior of all agents.[19] The conditions under which a single agent's behavior can accurately represent the behavior of an entire class are onerous and almost certainly never fulfilled in an actual economy.

One interpretation of the use of calibration methods in macroeconomics is that the practitioners recognize that highly aggregated theoretical models must be descriptively false, so that estimates of them are bound to fit badly in comparison to atheoretical econometric models, which are able to exploit large numbers of free parameters. The theoretical models are nonetheless to be preferred because policy evaluation is possible only within their structure. In this, they are exactly like Lucas's benchmark consumption model (see Section 3.1, above).

Calibractors appeal to microeconomic estimates of key parameters because information about individual agents is lost in the aggregatation process. Estimators, in contrast, could argue that the idealized representative-agent

[16] Aggregation and the problems it poses for macroeconomics are the subject of a voluminous literature. The present discussion is limited to a narrow set of issues most relevant to the question of idealization.

[17] For a full discussion of the relationship between new classical and Austrian economics see Hoover (1988, ch. 10).

[18] In Hoover (1984, pp. 64–6; and 1988, pp. 218–20), I refer to this as the 'Cournot problem' since it was first articulated by Augustin Cournot ([1838] 1927, p. 127).

[19] Some economists reserve the term 'representative-agent models' for models with a single, infinitely-lived agent. In a typical overlapping-generations model the new young are born at the start of every period, and the old die at the end of every period, and the model has infinitely many periods; so there are infinitely many agents. On this view, the overlapping-generations model is not a representative-agent model. I, however, regard it as one, because within any period one type of young agent and one type of old agent stand in for the enormous variety of people, and the same types are repeated period after period.

model permits better use of other information. Lars Peter Hansen and Sargent (1980, pp. 91, 92), for example, argue that the strength of their estimation method is that it accounts consistently for the interrelationships between constituent parts of the model—i.e. that is a general-equilibrium method. Calibrators respond, however, that it is precisely the importance of general equilibrium that supports their approach. Kydland and Prescott write:

> . . . it is in the stage of calibration where the power of the general equilibrium approach shows up most forcefully. The insistence on internal consistency implies that parsimoniously parameterized models of the household and business sector display rich dynamic behavior through the intertemporal substitution arising from capital accumulation and from other sources. (1991, p. 170)

The trade-off between the gains and losses of the two methods is not clear cut. Lucas (1987, pp. 46, 47) and Prescott (1986, p. 15) argue that the strength of calibration is that it uses multiple sources of information, supporting the belief that it is structured around true invariants. This argument would appear to appeal to the respectable, albeit philosophically controversial view, that a theory is better supported when tested on information not used in its formulation (see Lipton 1991, ch. 8; Hoover 1994b). Unfortunately, it is not clear that calibration relies on independent information nor that it avoids estimation altogether. Parameters are sometimes chosen for calibrated business-cycle models because they mimic so-called 'stylized facts'. That the models then faithfully reproduce such facts is not independent information. Other parameters are chosen from microeconomic studies. This introduces estimation through the back door, but without any but a subjective, aesthetic metric to judge model performance.

Furthermore, since all new classical, equilibrium business-cycle models rely on the idealization of the representative agent, both calibrated and estimated versions share a common disability: using the representative-agent model in any form begs the question by assuming that aggregation does not fundamentally alter the structure of the aggregate model. Physics may provide a useful analogy. The laws that relate pressure, temperature, and volumes of gases are macrophysics. The 'ideal-gas laws' can be derived from a micromodel: gas molecules are assumed to be point masses, subject to conservation of momentum, with a distribution of velocities. An aggregate assumption is also needed: the probability of the gas molecules moving in any direction is taken to be equal.

Direct estimation of the ideal gas laws shows that they tend to break down—and must be corrected with fudge factors—when pushed to extremes. For example, under high pressures or low temperatures the ideal laws must be corrected according to van der Waals' equation. This phenomenal law, a law in macrophysics, is used to justify alterations of the micromodel: when pressures are high one must recognize that forces operate between individual molecules.

The inference of the van der Waals' force from the macrophysical behavior of gases has an analogue in the development of real-business-cycle models. Gary Hansen (1985), for example, introduces the microeconomic device of indivisible labor into a real-business cycle model, not from any direct reflection on the

nature of labor markets at the level of the individual firm or worker, but as an attempt to account for the macroeconomic failure of Kydland and Prescott's (1982) model to satisfactorily reproduce the relative variabilities of hours and productivity in the aggregate data.[20] Of course, direct estimation of Kydland and Prescott's model rather than calibration may have pointed in the same direction.[21]

Despite examples of macro to micro inference analogous to the gas laws, Lucas's (1980, p. 291) more typical view is that we must build our models up from the microeconomic to the macroeconomic. Unlike gases, human society does not comprise homogeneous molecules, but rational people, each choosing constantly. To understand (*verstehen*) their behavior, one must model the individual and his situation. This insight is clearly correct, it is not clear in the least that it is adequately captured in the heroic aggregation assumptions of the representative-agent model. The analogue for physics would be to model the behavior of gases at the macrophysical level, not as derived from the aggregation of molecules of randomly distributed momenta, but as a single molecule scaled up to observable volume—a thing corresponding to nothing ever known to nature.[22]

5. Calibration and macroeconomic practice

The calibration methodology has both a wide following and a substantial opposition within the new classical school. I have attempted to give it a sympathetic reading—both in general and in its specific application to real-business-cycle models. I have concentrated on Kydland and Prescott, as its most prolific practitioners, and on Lucas, an articulate advocate. Although calibration is consistent with appealing accounts of the nature and role of models in science and economics, of their quanfication and idealization, its practical implementation in the service of real-business-cycle models with representative agents is less than compelling.

Does the calibration methodology amount to a repudiation of econometric estimation altogether? Clearly not. At some level, econometrics still helps to supply the values of the parameters of the models. Beyond that, whatever has been said in favor of calibration methods to the contrary notwithstanding, the

[20] Canova (1991b, p. 33) suggests that the particular covariance that Hansen's modification of Kydland and Prescott's model was meant to capture is an artifact of the Hodrick-Prescott filter, so that Hansen's model may be a product of misdirected effort rather than a progressive adaptation.

[21] This, rather than collegiality, may account for Kydland and Prescott's (1982, p. 1369) tolerant remark about the future utility of Lars Peter Hansen and Sargent's (1980) econometric techniques as well as for Lucas's (1987, p. 45) view that there is something to be learned from Altug's estimations of the Kydland and Prescott model—a view expressed in the midst of arguing in favour of calibration.

[22] A notable, non-new-classical attempt to derive macroeconomic behavior from microeconomic behavior with appropriate aggregation assumptions is Durlauf (1989). In a different, but related context, Stoker (1986) shows that demand systems fit the data only if distributional variables are included in the estimating equations. He takes this macroeconomic evidence as evidence for the failure of the microeconomic conditions of exact aggregation.

misgivings of econometricians such as Sargent are genuine. The calibration methodology, to date, lacks any discipline as stern as that imposed by econometric methods. For Lucas (1980, p. 288) and Prescott (1983, p. 11), the discipline of the calibration method comes from the paucity of free parameters. But one should note that theory places only loose restrictions on the values of key parameters. In practice, they are actually pinned down from econometric estimation at the microeconomic level or accounting considerations. Thus, in some sense, the calibration method would appear to be a kind of indirect estimation. Thus, although as was pointed out earlier, it would be a mistake to treat calibration as simply an alternative form of estimation, it is easy to understand why some critics interpret it that way. Even were there less flexibility in the parameterizations, the properties ascribed to the underlying components of the idealized real-business-cycle models (the agents, their utility functions, production functions, and constraints) are not subject to as convincing cross checking as the analogous components in physical models usually are. The fudge factors that account for the discrepancies between the ideal model and the data look less like van der Waals' equation than like special pleading. Above all, it is not clear on what standards competing, but contradictory, models are to be compared and adjudicated.[23] Some such standards are essential if any objective progress is to be made in economics.[24]

ACKNOWLEDGEMENTS

I thank Thomas Mayer, Kevin Salyer, Steven Sheffrin, Roy Epstein, Nancy Cartwright, Gregor Smith, Edward Prescrott, Adrian Pagan, and two anonymous referees for helpful comments on an earlier draft. The earliest version of this paper, entitled 'Calibration versus Estimation: Standards of Empirical Assessment in the New Classical Macroeconomics', was presented at the American Economic Association meetings in Washington, DC, December 1990.

[23] Prescott (1983, p. 12) seems oddly, to claim that inability of a model to account for some real events is a positive virtue—in particular, that the inability of real-business-cycle models to account for the Great Depression is a point in their favour. He writes: 'If any observation can be rationalized with some approach, then that approach is not scientific'. This seems to be a confused rendition of the respectable Popperian notion that a theory is more powerful the more things it rules out. But one must not mistake the power of a theory with its truth. Aside from issues of tractability, a theory that rationalizes only and exactly those events that actually occur, while ruling out exactly those events that do not occur is the perfect theory. In contrast, Prescott seems inadvertently to support the view that the more exceptions the better rule.

[24] Watson (1993) develops a goodness-of-fit measure for calibrated models. It takes into account that, since idealization implies differences between model and reality that may be systematic, the errors-in-variables and errors-in-equations statistical models are probably not appropriate. Also see Gregory and Smith (1991, pp. 27–8), Canova (1991a), and Hoover (1994a).

REFERENCES

ADELMAN, I. and ADELMAN, F. L. (1959). 'The Dynamic Properties of the Klein-Goldberger Model', *Econometrica*, **27**, 596–625.

ALDRICH, J. (1989). 'Autonomy', *Oxford Economic Papers*, **41**, 15–34.

ALTUG, S. (1989). 'Time-to-Build and Aggregate Fluctuations: Some New Evidence', *International Economic Review*, **30**, 889–920.

BANSAL, R. and GALLANT, R. A., HUSSEY, R., and TAUCHEN, G. (1991). 'Nonparametric Structural Estimation of Models for High-Frequency Currency Market Data', unpublished typescript.

BLAUG, M. (1992). *The Methodology of Economics: Or How Economists Explain*, Cambridge University Press, Cambridge, 2nd ed.

CANOVA, F. (1991a). 'Sensitivity Analysis and Model Evalution in Simulated Dynamic General Equilibrium Economies', unpublished typescript, Department of Economics, European University Institute, Florence, Italy.

CANOVA, F. (1991b). 'Detrending and Business Cycle Facts', unpublished typescript, Department of Economics, European University Institute, Florence, Italy.

CANOVA, F., FINN, M., and PAGAN, A. R. (1992). 'Evaluating a Real Business Cycle Model', unpublished typescript.

COGLEY, T. and NASON, J. M. (1993). 'Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research', unpublished typescript.

COOLEY, T. F. and HANSEN, G. D. (1989). 'Welfare Consequences of Monetary and Fiscal Policy', paper presented to the Federal Reserve Bank of San Francisco's Fall Academic Conference, 1 December.

COURNOT, A. [1838] (1927). *Researches into the Mathematical Principles of the Theory of Wealth*, Nathaniel T. Bacon (trans,), Macmillan, New York.

DIAMOND, P. A. (1984). *A Search-Equilibrium Approach to the Micro Foundations of Macroeconomics: The Wicksell Lectures, 1982*, MIT Press, Cambridge, MA.

DURLAUF, S. N. (1989). 'Locally Interacting Systems, Coordination Failure, and the Behavior of Aggregate Activity', unpublished typescript.

FRIEDMAN, M. (1953). 'The Methodology of Positive Economics', in *Essays in Positive Economics*, Chicago University Press, Chicago, IL.

FRISCH, R. (1933). 'Propagation Problems and Impulse Response Problems in Dynamic Economics', in *Economic Essays in Honour of Gustav Cassel: October 20th 1933*, George Allen and Unwin, London.

GREGORY, A. W. and SMITH, G. W. (1990a). 'Calibration as Estimation', *Econometric Reviews*, **9**, 57–89.

GREGORY, A. W. and SMITH, G. W. (1990b). 'Calibration as Testing: Inference in Simulated Macroeconomic Models' in *Journal of Business and Economic Statistics*, forthcoming.

GREGORY, A. W. and SMITH, G. W. (1991). 'Calibration in Macroeconomics', in G. S. Maddala and C. R. Rao (eds), *Handbook of Statistics*, **10**: *Econometics*, forthcoming.

HAAVELMO, T. (1944). 'The Probability Approach in Econometrics', *Econometrica*, **12** (supplement).

HALL, R. E. (1977). 'Investment, Interest Rates, and the Effects of Stabilization Policies' *Brookings Papers on Economic Activity*, **6**, 61–100.

HANSEN, G. D. (1985). 'Indivisible Labor and the Business Cycle', *Journal of Monetary Economics*, **16**, 309–28.

HANSEN, G. D. and PRESCOTT, E. C. (1993). 'Did Technology Shocks Cause the 1990–1991 Recession', *American Economic Review*, **83**, 280–6.

HANSEN, G. D. and SARGENT, T. J. (1988), 'Straight Time and Overtime in Equilibrium', *Journal of Monetary Economics*, **21**, 281–308.

HANSEN, L. P. and SARGENT, T. J. (1980). 'Formulating and Estimating Dynamic Linear Rational Expectations Models', in R. E. Lucas, Jr. and T. J. Sargent (eds), *Rational Expectations and Econometric Practice*, George Allen & Unwin, London, 1981.

HARTLEY, J. , SALYER, K., and SHEFFRIN, S. (1992). 'Calibration and Real Business Cycle Models: Two Unorthodox Tests', unpublished typescript.

HOOVER, K. D. (1984). 'Two Types of Monetarism', *Journal of Economic Literature*, **22**, 58–76.

HOOVER, K. D. (1988). *The New Classical Macroeconomics: A Skeptical Inquiry*, Blackwell, Oxford.

HOOVER, K. D. (1992). 'The Rational Expectations Revolution: An Assessment', *The Cato Journal*, **12**, 81–96.

HOOVER, K. D. (1994a). 'Six Queries About Idealization in an Empirical Context', *Poznan Studies in the Philosphy of Science and the Humanities*, **38**, 43–53.

HOOVER, K. D. (1994b). 'In Defence of Datamining: Some Preliminary Throughts', in K. D. Hoover and S. M. Sheffrin (eds), *Monetarism and the Methodology of Econmics: Essays in Honor of Thomas Mayer*, Elgar, Aldershot, forthcoming.

KING, R. G. and PLOSSER, C. I. (1984). 'Money, Credit and Prices in a Real Business Cycle', *American Economic Review*, **74**, 363–80.

KING, R. G. and PLOSSER, C. I. (1989). 'Real Business Cycle Models and the Test of the Adelmans', paper presented to the Federal Reserve Bank of San Francisco's Fall Academic Conference, 1 December 1989.

KOOPMANS, T. C. (1947). 'Measurement without Theory', *Review of Economics and Statistics*, **29**, 161–72.

KYDLAND, F. E. (1984). 'Labor-force Heterogeneity and the Business Cycle' in K. Brunner and A. H. Meltzer (eds), *Essays on Macroeconomic Implications of Financial and Labor Markets and Political Processes*, Carnegie-Rochester Conference Series on Public Policy, **21**.

KYDLAND, F. E. and PRESCOTT, E. C. (1982). 'Time to Build and Aggregate Fluctuations', *Econometrica*, **50**, 1345–70.

KYDLAND, F. E. and PRESCOTT, E. C. (1988). 'The Workweek of Capital and Its Cyclical Implication' *Journal of Monetary Economics*, **21**, 243–60.

KYDLAND, F. E. and PRESCOTT, E. C. (1990). 'Business Cycles: Real Facts and a Monetary Myth', *Federal Reserve Bank of Minneapolis Quartly Review*, **14**, 3–18.

KYDLAND, F. E. and PRESCOTT, E. C. (1991). 'The Econometrics of the Great Equlibrium Approach to Business Cycles', *Scandinavian Journal of Economics*, **93**, 161–78.

LAKATOS, I. (1970). 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge*, Cambridge University Press, Cambridge.

LAKATOS, I. (1978). 'History of Science and Its Rational Reconstructions' in J. Worrall and G. Currie (eds), *The Methodology of Scientific Research Programmes: Philosphical Papers*, **I**, Cambridge University Press, Cambridge.

LAU, L. J. (1984). 'Comments', in H. E. Scarf and J. B. Shoven (eds), *Applied General Equilibrium Analysis*, Cambridge University Press, Cambridge.

LIPTON, P. (1991). *Inference to the Best Explanation*, Routlege, London.

LUCAS, R. E., JR. (1976). 'Econometric Policy Evaluaton: A Critique', in *Studies in Business-Cycle Theory*, Oxford, Blackwell, 1981.

LUCAS, R. E., JR. (1977). 'Understanding Business Cycles', in *Studies in Business-Cycle Theory*, Blackwell, Oxford, 1981.

LUCAS, R. E., JR. (1980) 'Methods and Problems in Business Cycle Theory', in *Studies in Business-Cycle Theory*, Blackwell, Oxford, 1981.

LUCAS, R. E. JR. (1986). 'Adapttive Behavior and Economic Theory', *Journal of Business*, **59**, S401–26.

LUCAS, R. E., JR. (1987), *Models of Business Cycles*, Blackwell, Oxford.

LUCAS, R. E., JR. and SARGENT, T. J. C. (eds) (1981). *Rational Expectations and Econometric Practice*, George Allen and Unwin, London.

MACKINNON, J. (1984). 'Comments', in H. E. Scarf and J. B. Shoven (eds), *Applied General Equilibrium Analysis*, Cambridge University Press, Cambridge.

MANSUR, A. and WALLEY, J. (1984). 'Numerical Specification of Applied General-equilibrium models: Estimation, Calibration, and Data', in H. E. Scarf and J. B. Shoven (eds.), *Applied General Equilibrium Analysis*, Cambridge University Press, Cambridge.

MANUELLI, R. and SARGENT, T. J. (1988). 'Models of Business Cycles: A Review Essay' *Journal of Monetary Economics*, **22**, 523–42.

MAYER, T. (1960). 'Plant and Equipment Lead Times' *Journal of Business*, **33**, 127–32.

MCCALLUM, B. T. (1986). 'On "Real" and "Sticky-Price" Theories of the Business Cycle', *Journal of Money, Credit and Banking*, **18**, 397–414.

MCCALLUM, B. T. (1989). 'Real Business Cycle Models', in R. J. Barro (ed.), *Modern Business Cycle Theory*, Blackwell, Oxford.

MORGAN, M. (1990). *The History of Econometric Ideas*, Cambridge University Press, Cambridge.
PRESCOTT, E. C. (1983). '"Can the Cycle be Reconciled with a Consistent Theory of Expectations" or A Progress Report on Business Cycle Theory', Working Paper, No. 239, Research Department, Federal Reserve Bank of Minneapolis.
PRESCOTT, E. C. (1986). 'Theory Ahead of Business Cycle Measurement', *Federal Reserve Bank of Minneapolis Quarterly Review*, **10**, 9–22.
SCHUMPETER, J. A. (1954). *History of Economic Analysis*, Oxford University Press, New York.
SHEFFRIN, S. M. (1989). *The Making of Economic Policy*, Blackwell, Oxford.
SIMON, H. A. (1969). *The Sciences of the Artificial*, MIT Press, Cambridge, MA.
SIMS, C. A. (1980). 'Macroeconomics and Reality', *Econometrica*, **48**, 1–48.
STOKER, T. M. (1986). 'Simple Tests of Distributional Effects on Macroeconomic Equations', *Journal of Political Economy*, **94**, 763–95.
WATSON, M. W. (1993). 'Measures of Fit for Calibrated Models', *Journal of Political Economy*, **101**, 1011–41.