Journal of Economic Perspectives—Volume 10, Number 1—Winter 1996—Pages 87–104

The Empirical Foundations of Calibration

Lars Peter Hansen and James J. Heckman

General equilibrium theory provides the intellectual underpinnings for modern macroeconomics, finance, urban economics, public finance and numerous other fields. However, as a paradigm for organizing and synthesizing economic data, it poses some arduous challenges. A widely accepted empirical counterpart to general equilibrium theory remains to be developed. There are few testable implications of general equilibrium theory for a time series of aggregate quantities and prices. There are a variety of ways to formalize this claim. Sonnenschein (1973) and Mantel (1974) show that excess aggregate demand functions can have "arbitrary shapes" as long as Walras' Law is satisfied. Similarly, Harrison and Kreps (1979) show that a competitive equilibria can always be constructed to rationalize any arbitrage-free specification of prices. Observational equivalence results are pervasive in economics. There are two responses to this state of affairs. One can view the flexibility of the general equilibrium paradigm as its virtue. Since it is hard to reject, it provides a rich apparatus for interpreting and processing data.[1] Alternatively, general equilibrium theory can be dismissed as being empirically irrelevant because it imposes no testable restrictions on market data.

Even if we view the "flexibility" of the general equilibrium paradigm as a virtue, identification of preferences and technology is problematic. Concern about the

[1] Lucas and Sargent (1988) make this point in arguing that early Keynesian critiques of classical economics were misguided by their failure to recognize this flexibility.

■ *Lars Peter Hansen is Homer J. Livingston Professor of Economics, and James Heckman is Henry Schultz Distinguished Service Professor of Economics and Director of the Center for Social Program Evaluation at the Irving B. Harris School of Public Policy Studies, all at the University of Chicago, Chicago, Illinois.*

lack of identification of aggregate models has long troubled econometricians (for example, Liu, 1960; Sims, 1980). The tenuousness of identification of many models makes policy analysis and the evaluation of the welfare costs of programs a difficult task and leads to distrust of aggregate models. Different models that "fit the facts" may produce conflicting estimates of welfare costs and dissimilar predictions about the response of the economy to changes in resource constraints.

Numerous attempts have been made to circumvent this lack of identification, either by imposing restrictions directly on aggregate preferences and technologies, or by limiting the assumed degree of heterogeneity in preferences and technologies. For instance, the constant elasticity of substitution specification for preferences over consumption in different time periods is one of workhorses of dynamic stochastic equilibrium theory. When asset markets are sufficiently rich, it is known from Gorman (1953) that these preferences can be aggregated into the preferences of a representative consumer (Rubinstein, 1974). Similarly, Cobb-Douglas aggregate production functions can be obtained from Leontief micro technologies aggregated by a Pareto distribution for micro productivity parameters (Houthakker, 1956). These results give examples of when simple aggregate relations can be deduced from relations underlying the micro behavior of the individual agents, but they do *not* justify using the constructed aggregate relations to evaluate fully the welfare costs and benefits of a policy.[2]

Micro data offer one potential avenue for resolving the identification problem, but there is no clear formal statement that demonstrates how access to such data fully resolves it. At an abstract level, Brown and Matzkin (1995) show how to use information on individual endowments to obtain testable implications in exchange economies. As long as individual income from endowments can be decomposed into its component sources, they show that the testability of general equilibrium theory extends to production economies. Additional restrictions and considerable price variation are needed to identify microeconomic preference relations for data sets that pass the Brown-Matzkin test.

Current econometric practice in microeconomics is still far from the nonparametric ideal envisioned by Brown and Matzkin (1995). As shown by Gorman (1953), Wilson (1968), Aigner and Simon (1970) and Simon and Aigner (1970), it is only under very special circumstances that a micro parameter such as the intertemporal elasticity of substitution or even a marginal propensity to consume out of income can be "plugged into" a representative consumer model to produce an empirically concordant aggregate model. As illustrated by Houthakker's (1956) result, microeconomic technologies can look quite different from their aggregate counterparts. In practice, microeconomic elasticities are often estimated by reverting to a partial

[2] Gorman's (1953) results provide a partial justification for using aggregate preferences to compare alternative aggregate paths of the economy. Even if one aggregate consumption-investment profile is preferred to another via this aggregate preference ordering, to convert this into a Pareto ranking for the original heterogeneous agent economy requires computing individual allocations for the path—a daunting task.

equilibrium econometric model. Cross-market price elasticities are either assumed to be zero or are collapsed into constant terms or time dummies as a matter of convenience. General equilibrium, multimarket price variation is typically ignored in most microeconomic studies.

Battle lines are drawn over the issue of whether the microeconometric simplifications commonly employed are quantitatively important in evaluating social welfare and assessing policy reforms. Shoven and Whalley (1972, 1992) attacked Harberger's use of partial equilibrium analysis in assessing the effects of taxes on outputs and welfare. Armed with Scarf's algorithm (Scarf and Hansen, 1973), they computed fundamentally larger welfare losses from taxation using a general equilibrium framework than Harberger computed using partial equilibrium analysis. However, these and other applications of general equilibrium theory are often greeted with skepticism by applied economists who claim that the computations rest on weak empirical foundations. The results of many simulation experiments are held to be fundamentally implausible because the empirical foundations of the exercises are not secure.

Kydland and Prescott are to be praised for taking the general equilibrium analysis of Shoven and Whalley one step further by using stochastic general equilibrium as a framework for understanding macroeconomics.[3] Their vision is bold and imaginative, and their research program has produced many creative analyses. In implementing the real business cycle program, researchers deliberately choose to use simple stylized models both to minimize the number of parameters to be "calibrated" and to facilitate computations.[4] This decision forces them to embrace a rather different notion of "testability" than used by the other general equilibrium theorists, such as Sonnenschein, Mantel, Brown and Matzkin. At the same time, the real business cycle community dismisses conventional econometric testing of parametric models as being irrelevant for their purposes. While Kydland and Prescott advocate the use of "well-tested theories" in their essay, they never move beyond this slogan, and they do not justify their claim of fulfilling this criterion in their own research. "Well tested" must mean more than "familiar" or "widely accepted" or "agreed on by convention," if it is to mean anything.

Their suggestion that we "calibrate the model" is similarly vague. On one hand, it is hard to fault their advocacy of tightly parameterized models, because such models are convenient to analyze and easy to understand. Aggregate growth coupled with uncertainty makes nonparametric identification of preferences and technology extremely difficult, if not impossible. Separability and homogeneity restrictions on preferences and technologies have considerable appeal as identifying assumptions. On the other hand, Kydland and Prescott never provide a coherent

[3] The earlier work by Lucas and Prescott (1971) took an initial step in this direction by providing a dynamic stochastic equilibrium framework for evaluating empirical models of investment.

[4] The term "real business cycle" originates from an emphasis on technology shocks as a source of business cycle fluctuations. Thus, real, as opposed to nominal, variables drive the process. In some of the recent work, both real and nominal shocks are used in the models.

framework for extracting parameters from microeconomic data. The same charge of having a weak empirical foundation that plagued the application of deterministic general equilibrium theory can be lodged against the real business cycle research program. Such models are often elegant, and the discussions produced from using them are frequently stimulating and provocative, but their empirical foundations are not secure. What credibility should we attach to numbers produced from their "computational experiments," and why should we use their "calibrated models" as a basis for serious quantitative policy evaluation? The essay by Kydland and Prescott begs these fundamental questions.

The remainder of our essay is organized as follows. We begin by discussing simulation as a method for understanding models. This method is old, and the problems in using it recur in current applications. We then argue that model calibration and verification can be fruitfully posed as econometric estimation and testing problems. In particular, we delineate the gains from using an explicit econometric framework. Following this discussion, we investigate current calibration practice with an eye toward suggesting improvements that will make the outputs of computational experiments more credible. The deliberately limited use of available information in such computational experiments runs the danger of making many economic models with very different welfare implications compatible with the evidence. We suggest that Kydland and Prescott's account of the availability and value of micro estimates for macro models is dramatically overstated. There is no filing cabinet full of robust micro estimates ready to use in calibrating dynamic stochastic general equilibrium models. We outline an alternative paradigm that, while continuing to stress the synergy between microeconometrics and macro simulation, will provide more credible inputs into the computational experiments and more accurate assessments of the quality of the outputs.

Simulation as a Method for Understanding Models

In a simple linear regression model, the effect of an independent variable on the dependent variable is measured by its associated regression coefficient. In the dynamic nonlinear models used in the Kydland-Prescott real business cycle research program, this is no longer true. The dynamic nature of such models means that the dependent variable is generated in part from its own past values. Characterizing the dynamic mechanisms by which exogenous impulses are transmitted into endogenous time series is important to understanding how these models induce fluctuations in economic aggregates. Although there is a reliance on linearization techniques in much of the current literature, for large impulses or shocks, the nonlinear nature of such models is potentially important. To capture the richness of a model, the analyst must examine various complicated features of it. One way to do this is to simulate the model at a variety of levels of the forcing processes, impulses and parameters.

The idea of simulating a complex model to understand its properties is not a new principle in macroeconomics. Tinbergen's (1939) simulation of his League of Nations model and the influential simulations of Klein and Goldberger (1955) and Goldberger (1959) are but three of a legion of simulation exercises performed by previous generations of economists.[5] Fair (1994) and Taylor (1993) are recent examples of important studies that rely on simulation to elucidate the properties of estimated models.

However, the quality of any simulation is only as good as the input on which it is based. The controversial part of the real business cycle simulation program is the method by which the input parameters are chosen. Pioneers of simulation and of economic dynamics like Tinbergen (1939) and Frisch (1933) often guessed at the parameters they used in their models, either because the data needed to identify the parameters were not available, or because the econometric methods were not yet developed to fully exploit the available data (Morgan, 1990). At issue is whether the state of the art for picking the parameters to be used for simulations has improved since their time.

Calibration versus Estimation

A novel feature of the real business cycle research program is its endorsement of "calibration" as an alternative to "estimation." However, the distinction drawn between calibrating and estimating the parameters of a model is artificial at best.[6] Moreover, the justification for what is called "calibration" is vague and confusing. In a profession that is already too segmented, the construction of such artificial distinctions is counterproductive. It can only close off a potentially valuable dialogue between real business cycle research and other research in modern econometrics.

Since the Kydland-Prescott essay is vague about the operating principles of calibration, we turn elsewhere for specificity. For instance, in a recent description of the use of numerical models in the earth sciences, Oreskes, Shrader-Frechette and Belitz (1994, pp. 642, 643) describe calibration as follows:

> In earth sciences, the modeler is commonly faced with the inverse problem: The distribution of the dependent variable (for example, the hydraulic head) is the most well known aspect of the system; the distribution of the independent variable is the least well known. The process of tuning the model—that is, the manipulation of the independent variables to obtain a

[5] Simulation is also widely used in physical science. For example, it is customary in the studies of fractal dynamics to simulate models in order to gain understanding of the properties of models with various parameter configurations (Peitgen and Richter, 1986).

[6] As best we can tell from their essay, Kydland and Prescott want to preserve the term "estimation" to apply to the outputs of their computational experiments.

> match between the observed and simulated distribution or distributions of a dependent variable or variables—is known as calibration.
>
> Some hydrologists have suggested a two-step calibration scheme in which the available dependent data set is divided into two parts. In the first step, the independent parameters of the model are adjusted to reproduce the first part of the data. Then in the second step the model is run and the results are compared with the second part of the data. In this scheme, the first step is labeled "calibration" and the second step is labeled "verification."

This appears to be an accurate description of the general features of the "calibration" method advocated by Kydland and Prescott. For them, data for the *first step* come from micro observations and from secular growth observations (see also Prescott, 1986a). Correlations over time and across variables are to be used in the *second step* of verification.

Econometricians refer to the first stage as *estimation* and the second stage as *testing*. As a consequence, the two-stage procedure described by Oreskes, Shrader-Frechette and Belitz (1994) has a straightforward econometric counterpart.[7]

From this perspective, the Kydland-Prescott objection to mainstream econometrics is simply a complaint about the use of certain loss functions for describing the fit of a model to the data or for producing parameter estimates. Their objection does not rule out econometric estimation based on other loss functions.

Econometric estimation metrics like least squares, weighted least squares or more general method-of-moments metrics are traditional measures of fit. Difference among these methods lie in how they weight various features of the data; for example, one method might give a great deal of weight to distant outliers or to certain variables, causing them to pull estimated trend lines in their direction; another might give less weight to such outliers or variables. Each method of estimation can be justified by describing the particular loss function that summarizes the weights put on deviations of a model's predictions from the data. There is nothing sacred about the traditional loss functions in econometrics associated with standard methods, like ordinary least squares. Although traditional approaches do have rigorous justifications, a variety of alternative loss functions could be explored that weight particular features of a model more than others. For example, one could estimate with a loss function that rewarded models that are more successful in predicting turning points. Alternatively, particular time series frequencies could be deemphasized in adopting an estimation criterion because misspecification of a model is likely to contaminate some frequencies more than others (Dunsmuir and Hannan, 1978; Hansen and Sargent, 1993; Sims, 1993).

[7] See Christiano and Eichenbaum (1992) for one possible econometric implementation of this two-step approach. They use a generalized method of moments formulation (for example, Hansen, 1982) in which parameters are estimated by a first stage, exactly identified set of moment relations, and the model is tested by looking at another set of moment restrictions. Not surprisingly, to achieve identification of the underlying set of parameters, they are compelled to include more than just secular growth relations in the first-stage estimation, apparently violating one of the canons of current calibration practice.

The real business cycle practitioners adopt implicit loss functions. In looking at economic aggregates, their implicit loss functions appear to focus on the model predictions for long-run means, to the exclusion of other features of the data, when selecting parameter estimates. It is unfortunate that we are forced to guess about the rationale for the loss functions implicit in their research. There is little emphasis on assessing the quality of the resulting calibration. Formalizing the criteria for calibration and verification via loss functions makes the principle by which a particular model is chosen easier to understand. A clear statement would lead to more fruitful and focused conversations about the sources and reliability of estimated parameters.

As Oreskes, Shrader-Frechette and Belitz (1994) emphasize, the distinction between calibration and verification is often contrived. In many circumstances the verification step should really be considered part of the "calibration" step. The absence of a sharp distinction between these two stages is consistent with the difficulty of obtaining testable implications from the general equilibrium paradigm. Model testing serves as a barometer for measuring whether a given parametric structure captures the essential features of the data. When cleverly executed, it can pinpoint defective features of a model. Applied statistical decision theory and conventional statistical practice provide a formalism for conducting this endeavor. While this theory can be criticized for its rigidity or its naiveté, it seems premature to scrap it altogether without putting in place some other clearly stated criterion for picking the parameters of a model and assessing the quality of that selection.

The rational agents in a model of the Kydland-Prescott type rely explicitly on loss functions. After all, their rational decision making is based on the application of statistical decision theory, and part of the Kydland-Prescott line of research is to welcome the application of this theory to modern macroeconomics. But the idea of a loss function is also a central concept in statistical decision theory (LeCam and Yang, 1990). The rational agents in real business cycle models use this theory and, as a consequence, are assumed to process information in a highly structured way. Why should the producers of estimates for the real business cycle models act differently?

Although Kydland and Prescott are not precise in this essay in stating how calibration should be done in practice, there is much more specificity in Prescott (1986a, p. 14), who writes: "The key parameters of the growth model are the intertemporal and intratemporal elasticities of substitution. As Lucas (1980, p. 712) emphasizes, 'On these parameters, we have a wealth of inexpensively available data from census and cohort information, from panel data describing market conditions and so forth.'"

It is instructive to compare Prescott's optimistic discussion of the ease of using micro data to inform calibration with the candid and informative discussion of the same issue by Shoven and Whalley (1992, p. 105), who pioneered the application of calibration methods in general equilibrium analysis. They write:

> Typically, calibration involves only one year's data or a single observation represented as an average over a number of years. Because of the reliance on a single observation, benchmark data typically does not identify a unique set of values for the parameters in any model. Particular values for the relevant

> elasticities are usually required, and are specified on the basis of other research. These serve, along with the equilibrium observation, to uniquely identify the other parameters of the model. This typically places major reliance on literature surveys of elasticities; as many modelers have observed in discussing their own work, it is surprising how sparse (and sometimes contradictory) the literature is on some key elasticity values. And, although this procedure might sound straightforward, it is often exceedingly difficult because each study is different from every other.

What is noteworthy about this quotation is that the authors are describing a deterministic general equilibrium model based on traditional models of factor demand, sectoral output, product supply, labor supply and demand for final products, which have been the focus of numerous micro empirical studies. There have been many fewer micro empirical studies of the sectoral components of the stochastic general equilibrium models used in real business cycle theory. If there are few well-tested models that Shoven and Whalley can pull off the shelf, is it plausible that the shelf is unusually rich in models estimated assuming the relevant economic agents are operating in the more general economic environments considered in real business cycle theory?

Shoven and Whalley (1992, p. 106) come close to acknowledging the fundamental underidentification of general equilibrium systems from time series data when they write:

> [I]n some applied models many thousands of parameters are involved, and to estimate simultaneously all of the model parameters using time-series methods would require either unrealistically large numbers of observations or overly severe identifying restrictions. . . . Thus far, these problems have largely excluded complete econometric estimation of general equilibrium systems in applied work.

Current real business cycle models often require many fewer parameters to be calibrated, because they are highly aggregated. However, the extraction of the required elasticities from microeconometric analyses is more problematic, because the implicit economic environments invoked to justify microeconometric estimation procedures seldom match the dynamic stochastic single-agent models for which the micro estimates act as inputs. Microeconomic studies rarely estimate models that can be directly applied to the aggregates used in real business cycle theory. Moreover, as the specification of the real business cycle models become richer, they will inevitably have to face up to the same concerns that plague Shoven and Whalley.[8]

[8] This problem has already surfaced in the work of Benhabib, Rogerson and Wright (1991). They try to identify the parameters of a household production function for the services from durable goods using Panel Survey of Income Dynamics data, but without data on one of the inputs (the stock of durable goods), poor data on the other input (time spent by the household required to make durable goods productive) and no data on the output.

The Real Business Cycle Empirical Method In Practice

Kydland and Prescott, along with other real business cycle practitioners, endorse the use of time series averages—but not correlations—in calibrating models. In their proposed paradigm for empirical research, correlations are to be saved and used to test models, but are not to be used as a source of information about parameter values. It has become commonplace in the real business cycle research program to match the steady-state implications of models to time series averages. To an outsider, this looks remarkably like a way of doing estimation without accounting for sampling error in the sample means. In fact, the real business cycle "calibration" estimator of the Cobb-Douglas share parameter is a classical econometric estimator due to Klein (Klein, 1953; Nerlove, 1965). The only difference is that the Klein estimator usually is presented with a standard error.

Why is it acceptable to use sample means as a valid source of information about model parameters and not sample autocorrelations and cross correlations? Many interesting parameters cannot be identified from population means alone. Although the real business cycle literature provides no good reason for not using other sample moments, some reasons could be adduced. For example, one traditional argument for using sample means is that they are robust to measurement error in a way that sample variances and covariances are not as long as the errors have mean zero. Another possible rationale is that steady-state relations are sometimes robust with respect to alternative specifications of the short-run dynamics of a model. In these cases, a calibration fit to sample means will be consistent with a class of models that differ in their implications for short-run dynamics. However, the other side of this coin is that long-term means identify the short-run dynamics of a model only in very special circumstances. Moreover, as pointed out by Sargent (1989), even with measurement error, time series correlations and cross correlations can still provide more information about a model than is conveyed in sample means.

Since the models considered by Kydland and Prescott are stochastic, it is not in general possible to calibrate all of the parameters of a model solely from the means of macro time series. Computational experiments make assumptions about the correlation among the stochastic inputs to the model. Information about shocks, such as their variances and serial correlations, are needed to conduct the computational experiments. In a related vein, macroeconomic correlations contain potentially valuable information about the mechanism through which shocks are transmitted to macro time series. For models with richer dynamics, including their original "time-to-build" model, Kydland and Prescott (1982) envision fully calibrating the transmission mechanisms from micro evidence; but they provide no defense for avoiding the use of macro correlations in that task.

Recently, Cogley and Nason (1995) have criticized models in the literature spawned by Kydland and Prescott for failing to generate business cycle dynamics (see also Christiano, 1988; Watson, 1993; Cochrane, 1994). Since matching the full set of dynamics of the model to the dynamics in the data is not an essential part of calibration methodology, these models survive the weak standards for verification imposed by the calibrators. A much more disciplined and systematic exploration of the intertemporal and cross correlations, in a manner now routine in time series econometrics, would have shifted the focus from the empirical successes to the empirical challenges. We agree with Oreskes, Shrader-Frechette and Belitz (1994) that the distinction between calibration and verification is commonly blurred in practice. In the case of real business cycle research, such blurring is likely to be all the more prevalent as the models are redesigned to incorporate richer transient dynamics and additional sources of uncertainty.

As Kydland and Prescott emphasize, one of the most important questions for macroeconometrics is the quantitative importance of alternative sources of business cycle fluctuations. This classical problem has not yet been definitively answered (Morgan, 1990, pt. I). Using intuition from factor analysis, it is impossible to answer this question from a single time series. From two time series, one can isolate a single common factor. (Intuitively, two random variables can always be decomposed into a common component and two uncorrelated components.) Only using multiple time series is it possible to sort out multiple sources of business cycle shocks. The current emphasis in the literature on using only a few "key correlations" to check a model's implications makes single-factor explanations more likely to emerge from real business cycle analyses.[9] The idiosyncratic way Kydland and Prescott quantify the importance of technology shocks unfortunately makes it difficult to compare their answers to those obtained from the "innovation accounting" methods advocated by Sims (1980) and used extensively in empirical macroeconomics or to those obtained using the dynamic factor models of Geweke (1977) and Sargent and Sims (1977). Kydland and Prescott's answer to the central question of the importance of technology shocks would be much more credible if it were reinforced by other empirical methodologies.

A contrast with John Taylor's approach to investigating the properties of models is instructive. Taylor's research program includes the use of computational experiments. It is well summarized in his recent book (Taylor, 1993). Like Kydland and Prescott, Taylor relies on fully specified dynamic models and imposes rational expectations when computing stochastic equilibria. However, in fitting linear models he uses all of the information on first and second moments available in the macro data when it is computationally possible to do so. The econometric methods used in parameter estimation are precisely described. Multiple sources of business cycle shocks are admitted into the model at the outset, and rigorous empirical testing of models appears throughout his analyses.[10]

[9] In private correspondence, John Taylor has amplified this point: "I have found that the omission of aggregate price or inflation data in the Kydland-Prescott second moment exercise creates an artificial barrier between real business cycle models and monetary models. To me, the Granger causality from inflation to output and vice versa are key facts to be explained. But Kydland and Prescott have ignored these facts because they do not fit into their models."

[10] Fair (1994) presents an alternative systematic approach to estimation and simulation, but unlike Taylor, he does not impose rational expectations assumptions.

If the Kydland and Prescott real business cycle research program is to achieve empirical credibility, it will have to provide a much more comprehensive assessment of the successes and failures of its models. To convince a wide audience of "outsiders," the proclaimed successes in real business cycle calibration should not be intertwined with an idiosyncratic and poorly justified way of evaluating models. We sympathize with Fair (1992, p. 141), who writes:

> Is the RBC [real business cycle] approach a good way of testing models? At first glance it might seem so, since computed paths are being compared to actual paths. But the paths are being compared in a very limited way in contrast to the way that the Cowles Commission approach would compare them. Take the simple RMSE [root mean square error[11]] procedure. This procedure would compute a prediction error for a given variable for each period and then calculate the RMSE from another structural model or from an autoregressive or vector autoregressive model.
>
> I have never seen this type of comparison done for a RBC model. How would, say, the currently best-fitting RBC model compare to a simple first-order autoregressive equation for real GNP in terms of the RMSE criterion? My guess is very poorly. Having the computed path mimic the actual path for a few selected moments is a far cry from beating even a first-order autoregressive equation (let alone a structural model) in terms of fitting the observations well according to the RMSE criterion. The disturbing feature of the RBC literature is there seems to be no interest in computing RMSEs and the like. People generally seem to realize that the RBC models do not fit well in this sense, but they proceed anyway.

Specification Uncertainty Underlies the Estimates

One of the most appealing features of a research program that builds dynamic macroeconomic models on microeconomic foundations is that it opens the door to the use of micro empirical evidence to pin down macro parameter values. Kydland and Prescott and the entire real business cycle community pay only lip service to the incompatibility between the macroeconomic model used in their computational experiments and the microeconometric models used to secure the simulation parameters.

It can be very misleading to plug microeconometric parameter estimates into a macroeconomic model when the economic environments for the two models are fundamentally different. In fact, many of the micro studies that the "calibrators" draw upon do not estimate the parameters required by the models being simulated.

[11] RMSE is the square root of the mean of the squared discrepancies between predicted and actual outcomes.

This creates specification uncertainty (Leamer, 1978). To adequately represent this uncertainty, it is necessary to incorporate the uncertainty about model parameters directly into the outputs of simulations. Standard errors analogous to those presented by Christiano and Eichenbaum (1992) and Eichenbaum (1991) are a useful first step, but do not convey the full picture of model uncertainty. What is required is a sensitivity analysis to see how the simulation outputs are affected by different choices of simulation parameters. Trostel (1993) makes effective use of such a methodology.

Consider using the estimates of intertemporal labor supply produced by Ghez and Becker (1975) for simulation purposes.[12] Ghez and Becker (1975) estimate the intertemporal substitution of leisure time assuming perfect credit markets, no restrictions on trade in the labor market and no fixed costs of work. This study is important, but like all empirical work in economics, the precise estimates are enveloped by some uncertainty. Moreover, different estimation schemes are required to secure this parameter if there is uninsurable uncertainty in the environment (MaCurdy, 1978). Even looking only at estimates of the intertemporal substitution of leisure based on models that assume that workers can perfectly insure, the point estimates reported in the literature are very imprecisely determined (MaCurdy, 1981; Altonji, 1986). Further, it is not clear how the estimates should be modified to be compatible with the other economic environments including settings that allow for uninsurable uncertainty, transactions costs and restrictions on trades in the market.

Current practices in the field of calibration and simulation do not report either estimation error and/or model-specification error. Nor is it a standard feature of real business cycle practice to present formal analyses that explore how sensitive the simulations are to different parameter values. Precise numerical outputs are reported, but with no sense of the confidence that can be placed in the estimates. This produces a false sense of precision.

Observationally Equivalent Models Offer Different Predictions about Policy Interventions

While putting on empirical "blinders" permits a particular line of research to proceed, looking at too narrow of a range of data makes identification problems more severe. A disturbing feature of current practice in the real business cycle

[12] Kydland and Prescott cite Ghez and Becker (1975) as a prime example of the value of microeconomic empirical work. However, their citation misses two central aspects of that work. First, Ghez and Becker (1975) use synthetic cohort data, not panel data as stated by Kydland and Prescott. Second, the interpretation of the Ghez-Becker estimates as structural parameters is predicated on a list of identifying assumptions. These assumptions coupled with the resulting estimates are the most important part of their investigation, not their observation that people sleep eight hours a day, which is what Kydland and Prescott emphasize.

literature is that models with the same inputs can produce fundamentally different computed values of welfare loss and quantitative assessments of alternative economic policies.

Consider the following developments in the field of empirical finance. A frequently noted anomaly is that the observed differences in returns between stocks and bonds are too large to be consistent with models of the preferences commonly used in real business cycle analysis (Hansen and Singleton, 1983; Mehra and Prescott, 1985; Cochrane and Hansen, 1992; Kocherlakota, 1996). One response to these asset-pricing anomalies has been the modification to preferences developed by Epstein and Zin (1989), which breaks the tight link between intertemporal substitution and risk aversion that was maintained in the preceding literature. A parallel advance has been the introduction of intertemporal complementarities such as habit persistence in the preference orderings of consumers (Constantinides, 1990). Hansen, Sargent and Tallarini (1995) find that models with Epstein-Zin type preferences and models without this form of risk sensitivity explain the same quantity data but have fundamentally different implications for the market price of risk (the slope of the mean-standard deviation frontier for asset returns).[13] These "observationally equivalent" preference specifications produce very different estimates of the welfare losses associated with hypothetical policy interventions. The decision by other researchers such as Epstein and Zin to look more broadly at available data and to emphasize model defects instead of successes provoked quantitatively important advances in economic theory.

Another competing explanation for the equity premium puzzle is the presence of incomplete markets and transactions costs in asset markets. This explanation is consistent with Prescott's (1986b, p. 29) earlier argument for ignoring asset market data in real business cycle calibrations: "That the representative agent model is poorly designed to predict differences in borrowing and lending rates . . . does not imply that this model is not well suited for other purposes—for predicting the consequences of technology shocks for fluctuations in business cycle frequencies, for example."

Heaton and Lucas (1995) quantify the magnitude of transaction costs needed to address the equity-premium puzzle (see also Aiyagari and Gertler, 1991). Prescott may be correct that such models will not help to match "key" correlations in economic aggregates, but this requires documentation. Even if there is robustness of the form hoped for by Prescott, the presence of transactions costs of the magnitude suggested by Heaton and Lucas (1995) are likely to alter the welfare comparisons across different policy experiments in a quantitatively important way.[14] This is so because transactions costs prevent heterogeneous consumers from equating marginal rates of substitution and put a wedge between marginal rates of substitution and marginal rates of transformation.

[13] Recent work by Campbell and Cochrane (1995) and Boldrin, Christiano and Fisher (1995) suggests a similar conclusion for models with strong intertemporal complementarities.

[14] This sensitivity actually occurs in the "bothersome" experiment of İmrohorğlu (1992) mentioned by Kydland and Prescott.

A More Constructive Research Program

The idea of using micro data to enrich the information in macro time series dates back at least to the writings of Tobin (1950). A careful reading of the literature that accompanied his suggestion reveals that his idea was inherently controversial, especially if the micro information is based on cross-section data, and if the behavioral equations are dynamic (Aigner and Simon, 1970; Simon and Aigner, 1970). This issue was revived in the late 1970s and early 1980s when numerous economists attempted to estimate micro labor supply equations to test the Lucas and Rapping (1969) intertemporal substitution hypothesis. The hypothesis rests critically on consumer responses to expected real discounted future wage movements relative to current wages. By providing well-focused economic questions, the Lucas-Rapping model advanced the development of empirical microeconomics by challenging economists to supply answers. Numerous micro studies of labor supply were conducted with an eye toward confirming or disconfirming their hypothesis (Altonji and Ashenfelter, 1980; MaCurdy, 1981; Ham, 1982; Altonji, 1986).

However, these studies reveal that even with large micro samples, it is not possible to estimate the parameter of interest precisely. Measurement error in micro data and selection problems often limit the value of the information in the micro data. Macro time series or aggregated cross sections can sometimes solve selection problems that are intractable in micro data (Heckman and Robb, 1985, pp. 168–169, 210–213). Different micro survey instruments produce fundamentally different descriptions of the same phenomena (Smith, 1995). Micro data are no panacea. Moreover, the recent movement in empirical microeconomics away from economic models to "simple descriptive" estimation schemes has reduced the supply of new structural parameters.

It is simply not true that there is a large shelf of micro estimates already constructed for different economic environments that can be plugged without modification into a new macro model. In many cases, estimators that are valid in one economic environment are not well suited for another. Given the less-than-idyllic state of affairs, it seems foolish to look to micro data as the primary source for many macro parameters required to do simulation analysis. Many crucial economic parameters—for example, the effect of product inputs on industry supply—can only be determined by looking at relationships among aggregates. Like it or not, time series evidence remains essential in determining many fundamentally aggregative parameters.

A more productive research program would provide clearly formulated theories that will stimulate more focused microeconomic empirical research. Much recent micro research is atheoretical in character and does not link up well with macro general equilibrium theory. For example, with rare exceptions, micro studies treat aggregate shocks as nuisance parameters to be eliminated by some trend or dummy variable procedure.[15] A redirection of micro empirical work toward providing input

[15] For an exception see Heckman and Sedlacek (1985), who show how cross-section time dummies can be used to estimate the time series of unobserved skill prices in a market model of self-selection.

into well-defined general equilibrium models would move discussions of micro evidence beyond discussions of whether wage or price effects exist, to the intellectually more important questions of what the micro estimates mean and how they can be used to illuminate well-posed economic questions. "Calibrators" could make a constructive contribution to empirical economics by suggesting a more symbiotic relationship between the macro general equilibrium model as a synthesizing device and motivating vehicle and the micro evidence as a source of robust parameter values.

Recently there has been considerable interest in heterogeneous agent models in the real business cycle literature; Ríos-Rull (1995) offers a nice summary. To us, one of the primary reasons for pushing this line of research is to narrow the range of specification errors in calibrating with microeconomic data. Microeconometric estimates routinely incorporate heterogeneity that is often abstracted from the specification of dynamic, stochastic general equilibrium models. It is remarkable to us that so little emphasis has been given to the transition from micro to macro in the real business cycle literature, given that understanding the distribution of heterogeneity is central to making this transition (Stoker, 1993).

The Kydland and Prescott program is an intellectually exciting one. To date, however, the computations produced from it have only illustrated some of the qualitative properties of some dynamic stochastic models and demonstrated the possibility of executing an array of interesting calculations. The real business cycle modeling effort would be more beneficial if it shifted its focus to micro predictions and in this way helped to stimulate research on empirical models that would verify or contradict the macro models.

We envision a symbiotic relationship between calibrators and empirical economists in which calibration methods like those used by Frisch, Tinbergen, and Kydland and Prescott stimulate the production of more convincing micro empirical estimates by showing which gaps in our knowledge of micro phenomenon matter and which gaps do not. Calibration should only be the starting point of an empirical analysis of general equilibrium models. In the absence of firmly established estimates of key parameters, sensitivity analyses should be routine in real business cycle simulations. Properly used and qualified simulation methods can be an important source of information and an important stimulus to high-quality empirical economic research.

The research program we advocate is not an easy one. However, it will be an informative one. It will motivate micro empirical researchers to focus on economically interesting questions; it will secure the foundations of empirical general equilibrium theory; and, properly executed, it will demonstrate both the gaps and strengths of our knowledge on major issues of public policy.

■ *We thank Jennifer Boobar, John Cochrane, Marty Eichenbaum, Ray Fair, Chris Flinn, John Heaton, Bob Lucas, Tom Sargent, Jeff Smith, Nancy Stokey, John Taylor and Grace Tsiang for their valuable comments on this draft. Hansen's research is supported in part by NSF SBR-94095–01; Heckman is supported by NSF 93-048-0211 and grants from the Russell Sage Foundation and the American Bar Foundation.*

References

Aigner, Dennis, and Julian L. Simon, "A Specification BIAS Interpretation of Cross-Section vs. Time-Series Parameter Estimates," *Western Economic Journal,* June 1970, *22,* 144–61.

Aiyagari, S. Rao, and Mark Gertler, "Asset Returns with Transitions Costs and Uninsured Individual Risk," *Journal of Monetary Economics,* June 1991, *27,* 311–31.

Altonji, Joseph G., "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy,* June 1986, *94*:3, Part 2, S176–S215.

Altonji, Joseph G., and Orley Ashenfelter, "Wage Movements and the Labour Market Equilibrium Hypothesis," *Economica,* August 1980, *49,* 783–824.

Benhabib, J., R. Rogerson, and R. Wright, "Homework in Macroeconomics: Household Production and Aggregate Fluctuations," *Journal of Political Economy,* December 1991, *99,* 1166–87.

Boldrin, Michele, Lawrence J. Christiano, and J. D. M. Fisher, "Asset Pricing Lessons for Modeling Business Cycles," manuscript, August 1995.

Brown, D. J., and R. L. Matzkin, "Testable Restrictions on the Equilibrium Manifold," manuscript, April 1995.

Campbell, John, and John H. Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." National Bureau of Economic Research Working Paper No. 4995, 1995.

Christiano, Lawrence J., "Why Does Inventory Investment Fluctuate So Much?," *Journal of Monetary Economics,* March/May 1988, *21,* 247–80.

Christiano, Lawrence J., and Martin Eichenbaum, "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations," *American Economic Review,* June 1992, *82,* 430–50.

Cochrane, John H., "Shocks," *Carnegie-Rochester Conference. Series on Public Policy,* Spring 1994, *41,* 295–364.

Cochrane, John H., and Lars Peter Hansen, "Asset Pricing Explorations for Macroeconomics." In Blanchard, O. J., and S. Fischer, eds., *NBER Macroeconomics Annual 1992.* Cambridge, Mass.: Massachusetts Institute of Technology Press, 1992, pp. 115–69.

Cogley, T., and J. M. Nason, "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series Implications for Business Cycle Research," *Journal of Economic Dynamics and Control,* January/February 1995, *19,* 253–78.

Constantinides, George M., "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy,* June 1990, *98,* 519–43.

Dunsmuir, W., and E. J. Hannan, "Vector Linear Time Series Models," *Advances in Applied Probability,* 1978, *8:2,* 339–64.

Eichenbaum, M., "Real Business Theory: Wisdom or Whimsy," *Journal of Economic Dynamics and Control,* October 1991, *15,* 607–26.

Epstein, Larry G., and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica,* July 1989, *57,* 937–69.

Fair, Ray, "The Cowles Commission Approach, Real Business Cycle Theories, and New-Keynesian Economics." In Belongia, M. T., and M. R. Garfinkel, eds., *The Business Cycle: Theories and Evidence.* Boston: Kluwer Academic, 1992, pp. 133–47.

Fair, Ray, *Testing Macroeconomic Models.* Cambridge: Harvard University Press, 1994.

Frisch, Ragnar, "Propagation Problems and Impulse Problems in Dynamic Economics." In *Economic Essays in Honor of Gustav Cassel.* London: Allen and Unwin, 1933, pp. 171–205.

Geweke, John, "The Dynamic Factor Analysis of Economic Time Series Models." In Aigner, D. J., and A. S. Goldberger, eds., *Latent Variables in Socio-Economic Models.* Amsterdam: North-Holland, 1977, pp. 365–83.

Ghez, G., and Gary Becker, *The Allocation of Time and Goods Over the Life Cycle.* New York: National Bureau of Economic Research, 1975.

Goldberger, Arthur, *Impact Multipliers and Dynamic Properties of the Klein-Golberger Model.* Amsterdam: North-Holland, 1959.

Gorman, William M., "Community Preference Fields," *Econometrica,* January 1953, *21,* 63–80.

Ham, John C., "Estimation of a Labour Supply Model with Censoring Due to Unemployment and Underemployment," *Review of Economic Studies,* July 1982, *49,* 335–54.

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* July 1982, *50,* 1029–54.

Hansen, Lars Peter, and Thomas J. Sargent, "Seasonality and Approximation Errors in Rational Expectations Models," *Journal of Econometrics,* January/February 1993, *55,* 21–55.

Hansen, Lars Peter, and K. J. Singleton, "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy,* 1983, *91*:2, 249–65.

Hansen, Lars Peter, Thomas J. Sargent, and

T. D. Tallarini, "Pessimism and Risk in General Equilibrium," manuscript, 1995.

Harrison, J. Michael, and David M. Kreps, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, June 1979, *20*, 381–408.

Heaton, John, and D. Lucas, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy*, forthcoming, June 1996.

Heckman, James, and Richard Robb, Jr., "Alternative Methods for Evaluating the Impact of Interventions." In Heckman, J., and Burton Singer, eds., *Longitudinal Analysis of Labor Market Data.* Cambridge, U.K.: Cambridge University Press, 1985, pp. 156–245.

Heckman, James J., and Guilherme Sedlacek, "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, December 1985, *93*, 1077–125.

Houthakker, Hendrik, "The Pareto Distribution and the Cobb-Douglas Production Function," *Review of Economic Studies*, 1956, *23*:1 27–31.

İmrohoroğlu, Ayşe, "The Welfare Costs of Inflation Under Imperfect Insurance," *Journal of Economic Dynamics and Control*, January 1992, *16*, 79–91.

Klein, L. R., *A Textbook of Econometrics.* Evanston, Ill.: Row, Peterson and Company, 1953.

Klein, L. A., and A. S. Goldberger, *An Econometric Model of the United States, 1929–52.* Amsterdam: North-Holland, 1955.

Kocherlakota, N., "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, forthcoming 1996.

Kydland, Finn E., and Edward C. Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica*, 1982, *50*:6, 1345–70.

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

LeCam, L., and G. Yang, *Asymptotics in Statistics: Some Basic Concepts.* Berlin: Springer-Verlag, 1990.

Liu, T. C., "Underidentification, Structural Estimation, and Forecasting," *Econometrica*, October 1960, *28*, 855–65.

Lucas, Robert E., Jr., "Methods and Problems in Business Cycle Theory," *Journal of Money, Credit and Banking*, November 1980, *12*, 696–715; reprinted in Lucas, R. E., ed., *Studies in Business Cycle Theory.* Cambridge, Mass.: Massachusetts Institute of Technology Press, 1981, pp. 271–96.

Lucas, Robert E., Jr., and Edward C. Prescott, "Investment Under Uncertainty," *Econometrica*, 1971, *39*:5, 659–81.

Lucas, Robert E., Jr., and Leonard A. Rapping, "Real Wages, Employment, and Inflation," *Journal of Political Economy*, September/October 1969, *77*, 721–54.

Lucas, Robert E., Jr., and Thomas J. Sargent, "After Keynesian Macroeconomics," *Quarterly Review, Federal Reserve Bank of Minneapolis*, Spring 1979, *3*, 1–16.

MaCurdy, Thomas, "Econometric Model of the Labor Supply in a Life Setting," Ph.D. dissertation, University of Chicago, 1978.

MaCurdy, Thomas, "An Empirical Model of Labor Supply in a Life Cycle Setting," *Journal of Political Economy*, 1981, *89*:6, 1059–85.

Mantel, Rolf R., "On the Characterization of Aggregate Excess Demand," *Journal of Economic Theory*, March 1974, *7*, 348–55.

Mehra, Rajnish, and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, March 1985, *15*, 145–61.

Morgan, M. S., *The History of Econometrics Ideas.* Cambridge: Cambridge University Press, 1990.

Nerlove, M., *Estimation and Identification of Cobb-Douglas Production Functions.* Chicago: Rand McNally, 1965.

Oreskes, Naomi, Kristen Shrader-Frechette, and Kenneth Belitz, "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences," *Science*, February 4, 1994, *263*, 641–46.

Peitgen, H.-O., and P. Richter, *The Beauty of Fractals.* Berlin: Springer-Verlag, 1986.

Prescott, Edward, "Theory Ahead of Business-Cycle Measurement," *Carnegie-Rochester Conference on Public Policy*, 1986a, *25*, 11–44; reprinted in *Quarterly Review, Federal Reserve Bank of Minneapolis*, Fall 1986, *10*, 9–22.

Prescott, Edward, "Response to a Skeptic," *Quarterly Review, Federal Reserve Bank of Minneapolis*, Fall 1986b, *10*, 28–33.

Ríos-Rull, José-Víctor, "Models with Heterogeneous Agents." In Cooley, T. F., ed., *Frontiers of Business Cycle Research.* Princeton, N.J.: Princeton University Press, 1995, pp. 98–125.

Rubinstein, Mark, "An Aggregation Theorem for Securities Markets," *Journal of Financial Economics*, September 1974, *1*, 225–44.

Sargent, Thomas J., "Two Models of Measurements, and the Investment Accelerator," *Journal of Political Economics*, 1989, *97*:2, 251–87.

Sargent, Thomas J., and Christopher A. Sims, "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory." In Sims, Christopher, ed., *New Methods in Business Cycle Research.* Minneapolis: Federal Reserve Bank, 1977, pp. 45–110.

Scarf, Herbert, and T. Hansen, *The Computation of Economic Equilibria.* New Haven: Yale University Press, 1973.

Shoven, John B., and John Whalley, "A General Equilibrium Calculation of the Effects of Differential Taxation of Income from Capital in the U.S.," *Journal of Public Economics,* November 1972, *1,* 281–322.

Shoven, John B., and John Whalley, *Applying General Equilibrium.* New York: Cambridge University Press, 1992.

Simon, Julian L., and Dennis J. Aigner, "Cross Sectional Budget Studies, Aggregate Time Series Studies and the Permanent Income Hypothesis," *American Economic Review,* June 1970, *60:*2, 526–41.

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica,* 1980, *48:*1, 1–48.

Sims, Christopher A., "Rational Expectations Modeling with Seasonally Adjusted Data," *Journal of Econometrics,* January/February 1993, *55,* 9–19.

Smith, J., "A Comparison of the Earnings Patterns of Two Samples of JTPA Eligibles," unpublished paper, University of Western Ontario, London, Canada, August 1995.

Sonnenschein, H., "Do Walres Identity and Continuity Characterize the Class of Community Excess Demand Functions?," *Journal of Economic Theory,* August 1973, 6, *345–54.*

Stoker, T. M., "Empirical Approaches to the Problem of Aggregation over Individuals," *Journal of Economic Literature,* December 1993, *31:*4, 1827–74.

Taylor, John B., *Macroeconomic Policy in a World Economy: From Econometric Design to Practical Operation.* New York: W. W. Norton and Company, 1993.

Tinbergen, J., *Statistical Testing of Business Cycle Theories.* Vol. 2, Business Cycles in the USA, 1919-1932, Geneva: League of Nations, 1939.

Tobin, James, "A Statistical Demand Function for Food in the USA," *Journal of the Royal Statistical Society,* 1950, *113,* Series A, Part II, 113–41.

Trostel, P. A., "The Effect of Taxation on Human Capital," *Econometrica,* April 1993, *101,* 327–50.

Watson, M. W., "Measures of Fit for Calibrated Models," *Journal of Political Economy,* 1993, *101:*6, 1011–41.

Wilson, Robert B., "The Theory of Syndicates," *Econometrica,* January 1968, *36,* 119–32.

This article has been cited by:

1. Nicolai V. Kuminoff, V. Kerry Smith, Christopher Timmins. 2013. The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets. *Journal of Economic Literature* **51**:4, 1007-1062. [Abstract] [View PDF article] [PDF with links]
2. Hervé Le Bihan,, Jérémi Montornès,, Thomas Heckel. 2012. Sticky Wages: Evidence from Quarterly Microeconomic Data. *American Economic Journal: Macroeconomics* **4**:3, 1-32. [Abstract] [View PDF article] [PDF with links]
3. Costas Arkolakis,, Arnaud Costinot,, Andrés Rodríguez-Clare. 2012. New Trade Models, Same Old Gains?. *American Economic Review* **102**:1, 94-130. [Abstract] [View PDF article] [PDF with links]
4. James J. Heckman. 2010. Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy. *Journal of Economic Literature* **48**:2, 356-398. [Abstract] [View PDF article] [PDF with links]
5. 2008. Comments. *Journal of Economic Perspectives* **22**:1, 243-249. [Citation] [View PDF article] [PDF with links]
6. Yongsung Chang, Joao F. Gomes, Frank Schorfheide. 2002. Learning-by-Doing as a Propagation Mechanism. *American Economic Review* **92**:5, 1498-1520. [Abstract] [View PDF article] [PDF with links]
7. John Piggott,, John Whalley. 2001. VAT Base Broadening, Self Supply, and the Informal Sector. *American Economic Review* **91**:4, 1084-1094. [Citation] [View PDF article] [PDF with links]
8. Charles F. Manski,. 2000. Economic Analysis of Social Interactions. *Journal of Economic Perspectives* **14**:3, 115-136. [Abstract] [View PDF article] [PDF with links]
9. Francis X. Diebold. 1998. The Past, Present, and Future of Macroeconomic Forecasting. *Journal of Economic Perspectives* **12**:2, 175-192. [Abstract] [View PDF article] [PDF with links]
10. Richard Rogerson. 1997. Theory Ahead of Language in the Economics of Unemployment. *Journal of Economic Perspectives* **11**:1, 73-92. [Abstract] [View PDF article] [PDF with links]