Journal of Economic Perspectives—Volume 10, Number 1—Winter 1996—Pages 69–85

The Computational Experiment: An Econometric Tool

Finn E. Kydland and Edward C. Prescott

In a computational experiment, the researcher starts by posing a well-defined quantitative question. Then the researcher uses both theory and measurement to construct a model economy that is a computer representation of a national economy. A model economy consists of households, firms and often a government. The people in the model economy make economic decisions that correspond to those of their counterparts in the real world. Households, for example, make consumption and savings decisions, and they decide how much to work in the market. The researcher then calibrates the model economy so that it mimics the world along a carefully specified set of dimensions. Finally, the computer is used to run experiments that answer the question.[1]

Such experiments have become invaluable tools in quantitative aggregate theory.[2] They are being used, for example, to estimate the quantitative effects of trade liberalization policies, measure the welfare consequences of changes in the tax system and quantify the magnitude and nature of business cycle fluctuations induced by different types of shocks. In this paper, we review the use of the computational experiment in economics.

[1] Lucas (1980), in his paper on methods and problems in business cycle theory, explains the need for computational experiments in business cycle research.

[2] Shoven and Whalley (1972) were the first to use what we call the computational experiment in economics. The model economies that they used in their experiments are static and have many industrial sectors.

■ *Finn E. Kydland is Professor of Economics, Carnegie-Mellon University, Pittsburgh, Pennsylvania, and Research Associate, Federal Reserve Bank of Cleveland, Cleveland, Ohio. Edward C. Prescott is Professor of Economics, University of Minnesota, and Research Associate, Federal Reserve Bank of Minneapolis, both in Minneapolis, Minnesota.*

One immediate question that arises is whether the computational experiment should be regarded as an econometric tool (for example, Gregory and Smith, 1993). In the modern (narrow) sense of the term it is not, since it isn't used in the "measurement of economic relations" (Marschak, 1948, p. 1). Yet it is an econometric tool in the original (broad) sense of the term (which we prefer), since computational experiments are used to derive the quantitative implications of economic theory (Frisch, 1933a, p. 1). In Kydland and Prescott (1991a), we develop the position that the computational experiment is an econometric tool, but here, we avoid this largely semantic debate. Instead, we will simply state that the task of deriving the quantitative implications of theory differs from that of measuring economic parameters.

Computational experiments are not unique to economic science—they are heavily used in the physical sciences as well. In one crucial respect, however, they do differ across the two disciplines. Unlike theory in the physical sciences, theory in economics does not provide a law of motion. Rather, economic theory provides a specification of people's ability and willingness to substitute among commodities. Consequently, computational experiments in economics include the additional step of computing the equilibrium process in which all of the model's people behave in a way that is in each person's best interest—that is, economists must *compute* the equilibrium law of motion or process of the model economy. Given the process governing the system, both economic and physical science use the computer to generate realizations of this process.

If the model is deterministic, only one possible equilibrium realization exists for the path of the model economy. If the model economy has aggregate uncertainty—as it must, for example, if the phenomena of interest are business cycle fluctuations—then the model will imply a process governing the random evolution of the economy. In the case of uncertainty, the computer can generate any number of independent realizations of the equilibrium stochastic process, and these realizations, along with statistical estimation theory, are then used to measure the sampling distribution of any desired set of statistics of the model economy.

Steps in an Economic Computational Experiment

Any economic computational experiment involves five major steps: pose a question; use a well-tested theory; construct a model economy; calibrate the model economy; and run the experiment. We will discuss each of these steps in turn.

Pose a Question

The purpose of a computational experiment is to derive a quantitative answer to some well-posed question. Thus, the first step in carrying out a computational experiment is to pose such a question. Some of these questions are concerned with policy evaluation issues. These questions typically ask about the welfare and distributive consequences of some policy under consideration. Other questions are con-

Table 1
Examples of Well-Posed Questions in Studies Using the Computational Experiment

Studies Using Theory	*Question*
Auerbach and Kotlikoff (1987)	What are the effects of the current U.S. social security system on capital formation and intergenerational equity?
Brown, Deardorff and Stern (1994)	What are the potential welfare, wage rate and terms-of-trade effects of NAFTA on Canada, Mexico and the United States?
Gravelle and Kotlikoff (1995)	What are the welfare consequences of the 1986 U.S. tax reform?
Harris (1984)	What are the welfare gains for a small open economy with scale effects and imperfect competition?
Hopenhayn and Rogerson (1993)	What are the welfare costs of a job destruction tax on firms equal to one year's wages?
İmrohoroğlu (1992)	What are the welfare costs of inflation if insurance is imperfect?
Studies Developing Theory	***Question***
Backus, Kehoe and Kydland (1994)	Does theory imply the J-curve pattern of covariance between terms of trade and the trade balance?
Christiano and Eichenbaum (1992b) and Chang (1995)	What is the contribution of public finance shocks to aggregate fluctuations?
Finn (1995)	What is the contribution of oil shocks to business cycle fluctuations?
Greenwood, Hercowitz and Huffman (1988)	Does nonneutrality of technology shocks with respect to the consumption and investment good change the estimate of the contribution of technology shocks to business cycle fluctuations?
Hornstein (1993)	Does the introduction of monopolistic competition into real business cycle models alter the estimate of technology shocks contribution?
Kydland and Prescott (1982)	What is the quantitative nature of fluctuations induced by technology shocks?

cerned with the testing and development of theory. These questions typically ask about the quantitative implications of theory for some phenomena. If the answer to these questions is that the predictions of theory match the observations, theory has passed that particular test. If the answer is that there is a discrepancy, a deviation from theory has been documented. Still, other experiments are concerned with the sensitivity of previous findings to the introduction of some feature of reality from which previous studies have abstracted. Table 1 offers some examples of computational experiments that seek to answer each of these types of questions. That table highlights the fact that judging whether a model economy is a "good" abstraction can be done only relative to the question posed.

Use Well-Tested Theory

With a particular question in mind, a researcher needs some strong theory to carry out a computational experiment: that is, a researcher needs a theory that has been tested through use and found to provide reliable answers to a class of questions. Here, by theory we do not mean a set of assertions about the actual economy. Rather, following Lucas (1980), economic theory is defined to be "an explicit set of instructions for building . . . a mechanical imitation system" to answer a question. If the question is quantitative in nature, a computer representation of the imitation system or economy is used, and extensive computations are required to answer the posed question.

As one example, the computational experiments often carried out in modern business cycle theory build upon the neoclassical growth framework. Central to neoclassical growth theory is its use of an aggregate production function, with the output of goods resulting from inputs of labor and capital.[3] This framework has served well when dealing with growth within reasonably stable economic institutions. With an explicit description of the household sector, including its focus on the time-allocation decision, the neoclassical growth model becomes an internally consistent framework for addressing business cycle questions, as well as other questions of interest to macroeconomists. The theory implies that when a model economy is confronted with technology, public finance and terms-of-trade shocks, it should display business cycle fluctuations of a quantitative nature similar to those actually observed. In other words, modern business cycle models are stochastic versions of neoclassical growth theory. And the fact that business cycle models do produce normal-looking fluctuations adds dramatically to our confidence in the neoclassical growth theory model—including the answers it provides to growth accounting and public finance questions.

We recognize, of course, that although the economist should choose a well-tested theory, every theory has some issues and questions that it does not address well. In the case of neoclassical growth theory, for example, it fails spectacularly when used to address economic development issues. Differences in stocks of reproducible capital stocks cannot account for international differences in per capita incomes. This does not preclude its usefulness in evaluating tax policies and in business cycle research.

Construct a Model Economy

With a particular theoretical framework in mind, the third step in conducting a computational experiment is to construct a model economy. Here, key issues are the amount of detail and the feasibility of computing the equilibrium process. Often, economic experimenters are constrained to deal with a much simpler model

[3] Neoclassical growth theory also represents a good example of the importance of interaction between factual studies and theory development. Solow (1970) lists several growth facts that influenced the development of neoclassical growth theory. Once the main ingredients of the theory were established—such as the production function—new light was thrown on the data.

economy than they would like because computing the equilibrium of a more complicated model would be impossible, given currently available tools.

This situation is no different from that in the physical sciences, where, as in economics, the computational experiment has become accepted as an invaluable scientific tool. For example, in his overview of climate modeling, Schneider (1987, p. 72) states: "Although all climate models consist of mathematical representations of physical processes, the precise composition of a model and its complexity depend on the problem it is designed to address." And later (p. 72): "Often it makes sense to attack a problem first with a simple model and then employ the results to guide research at higher resolution." In the physical sciences, as in economics, confidence in a particular framework or approach is gained through successful use.

So far, most of the model environments that macroeconomists have used share certain characteristics. The environments are inhabited by a large number of people whose decision problems are described explicitly. Both the household and business sectors play a central role. For some questions, government or foreign sectors must be included as well. In some models everyone is alike; in others, such as those designed to address questions where demographic changes are important, heterogeneous people must be used.

This description may sound somewhat indefinite or abstract, but we reemphasize that an abstraction can be judged only relative to some given question. To criticize or reject a model because it is an abstraction is foolish: all models are necessarily abstractions. A model environment must be selected based on the question being addressed. For example, heterogeneity of people is crucial in the Auerbach and Kotlikoff (1987) model, which predicts the consequences of the population's changing age distribution on savings. However, Ríos-Rull (1994) demonstrates that such life cycle features, even when combined with elements of market incompleteness, are not quantitatively important to business cycle findings regarding issues such as the contribution of technology shocks to business cycle fluctuations. The features of a given model may be appropriate for some question (or class of questions) but not for others.[4]

The selection and construction of a particular model economy should not depend on the answer provided. In fact, searching within some parametric class of economies for the one that best fits a set of aggregate time series makes little sense, because it isn't likely to answer an interesting question. For example, if the question is of the type, "how much of fact *X* is accounted for by *Y*," then choosing the parameter values in such a way as to make the amount accounted for as large as

[4] We will not debate the legitimacy of these methods, since such debates generally serve to define schools rather than to produce agreement. Such debates are almost nonexistent during normal science, but tend to recur during scientific revolutions. As stated by Kuhn (1962, p. 145), "Few philosophers of science still seek absolute criteria for the verification of scientific theories." All historically significant theories have agreed with the facts, but only to a degree. No more precise answer can be found to the question of how well an individual theory fits the facts. Using probabilistic verification theories that ask us to compare a given scientific theory with all others that might fit the same data is a futile effort. We agree with Kuhn (p. 146) that "probabilistic theories disguise the verification situation as much as they illuminate it."

possible according to some metric is an attempt to get a particular—not a good—answer to the question.

Calibrate the Model Economy

Now that a model has been constructed, the fourth step in carrying out a computational experiment is to calibrate that model. Originally, in the physical sciences, *calibration* referred to the graduation of measuring instruments. For example, a Celsius thermometer is calibrated to register zero degrees when immersed in water that contains ice and 100 degrees when immersed in boiling water. A thermometer relies on the theory that mercury expands (approximately) linearly within this range of temperatures. Related theory also tells us how to recalibrate the thermometer if the measurements are made in Denver or Mexico City rather than at sea level. In a sense, model economies, like thermometers, are measuring devices. Generally, some economic questions have known answers, and the model should give an approximately correct answer to them if we are to have any confidence in the answer given to the question with unknown answer. Thus, data are used to calibrate the model economy so that it mimics the world as closely as possible along a limited, but clearly specified, number of dimensions.

Note that calibration is not an attempt at assessing the size of something: it is not estimation. *Estimation* is the determination of the approximate quantity of something. To estimate a parameter, for example, a researcher looks for a situation in which the signal-to-noise ratio is high. Using the existing data and some theory, the researcher then constructs a probability model. An estimator is developed that is robust, relative to the parameter that is to be estimated, to the questionable features of the maintained hypothesis. As a second example of estimation, a computational experiment itself is a type of estimation, in the sense that the quantitative answer to a posed question is an estimate. For example, the quantitative theory of a computational experiment can be used to measure the welfare implications of alternative tax policies.

It is important to emphasize that the parameter values selected are not the ones that provide the best fit in some statistical sense. In some cases, the presence of a particular discrepancy between the data and the model economy is a test of the theory being used. In these cases, absence of that discrepancy is grounds to reject the use of the theory to address the question.

One such example is the use of standard theory to answer the question of how volatile the postwar U.S. economy would have been if technology shocks had been the only contributor to business cycle fluctuations. To answer this question, a model economy with only technology shocks was needed. Using the standard neoclassical production function, standard preferences to describe people's willingness to substitute intra- and intertemporally between consumption and leisure, and an estimate of the technology shock variance, we found that the model economy displays business cycle fluctuations 70 percent as large as did the U.S. economy (Kydland and Prescott, 1991b). This number is our answer to the posed question.

Some have questioned our finding, pointing out that on one key dimension real business cycle models and the world differ dramatically: the correlation between hours worked and average labor productivity is near one in the model economy and approximately zero in U.S. postwar observations (McCallum, 1989). The detractors of the use of standard theory to study business cycles are correct in arguing that the magnitude of this correlation in the world provides a test of the theory. They are incorrect in arguing that passing this test requires the value of this correlation in the model and in the real world to be approximately equal. An implication of the theory is that this correlation is a function of the importance of technology shocks relative to other shocks. In particular, the less is the relative importance of technology shocks, the smaller this correlation should be. The reason for this is that the factors other than technology shocks that give rise to variation in the labor input result in productivity being low when hours are high.[5] Given that the estimated contribution of technology shocks to fluctuations is 70 percent, the correlation between hours and labor productivity being near one in the data would have been grounds for dismissing our answer. (For further elaboration on this point, see Kydland and Prescott, 1991b, p. 79; Aiyagari, 1994.)

Run the Experiment

The fifth and final step in conducting a computational experiment is to run the experiment. Quantitative economic theory uses theory and measurement to estimate how big something is. The instrument is a computer program that determines the equilibrium process of the model economy and uses this equilibrium process to generate equilibrium realizations of the model economy. The computational experiment, then, is the act of using this instrument. These equilibrium realizations of the model economy can be compared with the behavior of the actual economy in some period as follows.

If the model economy has no aggregate uncertainty, then it is simply a matter of comparing the equilibrium path of the model economy with the path of the actual economy. If the model economy has aggregate uncertainty, first a set of statistics that summarize relevant aspects of the behavior of the actual economy is selected. Then the computational experiment is used to generate many independent realizations of the equilibrium process for the model economy. In this way, the sampling distribution of this set of statistics can be determined to any degree of accuracy for the model economy and compared with the values of the set of statistics for the actual economy. In comparing the sampling distribution of a statistic for the model economy to the value of that statistic for the actual data, it is crucial that the same statistic be computed for the model and the real world. If, for

[5] Christiano and Eichenbaum (1992a) have established formally this possibility in the case where the other shock is variations in public consumption, but the result holds for any shock that is approximately orthogonal to the technology shocks.

example, the statistic for the real world is for a 50-year period, then the statistic for the model economy must also be for a 50-year period.

The Computational Experiment in Business Cycle Research

Business cycles, that is, the recurrent fluctuations of output and employment about trend, puzzled economists for a long time. Understanding business cycles required the development of methods that made possible the use of computational experiments to answer questions concerned with the behavior of dynamic economies with uncertainty. Prior to the development of these methods, business cycle fluctuations were viewed as deviations from theory, and very little progress was made in understanding them. Subsequent to the development of those methods, computational experiments have been extensively used in business cycle research. The results of these experiments forced economists to revise their old views, and business cycle fluctuations are now seen as being, in fact, predicted by standard theory. For these reasons, we choose business cycle theory to illustrate the use of computational experiments in economic research.

Posing Questions about the Business Cycle

In the 1970s, a common assumption behind research on the business cycle was that one set of factors, most likely monetary shocks, was behind the cyclical component and that an entirely different set of factors, mainly the growth of productivity and inputs summarized by the neoclassical growth model, accounted for the movement of the long-run growth component.

But there was also an earlier view, tracing as far back as work by Wicksell (1907), that suggested that fluctuation in technological growth could produce broad economic fluctuations. In the 1960s and 1970s, this source was dismissed by many as being unlikely to play much of a role in the aggregate. Most researchers believed that considerable variation could exist in productivity at the industry level, but they believed that industry-level shocks would average out in the aggregate. During the 1980s, however, technology shocks gained renewed interest as a major source of fluctuations, supported largely by computational experiments and quantitative economic theory. As a consequence, business cycle theory treats growth and cycles as being integrated, not as a sum of two components driven by different factors.[6]

[6] An operational way of defining the trend empirically is described in Hodrick and Prescott (1980), who used standard curve-fitting techniques to define a growth component as being the curve that best fits a time series in a least-square sense, subject to a penalty on the sum of the second differences squared. The larger this penalty parameter, the smoother the fitted curve. For quarterly series, they found that a penalty parameter of 1600 made the fitted curve mimic well the one that business cycle analysts would draw. Given the finding that business cycle fluctuations are quantitatively just what neoclassical growth theory predicts, the resulting deviations from trend are nothing more than well-defined statistics. We emphasize that given the way the theory has developed, these statistics measure nothing. Business cycle theory treats growth and cycles as being

Thus, the fundamental question here is the extent to which neoclassical growth theory can account for business cycle fluctuations, as well as long-term growth trends. A particular question addressed was, "How much would the U.S. postwar economy have fluctuated if technology shocks had been the only source of fluctuations?" Computational experiments are well suited to tackle this question.

The Theory Used in Model Selection

The basic theory used in the modern study of business cycles is the neoclassical growth model. The basic version of this model can best be understood as based on five relationships.

The first relationship is an aggregate production function that sets total output equal to $A_tF(K_t,H_t)$, where F is a constant returns to scale function where the inputs are capital and labor, and A_t is the technology level that grows at random rates. In the simplest case, aggregate output is divided between consumption C and investment I. Under the assumption that factors are paid their marginal product, we obtain the identity that GNP and income are equal: $C + I = wH + rK$, where w and r are factor rental prices.

The second relationship in the model economy describes the evolution of the capital stock. Each time period, the existing capital stock depreciates at a constant rate δ, but is replenished by new investment I_t. Thus $K_{t+1} = (1-\delta)K_t + I_t$.

The third relationship describes the evolution of the technology parameter A_t. Given that a structure that displays persistence is needed, a typical form would be $A_{t+1} = \rho A_t + \epsilon_{t+1}$, where ρ is large but less than one, and the shocks ϵ_{t+1} are identically and independently distributed. In other words, the technology level for any given period depends on the technology level in the previous period, plus a random disturbance. The technology described by these relations specifies people's ability to substitute.

The fourth relationship needed for a fully specified economy is a specification of people's willingness to substitute between consumption and leisure, both intertemporally and intratemporally. For this purpose, our model economy has a stand-in household with utility function that depends on consumption and leisure.[7] For simplicity, one can assume that the households own the capital stock directly and rent it to the firms.

---

integrated, not as a sum of two components driven by different factors. For that reason, talking about the resulting statistics as imposing spurious cycles makes no sense. The Hodrick-Prescott filter is simply a statistical decomposition that summarizes in a reasonable way what happens at business cycle frequencies. The representation has proven useful in presenting the findings and judging the reliability of the answer, as well as a way of demonstrating remaining puzzles or anomalies relative to theory.

[7] More explicitly, the function can take the general form

$$E \sum_{t=0}^{\infty} \beta^t U(C_t, 1-H_t),$$

where we normalize so that market and nonmarket productive time add to one. For a complete specification of the model, values of the parameters β, δ and ρ are needed, as well as the explicit utility function U and the production function F.

The final required element is an equilibrium concept. The one used is the competitive equilibrium, which equates marginal rates of substitution and transformation to price ratios. This involves equilibrium decision functions for consumption, investment and labor input as functions of the capital stock and productivity level during that time period: $C(K_t,A_t)$, $I(K_t,A_t)$ and $H(K_t,A_t)$, respectively. In other words, using dynamic programming methods, the decisions can be computed as functions of the list of state variables that provide sufficient information about the position of the economy.

Through this theory, business cycle theorists make contact with other fields of economics. Macroeconomics is no longer largely separate from the rest of economics. The utility and production functions used by business cycle theorists are similar to those used by public finance researchers (for example, Auerbach and Kotlikoff, 1987). The introduction of household production illustrates the close connection with the work of labor economists (for example, Benhabib, Rogerson and Wright, 1991; Greenwood and Hercowitz, 1991). When these models are expanded to consider international trade explicitly, they draw upon work in that field (Backus, Kehoe and Kydland, 1994).

The choice of a model, as already noted, must be governed both by the question at hand and by what is computable.

As an example of altering the model to suit the posed question, consider a contrast between two otherwise similar models. Benhabib, Rogerson and Wright (1991) and Greenwood and Hercowitz (1991) both consider household production in addition to market production, but the two studies are motivated by somewhat different questions. Both use capital and labor as inputs in nonmarket production. Benhabib and his coauthors divide the time allocation of households into three uses: market and nonmarket production time and leisure time. The model is designed to capture the household decision to combine its labor with machines, such as stoves and washing machines, to produce household consumption services. The authors argue that houses do not need to be combined with labor, at least not to the same extent that household machines do, so they exclude housing capital from their concept of household capital. Greenwood and Hercowitz, on the other hand, distinguish only between market and nonmarket time and include the stock of housing, along with consumer durables, in their concept of household capital.[8] This example illustrates the important point that even the definition of particular variables in relation to the model economy may depend on the question.

If a model environment is not computable, then it cannot be used for a computational experiment. This restriction can be a severe one, and the development of appropriate computable methods must therefore be given high priority. Fortunately, there has been considerable progress in this area over the last 30 or 40 years.

[8] To be consistent, they then subtract gross housing product (the measure of the service flow from the economy's housing stock) from GNP and add it to the consumer durables component of personal consumption expenditures.

In cases involving uncertain intertemporal behavior, the development of statistical decision theory has provided a consistent way for people to make decisions under uncertainty. Another significant development is the Arrow-Debreu general equilibrium theory, which extends equilibrium theory to uncertain environments.[9] More recently, Ríos-Rull (1995) offers an overview of the expansion in computable general equilibrium models that incorporate heterogeneity in the household sector—a category that has expanded dramatically over the last few years.

Calibration

Often, calibration involves the simple task of computing a few averages. For example, if the standard Cobb-Douglas production function is used—that is, if we let $F(K,H) = K^{1-\theta}H^{\theta}$—then a numerical value for the parameter θ can be obtained by computing the average labor share of total output over a period of years. Several other growth relations map more or less directly into parameter values for typical models within the neoclassical growth framework, at least if the functional forms have been chosen with calibration in mind. Most growth relations have not changed much, on average, from one cycle to the next for several decades. As a consequence, computational experiments replicate the key long-term or growth relations among model aggregates.

Exceptions do exist, where long-term relationships are not stable. For example, the inventory stock as a fraction of GNP has declined steadily. Durables expenditures as a fraction of total output have risen. For some purposes these changes can be ignored, since that feature does not significantly affect the answer to the question posed. At other times, depending on the associated pattern in the corresponding relative price, such information enables the researcher to obtain a quite precise estimate of some elasticity of substitution, which can then be built into the model.

A good example of this sort of issue is the fact that hours of work per household are about the same now as four decades ago, in spite of a large rise in the real wage rate over the same period. This fact indicates that the elasticity of substitution between consumption and nonmarket time is near one. Still, many business cycle models abstract from the long-run productivity growth that is required to imply this sort of wage growth, because the answer to the questions addressed in those studies would have been essentially the same, as shown by Hansen (1986).[10]

To calibrate a utility function for the household sector of the economy, it is common to rely on averages across large numbers of the relevant population in the actual economy. For example, some model environments employ a utility function in consumption and leisure that, like the Cobb-Douglas production function above,

[9] Also important is the development of recursive methods for the study of economic dynamics, because these methods allow economists to use the computational experiment to generate time series disciplined by factual studies (Stokey and Lucas, 1989).

[10] Hansen (1986) compares otherwise identical model economies and permits growth in one version and not in the other. The model without growth needs a slight adjustment in the capital depreciation rate to be calibrated to the investment share of output and the observed capital/output ratio. With this adjustment, both models estimate the same role of technological shocks (more precisely, Solow residuals) for cyclical fluctuations.

has a share parameter. In this case, the weight that should be placed on consumption turns out to be approximately equal to the average fraction of time spent in market activity. This fraction, in principle, can be obtained from panel data for large samples of individuals. Ghez and Becker (1975) offer a careful measurement study—making reasonable and thoughtful judgments about factors like age limits of the population sample and definition of total time allocated to market and non-market activities, including treatment of sleep and personal care.

In calibration, we sometimes make the model economy inconsistent with the data on one dimension so that it will be consistent on another. For example, İmrohoroğlu (1992) explores the welfare consequences of alternative monetary arrangements in worlds where agents are liquidity constrained, while Cooley and Hansen (1989) explore the welfare consequences in worlds where people use money for transaction purposes. These are two very different environments, each of which abstracts from the main feature of the other. İmrohoroğlu calibrates her model economy to yield a stock of money held per household that is in line with U.S. observations. In her model, however, people hold money because they do not have access to an insurance technology to insure against randomness in the market value of their time. Equivalently, if they do have access to such an insurance technology, they find it so costly that, in equilibrium, they do not employ it. This is the only reason, in her model, for people to hold money; if she had calibrated the model to the amount of variation in individual income found in panel data, the model would have implied that average household holdings of liquid assets were about half of those actually held.

Of course, households have other reasons for holding liquid assets that earn much less than the average return on physical capital. For instance, such assets can be used as a down payment on a house at some future date, as a substitute for insurance against sickness and accidents, or for transaction purposes, as in the Cooley and Hansen (1989) environment. These and other factors are abstracted from in the world of İmrohoroğlu (1992), which led her to introduce greater variation in the market value of households' time so as to make per capita holdings of money in the model match actual holdings. This calibration is reasonable, given the question she addresses. Her implicit assumption is that it is unimportant which liquidity factor gives rise to these holdings. Subsequent research will either support this working hypothesis or disprove it and, in the process, lead to better model economies for evaluating monetary and credit policy arrangements. This sequence is how economic science progresses.

Running Experiments

With explicit functional forms for the production and utility functions, with values assigned to the parameters, and with a probability distribution for the shocks, a researcher can use the model economy to perform computational experiments. The first step is to compute the equilibrium decision rules, which are functions of the state of the economy. The next step is to generate equilibrium realizations of the economy. The computer begins each period with a given level of the state variables, for example, the capital stock and the technology level. The values of the state variables along with the equilibrium decision and pricing functions determine the equilibrium realization for that period. Equilibrium investment and the new technology shocks determine next period's state. In the next and subsequent periods, this procedure is repeated until time series of the desired length are obtained. The resulting model time series can then be summarized by a suitable set of statistics.

In Kydland and Prescott (1982), we built a model economy where all fluctuations could be traced back to technology shocks. We began by extending the neoclassical growth model to include leisure as an argument of the stand-in household's utility function. Given that more than half of business cycle fluctuations are accounted for by variations in the labor input, introducing this element was crucial. We then calibrated the deterministic version of the model so that its consumption-investment shares, factor income shares, capital/output ratios, leisure/market-time shares and depreciation shares matched the average values for the U.S. economy in the postwar period. We abstracted from public finance considerations and consolidated the public and private sectors. We introduced Frisch's (1933b) assumption of time to build new productive capital. The construction period considered was four quarters, with new capital becoming productive only upon completion, but with resources being used up throughout the construction period. Given the high volatility of inventory investment, inventory stocks were included as a factor of production. In our computational experiments, using technology shock variance estimated from production function residuals (Prescott, 1986), we found that the model economy's output variance was 55 percent as large as the corresponding variance for the U.S. economy in the postwar period.

Probably the most questionable assumption of this theory is that of homogeneous workers, with the additional implication that all variation in hours occurs in the form of changes in hours per worker. According to aggregate data for the U.S. economy, only about one-third of the quarterly fluctuations in market hours are of this form, while the remaining two-thirds arise from changes in the number of workers (Kydland and Prescott, 1990, Table 1).

This observation led Hansen (1985) to introduce the Rogerson (1988) labor indivisibility construct into a business cycle model. In the Hansen world, all fluctuations in hours are in the form of employment variation. To deal with the apparent nonconvexity arising from the assumption of indivisible labor, Hansen makes the problem convex by assuming that the commodity points are contracts in which every agent is paid the same amount whether that agent works or not, and that a lottery randomly chooses who actually works in every period. He finds that with this labor indivisibility, his model economy fluctuates as much as did the U.S. economy. Our view is that with the extreme assumption of fluctuations only in employment, Hansen overestimates the amount of aggregate fluctuations accounted for by Solow residuals in the same way that our equally extreme assumption of fluctuations solely in hours per worker led to an underestimation.

In Kydland and Prescott (1991b), the major improvement on the 1982 version of the model economy is that variation is permitted in both the number of workers and the number of hours per worker. The number of hours in which a plant is operated in any given period is endogenous.

Because the cost of moving workers in and out of the labor force is not included, a property of the equilibrium is that all of the hours variation is in the form of employment change and none in hours per worker. In that respect, the Kydland and Prescott (1991b) model is identical to Hansen's (1985) model. Using the economy with no moving costs, technology shocks are estimated to account for about 90 percent of the aggregate output variance. For the economy with moving costs, we calibrated it so that relative variations in hours per worker and the number of workers matched U.S. data. The estimate of the fraction of the cycle accounted for by technology shocks is then reduced to 70 percent.

These estimates of the contribution of technology shocks to aggregate fluctuations have been found to be robust in several modifications of the model economy. For instance, Greenwood, Hercowitz and Huffman (1988) permit the utilization rate of capital to vary and affect its depreciation rate and assume all technology change is embodied in new capital. Danthine and Donaldson (1990) introduce an efficiency-wage construct, while Cho and Cooley (1995) permit nominal-wage contracting. Ríos-Rull (1994) uses a model calibrated to life cycle earnings and consumption patterns. Gomme and Greenwood (1995) incorporate heterogeneous agents with recursive preferences and equilibrium risk allocations. In none of these cases is the estimate of the contribution of technology shocks to aggregate fluctuations significantly altered.

The computational experiment is also being used to derive the quantitative implications of monetary shocks for business cycle fluctuations if money is used for transaction purposes only. In these experiments, money may be held either because it is required in advance of purchasing cash goods (Lucas and Stokey, 1987; Cooley and Hansen 1989, 1992) or because real cash balances economize on time (Kydland, 1989). Models of this type have been used to evaluate monetary policy.

At this stage, we are less confident in these model economies than those designed to evaluate the contribution of technology shocks. There are three related reasons. The first is that, unlike actual economies, these model economies fail to display the sluggishness of the inflation rate's response to changes in the growth rate of money (Christiano and Eichenbaum, 1992b). The second is that people seem to hold far larger monetary assets than are needed for transaction purposes—in the postwar period, for example, U.S. households' holdings of M2 have exceeded half of annual GNP—which implies that the transaction rationale for holding money is not well understood. The third reason is that the evaluation of monetary policy appears to be sensitive to the reason people hold these liquid assets. İmrohoroğlu (1992) has constructed a model economy in which people vary their holdings of liquid assets as their income varies to smooth their consumption.[11] She finds that if a transaction-cost model is calibrated to data generated by her model econ-

[11] İmrohoroğlu and Prescott (1991) introduce a banking technology to intermediate government debt.

omy and the calibrated economy is used to estimate the cost of inflation, this estimate would grossly underestimate the true cost of inflation for her model world. This result is surprising and bothersome. Typically, how some feature is introduced is unimportant as long as the aggregate substitution elasticities and quantities match. We currently do not have the tools for computing equilibria of models with both the features of the neoclassical growth model and the idiosyncratic shocks that result in people holding money for precautionary reasons. One may say that stronger theory is needed when it comes to evaluating the contribution of monetary policy shocks to business cycle fluctuations.

Summary

With the general equilibrium approach, empirical knowledge is organized around preferences and technologies. Given the question and given existing economic theory and measurement, a researcher creates a model economy. This researcher then determines a quantitative answer to the posed question for the model economy. If the theory is strong and the measurements good, we have confidence that the answer for the model economy will be essentially the same as for the actual economy.

Of course, sometimes measurement is not very good, and a series of computational experiments reveals that different plausible values of some parameter give very different answers to the posed question. If so, this parameter—which measures some aspect of people's willingness and ability to substitute—must be more accurately measured before theory can provide an answer in which we have confidence. Or sometimes the theory relative to the question is weak or nonexistent, and the answer depends upon which of the currently competing theories is used to construct the model economy. If so, these competing theories must be subjected to further tests before there is a good basis for choosing among them. At still other times, the computational tools needed to derive the implications of the theory do not exist, so better computational methods or more powerful computers are needed.

Earlier in this article, we quoted the Lucas (1980) definition of "theory" as being an explicit set of instructions for building an imitation economy to address certain questions and not a collection of assertions about the behavior of the actual economy. Consequently, statistical hypothesis testing, which is designed to test assertions about actual systems, is not an appropriate tool for testing economic theory.

One way to test a theory is to determine whether model economies constructed according to the instructions of that theory mimic certain aspects of reality. Perhaps the ultimate test of a theory is whether its predictions are confirmed—that is, did the actual economy behave as predicted by the model economy, given the policy rule selected? If a theory passes these tests, then it is tested further, often by conducting a computational experiment that includes some feature of reality not previously included in the computational experiments. More often than not, introducing this feature does not change the answers, and currently established theory becomes stronger. Occasionally, however, the new feature turns out to be important, and established theory must then be expanded and improved. In this way, economic science progresses.

Given the infeasibility of controlled experiments with national economies, the computational experiment is the tool of quantitative economic theory, whether the primary concern be with theory use or with theory development and testing.

■ *We thank Graham Candler, Javier Díaz-Giménez, Tryphon Kollintzas, Jim Schmitz and Nancy Stokey for helpful discussions and the National Science Foundation for financial support. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Banks of Cleveland or Minneapolis, nor of the Federal Reserve System.*

References

Aiyagari, S. Rao, "On the Contribution of Technology Shocks to Business Cycles," *Quarterly Review, Federal Reserve Bank of Minneapolis,* Winter 1994, *18,* 22–34.

Auerbach, Alan J., and Laurence J. Kotlikoff, *Dynamic Fiscal Policy.* Cambridge, U.K.: Cambridge University Press, 1987.

Backus, David K., Patrick J. Kehoe, and Finn E. Kydland, "Dynamics of the Trade Balance and the Terms of Trade: The J-curve?," *American Economic Review,* March 1994, *84,* 84–103.

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

Brown, Drussilla K., Alan V. Deardorff, and Robert M. Stern, "Estimates of a North American Free Trade Agreement," manuscript, Federal Reserve Bank of Minneapolis, 1994.

Chang, Ly-June, "Business Cycles with Distorting Taxes and Disaggregate Capital Markets," *Journal of Economic Dynamics and Control,* July/September 1995, *19,* 985–1010.

Cho, Jong Ok, and Thomas F. Cooley, "The Business Cycle with Nominal Contracts," *Economic Theory,* June 1995, *6,* 13–34.

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

Christiano, Lawrence J., and Martin Eichenbaum, "Liquidity Effects, Monetary Policy, and the Business Cycle." Working Paper No. 4129, National Bureau of Economic Research, 1992b.

Cooley, Thomas F., and Gary D. Hansen, "The Inflation Tax in a Real Business Cycle Model," *American Economic Review,* September 1989, *79,* 733–48.

Cooley, Thomas F., and Gary D. Hansen, "Tax Distortions in a Neoclassical Monetary Economy," *Journal of Economic Theory,* December 1992, *58,* 290–316.

Danthine, Jean-Pierre, and John B. Donaldson, "Efficiency Wages and the Business Cycle Puzzle," *European Economic Review,* November 1990, *34,* 1275–1301.

Finn, Mary G., "Variance Properties of Solow's Productivity Residual and their Cyclical Implications," *Journal of Economic Dynamics and Control,* July/September 1995, *19,* 1249–82.

Frisch, Ragnar, "Editorial," *Econometrica,* January 1933a, *1,* 1–5.

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

Ghez, Gilbert R., and Gary S. Becker, *The Allocation of Time and Goods over the Life Cycle.* New York: Columbia University Press, 1975.

Gomme, Paul, and Jeremy Greenwood, "On the Cyclical Allocation of Risk," *Journal of Economic Dynamics and Control,* January/February 1995, *19,* 91–124.

Gravelle, Jane G., and Laurence J. Kotlikoff, "Corporate Taxation and the Efficiency Gains of the 1986 Tax Reform Act," *Economic Theory,* June 1995, *6,* 51–82.

Greenwood, Jeremy, and Zvi Hercowitz, "The Allocation of Capital and Time Over the Business Cycle," *Journal of Political Economy,* December 1991, *99,* 1188–214.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman, "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review,* June 1988, *78,* 402–18.

Gregory, Allan W., and Gregory W. Smith, "Statistical Aspects of Calibration in Macroeconomics." In Maddala, G. S., C. R. Rao, and H. D. Vinod, eds., *Handbook of Statistics.* Vol. 2, New York: Elsevier Science, 1993, pp. 703–19.

Hansen, Gary D., "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics,* November 1985, *16,* 309–27.

Hansen, Gary D., "Growth and Fluctuations." In "Three Essays on Labor Indivisibility and the Business Cycle," Ph.D. dissertation, University of Minnesota, 1986.

Harris, Richard, "Applied General Equilibrium Analysis of Small Open Economies with Scale Economies and Imperfect Competition," *American Economic Review,* December 1984, *18,* 1016–32.

Hodrick, Robert J., and Edward C. Prescott, "Post-War U.S. Business Cycles: An Empirical Investigation." Discussion Paper No. 451, Carnegie-Mellon University, 1980.

Hopenhayn, Hugo, and Richard Rogerson, "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy,* October 1993, *101,* 915–38.

Hornstein, Andreas, "Monopolistic Competition, Increasing Returns to Scale and the Importance of Productivity Change," *Journal of Monetary Economics,* June 1993, *31,* 299–316.

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

İmrohoroğlu, Ayşe, and Edward C. Prescott, "Seigniorage as a Tax: A Quantitative Evaluation," *Journal of Money, Credit and Banking,* August 1991, *23,* 462–75.

Kuhn, Thomas S., *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press, 1962.

Kydland, Finn E., "The Role of Money in a Business Cycle Model." Discussion Paper No. 23, Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, 1989.

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

Kydland, Finn E., and Edward C. Prescott, "Business Cycles: Real Facts and a Monetary Myth," *Quarterly Review, Federal Reserve Bank of Minneapolis,* Spring 1990, *14,* 3–18.

Kydland, Finn E., and Edward C. Prescott, "The Econometrics of the General Equilibrium Approach to Business Cycles," *Scandinavian Journal of Economics,* 1991a, *93*:2, 161–78.

Kydland, Finn E., and Edward C. Prescott, "Hours and Employment Variation in Business Cycle Theory," *Economic Theory,* January 1991b, *1,* 63–81.

Lucas, Robert E., Jr., "Methods and Problems in Business Cycle Theory," *Journal of Money, Credit and Banking,* November 1980, *12,* 696–715.

Lucas, Robert E., Jr., and Nancy L. Stokey, "Money and Interest in a Cash-in-Advance Economy," *Econometrica,* May 1987, *55,* 491–513.

Marschak, Jacob, "Introduction to Econometrics," hectographed lecture notes, University of Buffalo, 1948.

McCallum, Bennett T., "Real Business Cycle Models." In Barro, Robert J., ed., *Modern Business Cycle Theory.* Cambridge, Mass.: Harvard University Press, 1989, pp. 16–50.

Prescott, Edward C., "Theory Ahead of Business Cycle Measurement," *Quarterly Review, Federal Reserve Bank of Minneapolis,* Fall 1986, *10,* 9–22; also in Brunner, Karl, and Allan H. Meltzer, eds., *Real Business Cycles, Real Exchange Rates, and Actual Policies, Carnegie-Rochester Conference Series on Public Policy.* Vol. 25, Amsterdam: North-Holland, 1986, pp. 11–44.

Ríos-Rull, José-Víctor, "On the Quantitative Importance of Market Completeness," *Journal of Monetary Economics,* December 1994, *34,* 462–96.

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.

Rogerson, Richard, "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics,* January 1988, *21,* 3–16.

Schneider, Stephen H., "Climate Modeling," *Scientific American,* May 1987, *256,* 72–80.

Shoven, John B., and J. Whalley, "A General Equilibrium Calculation of the Differential Taxation of Income from Capital in the U.S.," *Journal of Public Economics,* November 1972, *1,* 281–321.

Solow, Robert M., *Growth Theory: An Exposition.* Oxford: Clarendon Press, 1970.

Stokey, Nancy L., and Robert E. Lucas, Jr. (with Edward C. Prescott), *Recursive Methods in Economic Dynamics.* Cambridge, Mass.: Harvard University Press, 1989.

Wicksell, Knut, "Krisernas gåta," *Statsøkonomisk Tidsskrift,* 1907, *21,* 255–70.