I have been given the daunting task of discussing how the debate among the various schools of business cycle theorists might be resolved. People obviously differ in how they think the macroeconomy works. Will there ever be a winner? I am optimistic enough to think so, although I view the last two decades as making only modest progress in this direction. One problem is that there is too little testing of alternative models. There has been no systematic attempt to find the model that best approximates the macroeconomy. As disturbing, however, is the fact that macroeconomic research appears to be moving away from its traditional empirical emphasis. I will elaborate on both points in this chapter.

From Tinbergen's (1939) model building in the late 1930s through the 1960s, the dominant methodology of macroeconomics was what I will call the "Cowles Commission" approach. Structural econometric models were specified, estimated, and then analyzed and tested in various ways. One of the major macroeconometric efforts of the 1960s, building on the earlier work of Klein (1950) and Klein and Goldberger (1955), was the Brookings model (Duesenberry et al., 1965, 1969). This model was a joint effort of many individuals, peaking at nearly 400 equations. Although much was
learned from this exercise, the model never achieved the success initially expected and was laid to rest around 1972.

Two important events in the 1970s contributed to the decline in popularity of the Cowles Commission approach. The first was the commercialization of macroeconometric models. This changed the focus of research on the models. Basic research gave way to the day-to-day needs of updating the models, of subjectively adjusting the forecasts to make them "reasonable," and of meeting the special needs of clients. The second event was Lucas's (1976) critique, which argued that the models are not likely to be useful for policy purposes.

The Lucas critique led to a line of research that has cumulated in the real business cycle (RBC) theories. This in turn has generated a counter response in the form of new-Keynesian economics. I will argue that neither the RBC approach nor new-Keynesian economics is in the spirit of the Cowles Commission approach and that this is a step backward. The Cowles Commission approach is discussed in section 1; the RBC approach, in section 2; and new-Keynesian economics, in section 3. Suggestions for the future are then presented.

1 The Cowles Commission Approach

1.1 Specification

Some of the early macroeconometric models were linear, but this soon gave way to the specification of nonlinear models. Consequently, only the nonlinear case will be considered here. The model will be written as

$$f_i(y_t, x_t, \alpha_i) = u_{it} \quad (i = 1, \ldots, n) \quad (t = 1, \ldots, T),$$

where $y_t$ is an $n$-dimensional vector of endogenous variables, $x_t$ is a vector of predetermined variables (including lagged endogenous variables), $\alpha_i$ is a vector of unknown coefficients, and $u_{it}$ is the error term for equation $i$ for period $t$. For equations that are identities, $u_{it}$ is identically zero for all $t$.

Specification consists of choosing (1) the variables that appear in each equation with nonzero coefficients, (2) the functional form of each equation, and (3) the probability structure for $u_{it}$. (In modern times one has to make sufficient stationarity assumptions about the variables to make the time series econometricians happy. The assumption, either explicit or implicit, of most macroeconometric model building work is that the variables are trend stationary.) Economic theory is used to guide the choice of variables. In most cases there is an obvious left-side variable for the
equation, where the normalization used is to set the coefficient of this variable equal to one. This is the variable considered to be "explained" by the equation.

It will be useful to consider an example of how theory is used to specify an equation. Consider the following maximization problem for a representative household. Maximize

$$E_0 U(C_1, \ldots, C_T, L_1, \ldots, L_T),$$

subject to

$$S_t = W_t (H - L_t) + r_t A_{t-1} - P_t C_t$$
$$A_t = A_{t-1} + S_t$$
$$A_T = \bar{A},$$

where $C$ is consumption, $L$ is leisure, $S$ is saving, $W$ is the wage rate, $H$ is the total number of hours in the period, $r$ is the one-period interest rate, $A$ is the level of assets, $P$ is the price level, $\bar{A}$ is the terminal value of assets, and $t = 1, \ldots, T$. $E_0$ is the expectations operator conditional on information available through time 0. Given $A_0$ and the conditional distributions of the future values of $W$, $P$, and $r_t$, $t = 1, \ldots, T$, it is possible in principle to solve for the optimal values of $C$ and $L$ for period 1, denoted $C^*_1$ and $L^*_1$. In general, however, this problem is not analytically tractable. In other words, it is not generally possible to find analytic expressions for $C^*_1$ and $L^*_1$.

The approach that I am calling the Cowles Commission approach can be thought of as specifying and estimating approximations of the decision equations. In the context of the present example, this approach is as follows. First, the random variables, $W_t$, $P_t$, and $r_t$, $t = 1, \ldots, T$, are replaced by their expected values, $E_0 W_t$, $E_0 P_t$, and $E_0 r_t$, $t = 1, \ldots, T$. Given this replacement, one can write the expressions for $C^*_1$ and $L^*_1$ as

$$C^*_t = g_1(A_0, \bar{A}, E_0 W_1, \ldots, E_0 W_T, E_0 P_1, \ldots, E_0 P_T, E_0 r_1, \ldots, E_0 r_T, \beta)$$
$$L^*_t = g_2(A_0, \bar{A}, E_0 W_1, \ldots, E_0 W_T, E_0 P_1, \ldots, E_0 P_T, E_0 r_1, \ldots, E_0 r_T, \beta),$$

where $\beta$ is the vector of parameters of the utility function. Equations (4) and (5) simply state that the optimal values for the first period are a function of (1) the initial and terminal values of assets; (2) the expected future values of the wage rate, the price level, and the interest rate; and (3) the parameters of the utility function. The functional forms of equations (4) and (5) are not in general known. The aim of the empirical work is to try to estimate equations that are approximations of equations (4) and (5). Experimentation consists in trying different functional forms and in trying different assumptions about how expectations are formed. Because of the
large number of expected values in equations (4) and (5), the expectational assumptions usually restrict the number of free parameters to be estimated. For example, the parameters for $E_0 W_1, \ldots, E_0 W_T$ might be assumed to lie on a low-order polynomial or to be geometrically declining. The error terms are usually assumed to be additive, as specified in equation (1), and they can be interpreted as approximation errors.

It is often the case when equations like (4) and (5) are estimated that lagged dependent variables are used as explanatory variables. Since $C_0$ and $L_0$ do not appear in equations (4) and (5), how can one justify the use of lagged dependent variables? A common procedure is to assume that $C^*_t$ in equation (4) and $L^*_t$ in equation (5) are long-run “desired” values. It is then assumed that because of adjustment costs, there is only a partial adjustment of actual to desired values. The usual adjustment equation for consumption would be

$$C_t - C'_t = \lambda (C^*_t - C'_t) \quad 0 < \lambda < 1,$$

which adds $C_0$ to the estimated equation. This procedure is ad hoc in the sense that the adjustment equation is not explicitly derived from utility maximization. One can, however, assume that there are utility costs to large changes in consumption and leisure and thus put terms like $(C_1 - C_0)^2$, $(C_2 - C_1)^2$, $(L_1 - L_0)^2$, $(L_2 - L_1)^2$, $\ldots$ in the utility function, equation (2). This would add the variables $C_0$ and $L_0$ to the right-hand side of equations (4) and (5), which would justify the use of lagged dependent variables in the empirical approximating equations for (4) and (5).

This setup can handle the assumption of rational expectations in the following sense. Let $E_{t-1} y_{2t+1}$ denote the expected value of $y_{2t+1}$, where the expectation is based on information through period $t-1$, and assume that $E_{t-1} y_{2t+1}$ appears as an explanatory variable in equation (1). (Equation 1 might be the equation explaining consumption, and $y_2$ might be the wage rate.) If expectations are assumed to be rational, equation (1) can be estimated by either a limited information or a full information technique. In the limited information case $E_{t-1} y_{2t+1}$ is replaced by $y_{2t+1}$, and the equation is estimated by Hansen’s (1982) generalized method of moments (GMM) procedure. In the full information case the entire model is estimated at the same time by full information maximum likelihood, where the restriction is imposed that the expectations of future values of variables are equal to the model’s predictions of the future values. Again, the parameters of the expected future values might be restricted to lessen the number of free parameters to be estimated.

The specification just outlined does not allow the estimation of “deep structural parameters,” such as the parameters of utility functions, even
under the assumption of rational expectations. Only approximations of the
decision equations are being estimated. The specification is thus subject
to the Lucas (1976) critique. More will be said about this later. The
specification also uses the certainty equivalence procedure, which is strictly
valid only in the linear-quadratic setup.

1.2 Estimation

A typical macroeconometric model is dynamic, nonlinear, simultaneous,
and has error terms that may be correlated across equations and with their
lagged values. A number of techniques have been developed for the esti-
mation of such models. Techniques that do not consider the correlation of
the error terms across equations (limited information techniques) include
limited information maximum likelihood and two-stage least squares.
Techniques that do account for this correlation (full information tech-
niques) include full information maximum likelihood and three-stage least
squares. It is straightforward to modify these techniques to handle the case
in which the error terms follow autoregressive processes. Also, as already
noted, the techniques can be modified to handle the assumption of rational
expectations.

Computational advances, both in hardware and software, have made the
application of these techniques fairly routine. Even full information maxi-
mum likelihood when expectations are assumed to be rational appears
computationally feasible for most models.

1.3 Testing

Testing has always played a major role in applied econometrics. When
an equation is estimated, one examines how well it fits the data, if its
coefficient estimates are significant and of the expected sign, if the proper-
ties of the estimated residuals are as expected, and so on. Equations are
discarded or modified if they do not seem to approximate very well the
process that generated the data.

Complete models can also be tested, but here things are more compli-
cated. Given (1) a set of coefficient estimates, (2) values of the exogenous
variables, (3) values of the error terms, and (4) lagged values of the
endogenous variables, a model can be solved for the endogenous variables.
If the solution (simulation) is "static," the actual values of the lagged
endogenous variables are used for each period solved; if the solution is
"dynamic," the values of the lagged endogenous variables are taken to be the predicted values of the endogenous variables from the previous periods. If one set of values of the error terms is used, the simulation is said to be "deterministic." The expected values of the error terms are usually assumed to be zero, and so in most cases the error terms are set to zero for the solution. A "stochastic" simulation is one in which (1) the error terms are drawn from estimated distributions, (2) the model is solved for each set of draws, and (3) the predicted value of each endogenous variable is taken to be the average of the solution values across the sets of draws.

A standard procedure for evaluating how well a model fits the data is to solve the model by performing a dynamic, deterministic simulation and then compare the predicted values of the endogenous variables with the actual values using the root mean squared error (RMSE) criterion. Other criteria include mean absolute error and Theil's inequality coefficient. If two models are being compared and model A has lower RMSEs for most of the variables than model B, this is evidence in favor of model A over model B.

There is always a danger in this business of "data mining," which means specifying and estimating different versions of a model until a good fit has been achieved (say in terms of the RMSE criterion). The danger with this type of searching is that one finds a model that fits well within the estimation period but is in fact a poor approximation of the economy. To guard against this, predictions are many times taken to be outside of the estimation period. If a model is poorly specified, if should not predict well outside the period for which it was estimated, even though it may fit well within the period.5

One problem with the RMSE criterion is that it does not take account of the fact that forecast-error variances vary across time. Forecast-error variances vary across time because of nonlinearities in the model and because of variation in the exogenous variables. Although RMSEs are in some loose sense estimates of the averages of the variances across time, no rigorous statistical interpretation can be placed on them: They are not estimates of any parameters of a model.

A more serious problem with the RMSE criterion as a means of comparing models is that models may be based on different sets of exogenous variables. If one model takes investment as exogenous and a second does not, the first model has an unfair advantage when computing RMSEs.

I have developed a method, based on stochastic simulation, that accounts for the RMSE difficulties (Fair, 1980). The method accounts for the four main sources of uncertainty of a forecast from a model: uncertainty due to (1) the error terms, (2) the coefficient estimates, (3) the
exogenous variables, and (4) the possible misspecification of the model. The forecast-error variance for each variable and each period estimated by the method accounts for all four sources of uncertainty, and so it can be compared across models. The estimated variances from different structural models can be compared, or the estimated variances from one structural model can be compared to those from an autoregressive or vector autoregressive model. If a particular model's estimated variances are in general smaller than estimated variances from other models, this is evidence in favor of the particular model.

A by-product of the method is an estimate of the degree of misspecification of a model for each endogenous variable. Any model is likely to be somewhat misspecified, and the method can estimate the quantitative importance of the misspecification.

The method can handle a variety of assumptions about exogenous-variable uncertainty. One polar assumption is that there is no uncertainty attached to the exogenous variables. This might be true, for example, of some policy variables. The other polar assumption is that the exogenous variables are in some sense as uncertain as the endogenous variables. One can, for example, estimate autoregressive equations for each exogenous variable and add these equations to the model. This would produce a model with no exogenous variables, which could then be tested. An in-between case would include estimating the variance of an exogenous-variable forecast error from actual forecasting errors made by a forecasting service—say the errors made by DRI in forecasting defense spending.

Another method comparing models is to regress the actual value of an endogenous variable on a constant and forecasts of the variable from two or more models. This procedure is explained in Fair and Shiller (1990) and is related to the literature on encompassing tests—see, for example, Davidson and MacKinnon (1981), Hendry and Richard (1982), and Chong and Hendry (1986). Again, one can use autoregressive or vector autoregressive models as comparisons for structural models using these tests.

Testing models in the ways described here seems clearly in the spirit of the Cowles Commission approach. A model to the Cowles Commission was a null hypothesis to be tested.

1.4 A Digression on Monetarism

Laidler (1992) has written an interesting and useful chapter on the history of monetarism. From the perspective of the Cowles Commission approach, I have no complaints about this chapter. The Laidler and Bentley (1983)
model, for example, is a standard macroeconometric model and can be tested in the ways already discussed. In fact, in Laidler’s last sentence he states that he hopes to “reinstate empirical evidence as a factor more important than a priori principles” in the debate about business cycles. I would interpret this as an argument for the Cowles Commission approach.

This is not to say, however, that I would argue in favor of the Laidler and Bentley model. Although in simple tests the model does about the same as Barro’s (1978) model, I doubt that it would hold up well against larger structural models or even against autoregressive and vector autoregressive models. But the main point is that the model can be tested. The debate is simply between which macroeconometric model—a model in the monetarist spirit or some other model—best explains the data.

2 The Real Business Cycle Approach

In discussing the RBC approach, it will be useful to begin with the utility maximization model already considered. The RBC approach to this model would be to specify a particular functional form for the utility function in equation (2). The parameters of this function would then be either estimated or simply chosen (“calibrated”) to be in line with parameters estimated in the literature.

Although there is some parameter estimation in the RBC literature, most of the studies calibrate rather than estimate, in the spirit of the seminal article by Kydland and Prescott (1982). If the parameters are estimated, they are estimated from the first-order conditions. A recent example is Christiano and Eichenbaum (1990), where the parameters of their model are estimated using Hansen’s (1982) GMM procedure. Altug (1989) estimates the parameters of her model using a likelihood procedure. Chow (1991) and Canova, Finn, and Pagan (1991) contain interesting discussions of the estimation of RBC models. There is also a slightly earlier literature in which the parameters of a utility function, as in equation (2), were estimated from the first-order conditions—see, for example, Hall (1978), Hansen and Singleton (1982), and Mankiw, Rotemberg, and Summers (1985).

The RBC approach meets the Lucas critique; deep structural parameters are being estimated (or calibrated). It is hard to overestimate the appeal this has to many people. Anyone who doubts this appeal should read Lucas’s 1985 Jahnsson lectures (Lucas, 1987), which is an elegant argument for dynamic economic theory. The tone of these lectures is an exciting sense of progress in macroeconomics and hope that in the end
there will be essentially no distinction between microeconomics and macroeconomics. There will simply be economic theory applied to different problems.

Once the coefficients are chosen, by whatever means, the overall model can be solved. In the earlier example, one could solve the utility maximization problem for the optimal consumption and leisure paths. The properties of the computed paths of the decision variables are then compared to the properties of the actual paths of the variables. If the computed paths have similar properties to the actual paths (for example, similar variances, covariances, and autocovariances), this is judged to be a positive sign for the model. If the parameters are chosen by calibration, there is usually some searching over parameters to find that set that gives good results in matching the computed paths to the actual paths in terms of the particular criterion used. In this sense the calibrated parameters are also estimated.

Is the RBC 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 procedure. This procedure would compute a prediction error for a given variable for each period and then calculate the RMSE from these prediction errors. This RMSE might then be compared to 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.

If this literature proceeds anyway, it has in my view dropped out of the race. The literature may take a long time to play itself out, but it will eventually reach a dead end unless it comes around to developing models that can compete with other models in explaining the economy observation by observation.

One of the main reasons that individuals proceed anyway is undoubtedly the Lucas critique and the general excitement about deep structural
parameters. Why waste one's time in working with models whose coefficients change over time as policy rules and other things change? The logic of the Lucas critique is certainly correct, but the key question for empirical work is the quantitative importance of this critique. Even the best econometric model is only an approximation of how the economy works. Another potential source of coefficient change is the use of aggregate data. As the age and income distributions of the population change, the coefficients in aggregate equations are likely to change, and this is a source of error in the estimated equations. This problem may be quantitatively much more important than the problem raised by Lucas. Put another way, the representative agent model that is used so much in macroeconomics has serious problems of its own, which may completely swamp the problem of coefficients changing when policy rules change. The RBC literature has focused so much on solving one problem that it is likely in the process to have exacerbated the effects of a number of others.

There are a number of reasons the RBC models probably do not fit the data well. The RBC approach requires that a particular functional form for, say, the utility function be chosen, and errors made in this choice may lead to large prediction errors. Remember that this function represents the average of the utility functions of all the households in the economy, and it is unlikely that one is going to get this quite right. The advantage of estimating approximations of the decision equations, as discussed in section 1, is that it allows more flexibility in estimating functional forms. The data are allowed more play, if you will. Using the approach of estimating approximations of decision equations, one trades off estimating deep structural parameters for less sensitivity to functional-form errors and the like.

When deep structural parameters have been estimated from the first-order conditions, the results have not always been very good even when judged by themselves. The results in Mankiw, Rotemberg, and Summers (1985) for the utility parameters are not supportive of the approach. In a completely different literature—the estimation of production-smoothing equations—Krane and Braun (1989), whose study is based on quite good data, report that their attempts to estimate first-order conditions was unsuccessful. It may be that one is asking too much of the aggregate data to force them into estimating what one thinks are parameters from some postulated function.

Finally, one encouraging feature regarding the Lucas critique is that it can be tested. Assume that for an equation or set of equations the parameters change considerably when a given policy variable changes. Assume also that the policy variable changes frequently. In this case the model is obviously misspecified, and so methods like that discussed in
section 1 should be able to pick up this misspecification if the policy variable has changed frequently. If the policy variable has not changed or changed very little, the model will be misspecified but the misspecification will not be given a chance to be picked up in the data. But otherwise, models that suffer in an important way from the Lucas critique ought to be weeded out by various tests.

3 The New-Keynesian Economics

After reading or rereading a number of new-Keynesian articles for this chapter, I came away feeling uneasy. It's like coming out of a play that many of your friends liked and feeling that you did not really like it, but not knowing quite why. Given my views of how the economy works, many of the results of the new-Keynesian literature seem reasonable, but something seemed missing. One problem is that it is hard to get a big picture. There are many small stories, and it's hard to remember each one. In addition, many of the conclusions do not seem robust to small changes in the models.

On further reflection, however, I do not think this was my main source of uneasiness. The main problem is that this literature is not really empirical in the Cowles Commission sense. This literature has moved macroeconomics away from its econometric base. Consider, for example, the articles in the two volumes of New Keynesian Economics, edited by Mankiw and Romer (1991). By my count, of the 34 chapters in these two volumes, only 8 have anything to do with data. Of these 8, one (Carlton, "The Rigidity of Prices") is more industrial organization than macro and one (Krueger and Summers, "Efficiency Wages and the Interindustry Wage Structure") is more labor than macro. These two studies provide some interesting insights that might be of help to macroeconomists, but they are not really empirical macroeconomics.

It has been pointed out to me that the Mankiw and Romer volumes may be biased against empirical papers because of space constraints imposed by the publisher. Nevertheless, it seems clear that there is very little in the new-Keynesian literature in the nature of structural modeling of the kind outlined in section 1. As in the RBC literature, one does not see, say, predictions of real GNP from some new-Keynesian model compared to predictions of real GNP from an autoregressive equation using a criterion like RMSE. But here one does not see it because no econometric models of real GNP are constructed! So this literature has dropped out of the race not because it is necessarily uninterested in serious tests but because it is uninterested in constructing econometric models.
I should hasten to add that I do not mean by these criticisms that there is no interesting empirical work going on in macroeconomics. For example, the literature on production smoothing, which is largely empirical, has produced some important results and insights. It is simply that literature of this type is not generally classified as new-Keynesian. Even if one wanted to be generous and put some of this empirical work in the new-Keynesian literature, it is surely not the essence of new-Keynesian economics.

One might argue that new-Keynesian economics is just getting started and that the big picture (model) will eventually emerge to rival existing models of the economy. This is probably an excessively generous interpretation, given the focus of this literature on small theoretical models, but unless the literature does move in a more econometric and larger-model direction, it is not likely to have much long-run impact.

4 Looking Ahead

So I see the RBC and new-Keynesian literatures passing each other like two runners in the night, both having left the original path laid out by the Cowles Commission and its predecessors. To answer the question posed to me at the conference, I see no way to resolve the debates between these two literatures. The RBC literature is only interested in testing in a very limited way, and the new-Keynesian literature is not econometric enough to even talk about serious testing.

But I argue there is hope. Models can be tested, and there are procedures for weeding out inferior models. The RBC literature should entertain the possibility of testing models based on estimating deep structural parameters against models based on estimating approximations of decision equations. Also, the tests should be more than just observing whether a computed path mimics the actual path in a few ways. The new-Keynesian literature should entertain the possibility of putting its various ideas together to specify, estimate, and test structural macroeconometric models.

Finally, both literatures ought to consider bigger models. I have always thought it ironic that one of the consequences of the Lucas critique was to narrow the number of endogenous variables in a model from many (say a hundred or more) to generally no more than three or four. If one is worried about coefficients in structural equations changing, it seems unlikely that getting rid of the structural detail in large-scale models is going to get one closer to deep structural parameters.
Acknowledgments

I am indebted to my discussant, Arnold Zellner, and other participants at the conference for many helpful comments.

Notes

1. See Arrow (1991) and Malinvaud (1991) for interesting historical discussions of econometric research at the Cowles Commission (later Cowles Foundation) and its antecedents.

2. It should be noted that the commercialization of models has been less of a problem in the United Kingdom than in the United States. In 1983 the Macroeconomic Modeling Bureau of the Economic and Social Research Council was established at the University of Warwick under the direction of Kenneth F. Wallis. Various U.K. models and their associated databases are made available to academic researchers through the bureau.

3. If information for period 1 is available at the time the decisions are made, then $E_0W_1$, $E_0P_1$, and $E_0P_1$ should be replaced by the actual values in equations (4) and (5).

4. See Fair and Taylor (1983) for a description of this procedure. This procedure is based on the assumption of certainty equivalence, which is only an approximation for nonlinear models.

5. This is assuming that one does not search by (1) estimating a model up to a certain point, (2) solving the model for a period beyond this point, and (3) choosing the version that best fits the period beyond the point. If this were done, then one would have to wait for more observations to provide a good test of the model. Even if this type of searching is not formally done, it may be that information beyond the estimation period has been implicitly used in specifying a model. This might then lead to a better-fitting model beyond the estimation period than is warranted. In this case, one would also have to wait for more observations to see how accurate the model in fact is.

6. One might argue, as Okun's article, "Inflation: Its Mechanics and Welfare Costs," which I did not count in the eight, presents and briefly discusses data in one figure.

7. By Olivier Blanchard.

References


