Two Approaches to Macroeconomic Forecasting

Roy H. Webb

Following World War II, the quantity and quality of macroeconomic data expanded dramatically. The most important factor was the regular publication of the National Income and Product Accounts, which contained hundreds of consistently defined and measured statistics that summarized overall economic activity. As the data supply expanded, entrepreneurs realized that a market existed for applying that increasingly inexpensive data to the needs of individual firms and government agencies. And as the price of computing power plummeted, it became feasible to use large statistical macroeconomic models to process the data and produce valuable services. Businesses were eager to have forecasts of aggregates like gross domestic product, and even more eager for forecasts of narrowly defined components that were especially relevant for their particular firms. Many government policymakers were also enthusiastic at the prospect of obtaining forecasts that quantified the most likely effects of policy actions.

In the 1960s large Keynesian macroeconomic models seemed to be natural tools for meeting the demand for macroeconomic forecasts. Tinbergen (1939) had laid much of the statistical groundwork, and Klein (1950) built an early prototype Keynesian econometric model with 16 equations. By the end of the 1960s there were several competing models, each with hundreds of equations. A few prominent economists questioned the logical foundations of these models, however, and macroeconomic events of the 1970s intensified their concerns. At the time, some economists tried to improve the existing large macroeconomic models, but others argued for altogether different approaches. For example, Sims (1980) first criticized several important aspects of the large models and then suggested using vector autoregressive (VAR) models for macroeconomic forecasting. While many economists today use VAR models, many others continue to forecast with traditional macroeconomic models.

■ Roy.Webb@rich.frb.org. The views expressed are the author's and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

This article first describes in more detail the traditional and VAR approaches to forecasting. It then examines why both forecasting methods continue to be used. Briefly, each approach has its own strengths and weaknesses, and even the best practice forecast is inevitably less precise than consumers would like. This acknowledged imprecision of forecasts can be frustrating, since forecasts are necessary for making decisions, and the alternative to a formal forecast is an informal one that is subject to unexamined pitfalls and is thus more likely to prove inaccurate.

1. TRADITIONAL LARGE MACROECONOMIC MODELS

These models are often referred to as Keynesian since their basic design takes as given the idea that prices fail to clear markets, at least in the short run. In accord with that general principle, their exact specification can be thought of as an elaboration of the textbook IS-LM model augmented with a Phillips curve. A simple version of an empirical Keynesian model is given below:

$$C_t = \alpha_1 + \beta_{11}(Y_t - T_t) + \varepsilon_{1,t} \tag{1}$$

$$I_t = \alpha_2 + \beta_{21}(R_t - \pi^e_{t+1}) + \varepsilon_{2,t} \tag{2}$$

$$M_t = \alpha_3 + \beta_{31}Y_t + \beta_{32}R_t + \varepsilon_{3,t} \tag{3}$$

$$\pi_t = \alpha_4 + \beta_{41}\frac{Y_t}{Y^p_t} + \varepsilon_{4,t} \tag{4}$$

$$\pi^e_{t+1} = \theta_{51}\pi_t + \theta_{52}\pi_{t-1} \tag{5}$$

$$Y \equiv C_t + I_t + G_t. \tag{6}$$

Equation (1) is the consumption function, in which real consumer spending C depends on real disposable income $Y - T$. In equation (2), business investment spending I is determined by the real interest rate $R - \pi^e$. Equation (3) represents real money demand M, which is determined by real GDP Y and the nominal interest rate R.[1] In equation (4), inflation is determined by GDP relative to potential GDP Y^p; in this simple model, this equation plays the role of the Phillips curve.[2] And in equation (5), expected inflation π^e during the

---

[1] The same letter is used for GDP and personal income since in this simple model there are no elements such as depreciation or indirect business taxes that prevent gross national product from equaling national income or personal income.

[2] In this article the role of the Phillips curve is to empirically relate the inflation rate and a measure of slack in the economy. In a typical large Keynesian model, the Phillips curve would be an equation that relates wage growth to the unemployment rate, with an additional equation that relates wage growth to price changes and another relating the unemployment rate to GDP relative to potential.

next period is assumed to be a simple weighted average of current inflation and the previous period's inflation. Equation (6) is the identity that defines real GDP as the sum of consumer spending, investment spending, and government spending G. In the stochastic equations, ε is an error term and α and β are coefficients that can be estimated from macro data, usually by ordinary least squares regressions. The Θ coefficients in equation (5) are assumed rather than estimated.[3]

One can easily imagine more elaborate versions of this model. Each major aggregate can be divided several times. Thus consumption could be divided into spending on durables, nondurables, and services, and spending on durables could be further divided into purchases of autos, home appliances, and other items. Also, in large models there would be equations that describe areas omitted from the simple model above, such as imports, exports, labor demand, and wages. None of these additions changes the basic character of the Keynesian model.

To use the model for forecasting, one must first estimate the model's coefficients, usually by ordinary least squares. In practice, estimating the model as written would not produce satisfactory results. This could be seen in several ways, such as low R^2 statistics for several equations, indicating that the model fits the data poorly. There is an easy way to raise the statistics describing the model's fit, however. Most macroeconomic data series in the United States are strongly serially correlated, so simply including one or more lags of the dependent variable in each equation will substantially boost the reported R^2 values. For example, estimating equation (2) above from 1983Q1 through 1998Q4 yields an R^2 of 0.02, but adding the lagged dependent variable raised it to 0.97. What has happened is that investment has grown with the size of the economy. The inclusion of any variable with an upward trend will raise the reported R^2 statistic. The lagged dependent variable is a convenient example of a variable with an upward trend, but many other variables could serve equally well. This example illustrates that simply looking at the statistical fit of an equation may not be informative, and economists now understand that other means are necessary to evaluate an empirical equation or model. At the time the Keynesian models were being developed, however, this point was often not appreciated.

Once the model's coefficients have been estimated, a forecaster would need future time paths for the model's exogenous variables. In this case the exogenous variables are those determined by government policy—G, T, and M—and potential GDP, which is determined outside the model by technology. And although the money supply is ultimately determined by monetary policy,

---

[3] The coefficients are assumed, rather than estimated, due to the problematic nature of existing data on actual expectations of inflation.

the Federal Reserve's policy actions immediately affect the federal funds rate. Thus rather than specifying a time path for the money supply, analysts would estimate the money demand equation and then rearrange the terms in order to put the interest rate on the left side. The future time path for short-term interest rates then became a key input into the forecasting process, although its source was rarely well documented.

Next, one could combine the estimated model with the recent data for endogenous variables and future time paths for exogenous variables and produce a forecast. With most large Keynesian models that initial forecast would require modification.[4] The reason for modifying the forecast is to factor in information that was not included in the model. For example, suppose that the model predicted weak consumer spending for the current quarter, but an analyst knew that retail sales grew rapidly in the first two months of the quarter. Or suppose that the analyst observes that consumer spending had been more robust than the model had predicted for the last several quarters. Also, the model's forecast might display some other property that the analyst did not believe, such as a continuously falling ratio of consumer spending to GDP. These are all examples of information that could lead an analyst to raise the forecast for consumer spending above the model's prediction. To change the forecast an analyst would use "add factors," which are additions to the constant terms in the equations above. Thus if one wanted to boost predicted consumer spending by $100 billion in a particular quarter, the analyst would add that amount to the constant term for that quarter. In the model given above, there are four constant terms represented by the α coefficients. To forecast ahead eight quarters, one could consider 32 possible add factors that could modify the forecast. Add factors have long been a key part of the process that uses Keynesian models to produce forecasts and are still important. For example, an appendix to a recent forecast by Data Resources, a leading econometric forecasting service that uses a Keynesian model, lists over 10,000 potential add factors.

2. CRITICISMS OF KEYNESIAN MODELS FOR FORECASTING

One of the most critical components of an economywide model is the linkage between nominal and real variables. The Phillips curve relation between wage or price growth and unemployment rates provided that key linkage for Keynesian macroeconomic models. The Phillips curve was discovered, however, as an empirical relationship. Thus when it was first incorporated in Keynesian

---

[4] Not every Keynesian model required modification, however. Fair (1971), for example, presented a model that has evolved over time but has not made use of the add factors defined below.

models, it did not have a firm theoretical foundation in the sense that it was not derived from a model of optimizing agents. Milton Friedman (1968) criticized the simple Phillips curve, similar to equation (5), at the time that it appeared to be consistent with the unemployment and inflation rates that had been observed in the 1950s and the 1960s. His concern was that the Phillips curve may at times appear to give a stable relation between the amount of slack in the economy and the inflation rate. But suppose that the Federal Reserve were to ease monetary policy in an attempt to permanently raise output above potential. The model above ignores the fact that people would eventually figure out the new policy strategy, and thus, according to Friedman's logic, an expectations formation equation such as (5) would no longer hold. In the long run, he argued, an attempt to hold output above potential would fail; expectations would fully adjust to the new policy and output would return to potential, but inflation would be permanently higher.

Friedman's verbal exposition was very influential, but it did not contain a fully specified analytical model. Using a formal model that captured Friedman's insight, Lucas (1972) introduced rational expectations to macroeconomic analysis as a key element for constructing a dynamic macro model. Among the important conclusions of that paper, he demonstrated that a Phillips curve could fit previously observed data well but would not be valid if the monetary policy process were to change. The book that contained the Lucas paper also contained several papers that presented long-run Phillips curves from leading Keynesian models; a representative result of those models was that a 4 percent rate of unemployment corresponded to 3.5 percent inflation and that higher inflation would give lower unemployment (Christ 1972).

Propositions in economics are rarely tested decisively. In this case, though, it was soon clear that the simple Phillips curve was not a stable, dependable relation. In the fourth quarter of 1972 the unemployment rate averaged 5.4 percent and consumer inflation over the previous four quarters was 3.3 percent. By the third quarter of 1975, unemployment had risen to 8.9 percent; the inflation rate, however, did not fall but instead rose to 11.0 percent.

In retrospect, one can identify many problems with the Keynesian models of that period. Some could be resolved without making wholesale change to the models. For example, most models were changed to incorporate a natural rate of unemployment in the long run, thereby removing the permanent trade-off between unemployment and inflation. Also, most large Keynesian models were expanded to add an energy sector, so that exogenous oil price changes could be factored in. But some of the criticisms called for a fundamental change in the strategy of building and using macroeconomic models.

One of the most influential was the Lucas (1976) critique. Lucas focused on the use of econometric models to predict the effects of government economic policy. Rather than thinking of individual policy actions in isolation, he defined policy to mean a strategy in which specific actions are chosen in

order to achieve well-defined goals. As an example of this meaning of policy, consider the possibility that the Federal Reserve changed interest rates during the early 1960s in order to keep GDP close to potential and inflation low. That behavior could be represented as a reaction function such as equation (7):

$$R_t = R_{t-1} + \beta_{61}\frac{Y_t}{Y_t^p} + \beta_{62}\pi_t + \varepsilon_{6,t}. \tag{7}$$

Now suppose that the reaction function changed in the late 1960s and that less importance was placed on achieving a low rate of inflation. One can imagine replacing equation (7) with the new reaction function; however, Lucas argued that even with the new reaction function, a model would not give reliable policy advice. The reason is that the parameters of all the other equations reflect choices that were made when the previous policy rule was in effect. Under the new policy rule the parameters could well be significantly different in each equation above. This result is easiest to see in equation (6), which describes the formation of expectations of inflation in a manner that might be reasonable for a period when the monetary authority was stabilizing inflation. Individuals could do better, though, if the monetary policy strategy was in the process of changing substantially. During that period an analyst who wanted to produce reliable conditional forecasts would need to replace equation (6), even if the model as a whole continued to provide useful short-term forecasts of overall economic activity. As Lucas (1976, p. 20) put it, "the features which lead to success in short-term forecasting are unrelated to quantitative policy evaluation, . . . [T]he major econometric models are (well) designed to perform the former task only, and . . . simulations using these models can, in principle, provide *no* useful information as to the actual consequences of alternative economic policies."

This critique presented a difficult challenge for macroeconomic model builders. Every macroeconomic model is a simplification of a very complex economy, and the Keynesian models are no exception. One of the key elements of Keynesian models is that prices do no adjust instantaneously to equate supply and demand in every market. The reasons underlying sluggish price adjustment are not usually modeled, however. Thus the models cannot answer the question of to what extent, in response to a policy change, the sluggishness of price adjustment would change. The Lucas critique challenged the reliability of policy advice from models that could not answer such a basic question.

Analysts continue to offer policy advice based on Keynesian models and also other macroeconomic models that are subject to the Lucas critique. These analysts are in effect discounting the relevance of the possibility that their estimated coefficients could vary under the type of policy change analyzed by Lucas. For a succinct example of the reasoning that would allow the use of Keynesian models for policy analysis, consider the counterargument given by Tobin (1981, p. 392), "Lucas's famous 'critique' is a valid point . . . [but]

the critique is not so devastating that macroeconomic model-builders should immediately close up shop. The public's perception of policy regimes is not so precise as to exclude considerable room for discretionary policy moves that the public would see neither as surprises nor as signals of a systematic change in regime. Moreover, behavioral 'rules of thumb,' though not permanent, may persist long enough for the horizons of macroeconomic policy-makers." Sims (1982) gave a lengthier defense of traditional policy analysis.

Authors such as Lucas and Sargent (1979) and Sims (1980) also criticized Keynesian models for not being based on intertemporal optimizing behavior of individuals. At the time they recommended different strategies for model building. Since that time, however, there have been notable improvements in the economic theory embodied in Keynesian models. For example, in the Federal Reserve Board's FRB/US model, it is possible to simulate the model under the assumption that the expectations of individuals are the same as the entire model's forecasts (Brayton et al. 1997). And many modelers have successfully derived individual equations from optimizing dynamic models. Still, Keynesian models continue to be based on unmodeled frictions such as sluggish price adjustment. It is therefore not surprising that economists have explored alternative methods of forecasting and policy analysis. One important method was proposed by Sims (1980) and is discussed in the next section.

3. VAR MODELS

VAR models offer a very simple method of generating forecasts. Consider the simplest reasonable forecast imaginable, extrapolating the recent past. In practice, a reasonably accurate forecast for many data series from the United States over the past half century can be made by simply predicting that the growth rate observed in the previous period will continue unchanged. One could do better, though, by substituting a weighted average of recent growth rates for the single period last observed. That weighted average would be an autoregressive (AR) forecast, and these are often used by economists, at least as benchmarks. Only slightly more complicated is the idea that, instead of thinking of an autoregressive forecast of a single variable, one could imagine an autoregressive forecast of a vector of variables. The advantage of such a VAR relative to simpler alternatives would be that it allowed for the possibility of multivariate interaction. The simplest possible VAR is given below in equations (8) and (9), with only two variables and only one lagged value used for each variable; one can easily imagine using longer lag lengths and more variables:

$$R_t = a_{11}R_{t-1} + a_{12}p_{t-1} + u_{1,t} \tag{8}$$

$$p_t = a_{21}R_{t-1} + a_{22}p_{t-1} + u_{2,t}. \tag{9}$$

Because of the extreme simplicity of the VAR model, it may seem unlikely to produce accurate forecasts. Robert Litterman (1986), however, issued a

series of forecasts from small VAR models that incorporated from six to eight variables. The results, summarized in Table 1, are root mean squared errors (RMSEs), that is, $e = \sqrt{\sum_t (A_t - P_t)^2}$, where e is the RMSE, A is the actual value of a macroeconomic variable, and P is the predicted value. One caveat is that the data summarized in this table cover a relatively short time period, and thus it is a statistically small sample. Over that period, in comparison with forecasts from services using large Keynesian models, the VAR forecasts were more accurate for real GNP more than one quarter ahead, less accurate for inflation, and of comparable accuracy for nominal GNP and the interest rate.

In another study, Lupoletti and Webb (1986) also compared VAR forecasts to those of commercial forecasting services over a longer time period than in the previous comparison. A different caveat applies to their results, shown in Table 2. They studied simulated forecasts versus actual forecasts from the forecasting services. While the details[5] of the simple model were not varied to obtain more accurate forecasts, it is inevitable in such studies that if the VAR forecasts had been significantly less accurate, then the results probably would not have seemed novel enough to warrant publication. That said, their five-variable VAR model produced forecasts that, for four and six quarters ahead, were of comparable accuracy to those of the commercial forecasting services. The commercial services predicted real and nominal GNP significantly more accurately for one and two quarters ahead, which probably indicates the advantage of incorporating current data into a forecast by using add factors.[6]

The VAR model studied by Lupoletti and Webb has five variables, each with six lags. With a constant term, each equation contains 31 coefficients to be estimated—a large number relative to the length of postwar U.S. time series. Although there are methods to reduce the effective number of coefficients that need to be estimated, the number of coefficients still rises rapidly as the number of variables is increased. Thus as a practical matter, any VAR model will contain only a fairly small number of variables. As a result, a VAR model will always ignore potentially valuable data. How, then, is it possible for them

---

[5] In this case the authors could have changed the start date of the regressions used to estimate the VAR model's coefficients, the choice of variables (monetary base versus M1 or M2, for example), or the number of lag lengths. In addition, this model was unrestricted, whereas most VAR forecasters use restrictions to reduce the effective number of estimated coefficients; experimenting with methods of restricting parameters would have lowered the average errors of the VAR forecasts.

[6] For example, an analyst might note that labor input, measured as employee hours, was increasing rapidly in a quarter in which GDP was forecast to rise slowly. The unexpected increase in employee hours could indicate that labor demand had risen due to unexpectedly rapid GDP growth. If other data were consistent with that line of reasoning, the analyst would then increase the constant terms in the equations determining GDP for the current quarter and quite possibly the next quarter as well. Since statistical agencies release important new data every week, there are many such opportunities for skilled analysts to improve forecast accuracy by informally incorporating the latest data.

**Table 1 Average Forecast Errors from Forecasts Made
in the Early 1980s**

Variable: Forecast Horizon (Quarters)	Chase	DRI	WEFA	BVAR
Real GNP:				
1	2.4	2.0	3.1	2.8
2	2.6	2.3	2.6	2.1
4	2.7	2.5	2.4	1.9
8	2.0	2.0	1.7	1.3
GNP deflator:				
1	1.4	1.4	1.9	2.5
2	1.0	1.1	1.5	2.5
4	1.4	1.4	1.7	3.3
8	2.2	2.2	2.4	4.1
Nominal GNP:				
1	3.2	2.7	3.7	3.6
2	3.2	2.7	3.6	3.3
4	3.6	3.2	3.8	4.0
8	3.6	3.6	2.4	3.5
Treasury bill rate:				
1	0.2	0.1	0.4	0.1
2	1.8	1.9	1.8	1.7
4	3.3	3.2	3.2	2.9
8	2.9	3.7	1.1	3.7

Notes: Data are root mean squared errors (RMSEs) from postsample forecasts. Forecasts are from 1980Q2 to 1985Q1. Forecasts of real GNP, the GNP deflator, and nominal GNP are percentage changes from the previous quarter, and forecasts of the Treasury bill rate are cumulative changes in the quarterly average level. Data are from McNees (1986). Forecasts from WEFA were made in mid-quarter, and the others were made one month later.

to produce relatively accurate forecasts? One possibility is that there is only a limited amount of information in *all* macroeconomic time series that is relevant for forecasting broad aggregates like GDP or its price index and that a shrewdly chosen VAR model can capture much of that information.

At best, then, a VAR model is a satisfactory approximation to an under-lying structure that would be better approximated by a larger, more complex model. That more complex model would include how government policymak-ers respond to economic events. The VAR approximation will be based on the average response over a particular sample period. A forecast from a VAR model will thus be an unconditional forecast in that it is not conditioned on any particular sequence of policy actions but rather on the average behavior of policymakers observed in the past. A forecast from a Keynesian model,

Table 2 Average Forecast Errors from Simulated Forecasts

Variable: Forecast Horizon (Quarters)	Chase	DRI	WEFA	VAR
Real GNP:				
1	4.1	4.0	4.2	5.3
2	3.1	3.1	2.9	4.1
4	2.5	2.5	2.2	2.8
6	2.3	2.3	1.9	2.4
GNP deflator:				
1	1.8	2.0	1.9	2.3
2	1.9	2.0	1.9	2.0
4	2.2	2.1	2.0	2.1
6	2.5	2.4	2.2	2.4
Nominal GNP:				
1	5.1	4.6	4.9	6.0
2	4.1	3.6	3.8	4.7
4	3.5	3.0	3.0	3.3
6	3.3	2.7	2.6	3.1
Treasury bill rate:				
1	1.5	1.4	. . .	1.3
2	2.2	2.1	. . .	2.1
4	2.9	2.6	. . .	2.8
6	3.5	3.2	. . .	3.5

Notes: Data are root mean squared errors (RMSEs) from postsample forecasts. Ranges for RMSEs are: one-quarter forecasts, 1970:4–1983:4; two-quarter forecasts, 1971:1–1983:4; four-quarter forecasts, 1971:3–1983:4; and six-quarter forecasts, 1972:1–1983:4. The VAR forecasts are simulated forecasts, as described in the text. Forecasts of real GNP, the GNP deflator, and nominal GNP are cumulative percentage changes, and forecasts of the Treasury bill rate are for its quarterly average level.

however, usually is based on a particular sequence of policy actions and is referred to as a conditional forecast—that is, conditional on that particular sequence. Despite the Lucas critique, many users of Keynesian models seek to determine the consequences of possible policy actions by simulating their model with different time paths of policy actions. But, although the Lucas critique was discussed above in reference to Keynesian models, it is equally valid for VAR models. To help emphasize this point, the next section reviews some details of using a VAR model for forecasting.

Forecasting with VAR Models

To forecast with the VAR model summarized in equations (8) and (9), one would estimate the a_{ij} coefficients, usually by ordinary least squares, and

calculate period t values based on data for period $t - 1$. One can then use the period t forecasts to calculate forecasts for period $t + 1$; for example, inflation forecasts in the above model would be

$$\hat{p}_{t+1} = a_{21}\hat{R}_t + a_{22}\hat{p}_t + \hat{u}_{2,t+1}$$

$$= (a_{21}a_{11} + a_{22}a_{21})R_{t-1} + (a_{21}a_{12} + a_{22}^2)p_{t-1} + a_{21}\hat{u}_{1,t} + a_{22}\hat{u}_{2,t} + \hat{u}_{2,t+1}, \quad (10)$$

where the second line in (10) was obtained by taking the first line and substituting the right-hand sides of (8) and (9) for the estimated values of R_t and p_t, respectively. The above procedure can be repeated as many times as needed to produce as long a forecast as desired.

It is often assumed that the realizations of unknown error terms—$u_{1,t}, u_{2,t},$ and $u_{2,t+1}$—will all equal zero. One can discard that assumption to incorporate information that was not used to estimate the model. Suppose the above model uses monthly data, and at the beginning of a month one knows last month's average interest rate but not the inflation rate, which the Labor Department will release two weeks later. One could simply substitute the realized interest rate for the estimated rate in the calculations above; in equation (10) that would mean plugging in the realized value of $u_{1,t}$. Since the errors in a VAR are usually contemporaneously correlated, a realization of $u_{1,t}$ will also provide information about $u_{2,t}$. Specifically, the variances and covariances of the error terms are taken from the variance-covariance matrix that was estimated through period $t - 1$ when the a_{ij} coefficients were estimated; the expected value of $u_{2,t}$ is then the ratio of the estimated covariance of u_1 and u_2 to the estimated variance of u_1 times the realization of $u_{1,t}$. This expected value of $u_{2,t}$ can then also be included in equations (8) and (9) in order to forecast inflation in periods t and $t + 1$. One can easily apply this basic method for forecasting with a VAR, and the refinement for incorporating partial data for a period, to more complicated models with longer lags, more variables, and deterministic terms such as constants, time trends, and dummy variables.

With this background in mind, imagine that the true structure of the economy is given by the Keynesian model of equations (1) through (6) along with the monetary reaction function (7). Now suppose that the VAR model represented by equations (8) and (9) is estimated. Algebraic manipulation[7] yields the estimated coefficients of the VAR model as functions of the underlying structural coefficients and error terms in equations (8′) and (9′):

$$\pi_t = B_{1,t} + A_{11}\pi_{t-1} + A_{12}R_{t-1} + U_{1,t} \qquad (8')$$

---

[7] In brief, substitute equations (1) and (2) into (6), solve for Y, then substitute the resulting expression for Y into equation (3), and rearrange terms so that π_t is on the left. Next, solve equations (4) and (7) for Y/Y^p, equate the resulting expressions, and rearrange terms so that R_t is on the left. The resulting two equations for π_t and R_t can be solved for each variable as an expression containing lagged values of π and R, exogenous variables, structural error terms, and underlying structural coefficients.

$$R_t = B_{2,t} + A_{21}\pi_{t-1} + A_{22}R_{t-1} + U_{2,t}, \tag{9'}$$

where

$$B_{1,t} = [(\alpha_1 + \alpha_2 - \beta_{11}T_t + G_t + (1 - \beta_{11})(\alpha_3 - M_t))/\beta_{21}\theta_{51} + (\beta_{21} - \alpha_4\frac{\beta_{61}}{\beta_{41}})]/\delta$$

$$A_{11} = -\frac{\theta_{52}\pi_{t-1}}{\theta_{51}\delta}$$

$$A_{12} = \frac{\beta_{21} + (1 - \beta_{11})\beta_{32}}{\beta_{21}\delta}$$

$$U_{1,t} = [e_{1,t} + e_{2,t} + (1 - \beta_{11})e_{3,t}]/\beta_{21}\theta_{51}\delta$$

$$B_{2,t} = [\alpha_1 + \alpha_2 + \alpha_3(1 - \beta_{11}) + \frac{\beta_{61}}{\beta_{41}}\alpha_4 + G_t - T_t - (1 - \beta_{11})M_t]\frac{\beta_{41}\beta_{62} + \beta_{61}}{\beta_{21}\beta_{41}\theta_{51}\delta}$$

$$A_{21} = -\frac{\beta_{21}\theta_{51}(\beta_{41}\beta_{62} + \beta_{61})}{\beta_{21}\beta_{41}\theta_{51}\delta}$$

$$A_{22} = \frac{1}{\delta}$$

$$U_{2,t} = [\varepsilon_{1,t} + \varepsilon_{2,t} + \varepsilon_{3,t}(1 - \beta_{11})]\frac{\beta_{41}\beta_{62} + \beta_{61}}{\beta_{21}\beta_{41}\theta_{51}\delta} + \varepsilon_{4,t}\frac{\beta_{61}}{\beta_{41}\delta} + \varepsilon_{5,t}\delta^{-1}$$

$$\delta = (1 - \frac{\beta_{21} + (1 - \beta_{11})\beta_{32}}{\beta_{21}})(\beta_{62} + \frac{\beta_{61}}{\beta_{41}}).$$

Viewing the model as equations (8′) and (9′) reveals the problematic nature of conditional forecasting with the model. Suppose an analyst wishes to study the effect of a tighter monetary policy on the inflation rate by first obtaining a baseline forecast from the VAR model and then raising the interest rate prediction by a full percentage point for the next quarter. This step would be accomplished by feeding in a particular nonzero value for $u_{2,t+1}$ in equation (10). However, note that in terms of the underlying structure, the error term $U_{2,t}$ is a complicated composite of the five error terms from the equations of the underlying model. Yet for policy analysis it would be necessary to identify that composite error term as a monetary policy disturbance.[8]

An identification that ignores the distinction between VAR errors, the $u_{i,t}$s, and the underlying structural errors, such as the $\varepsilon_{j,t}$'s in the example above, can lead to absurd results. Suppose one simulates a tighter monetary policy in the model presented above by forcing the VAR model to predict higher interest rates; the outcome is a *higher* inflation prediction. The reason is that,

---

[8] This point is not new—see Cooley and LeRoy (1985).

in the quarterly macroeconomic time series of the last 50 years, the dominant shocks to interest rates and inflation have been aggregate demand shocks, and a positive aggregate demand shock raises interest rates, inflation, output, and employment. The VAR model captures these correlations. Asking the model to simulate a higher interest rate path will lead it to predict a higher inflation path as well. Now a clever user can tinker with the model—adding variables, changing the dates over which the model was estimated, and so forth—and eventually develop a VAR model that yields a lower inflation path in response to higher interest rates. At this point, though, the model would add little value beyond reflecting the user's prior beliefs.

To recap, VAR models are unsuited to conditional forecasting because a VAR residual tends to be such a hodgepodge. In addition, the models are vulnerable to the Lucas critique. Suppose that the monetary authority decided to put a higher weight on its inflation target and a lower weight on its output target and that its new reaction function could be represented by (7′):

$$R_t = R_{t-1} + (\beta_{61} - \phi)\frac{Y_t}{Y_t^p} + (\beta_{62} + \phi)\pi_t + \varepsilon_{6,t}. \tag{7′}$$

The interpretation of the VAR's coefficients in terms of the underlying structural coefficients would also change, with each instance of β_{61} changing to $\beta_{61} - \phi$ and each instance of β_{62} changing to $\beta_{62} + \phi$. Thus following a discrete change in the monetary strategy, the VAR's coefficients would be systematically biased and even the accuracy of its unconditional forecasts would be compromised.

Some authors, including Sims (1982), have questioned whether large policy changes in the United States have resulted in meaningful parameter instability in reduced forms such as VARs. One of the most dramatic changes in estimated coefficients in VAR equations for U.S. data occurred in an inflation equation. Table 3 is reproduced from Webb (1995) and shows significant changes in an inflation equation's coefficients estimated in different subperiods.[9] The subperiods, moreover, were determined by the author's review of minutes of the Federal Open Market Committee in order to find monetary policy actions that could indicate a discrete change in the monetary strategy. The results are thus consistent with the view that the monetary reaction function changed substantially in the mid-1960s and again in the early 1980s and that the changes in the economic structure played havoc with a VAR price equation's coefficients.

This section has thus presented two separate reasons for distrusting conditional forecasts from VAR models. First, their small size guarantees that residuals will be complicated amalgamations, and no single residual can be meaningfully interpreted as solely resulting from a policy action. Second,

---

[9] Consider, for example, the sum of coefficients on the nominal variables—inflation, the monetary base, and the nominal interest rate. In the early period the sum is 0.17, rising to 1.23 in the middle period, and then falling to 0.80 in the final period.

Table 3 Regression Results for Several Time Periods

1952Q2 to 1966Q4 $\bar{R}^2 = -0.08$

$\hat{p}_t = 0.28 - 0.08p_{t-1} + 0.08p_{t-2} + 0.11p_{t-3} + 0.07r_{t-1} + 0.02c_{t-1} - 0.01m_{t-1} + 0.05y_{t-1}$
$\quad\ \ (0.06)\ (-0.54)\quad (0.51)\qquad (0.72)\qquad (0.24)\qquad (0.28)\qquad (-0.01)\qquad (0.76)$

1967Q1 to 1981Q2 $\bar{R}^2 = 0.57$

$\hat{p}_t = -2.78 + 0.30p_{t-1} - 0.04p_{t-2} + 0.04p_{t-3} + 0.33r_{t-1} + 0.02c_{t-1} + 0.60m_{t-1} - 0.08y_{t-1}$
$\quad\ (-0.54)\ (2.30)\quad (-0.28)\qquad (0.27)\qquad (2.56)\qquad (0.26)\qquad (4.87)\qquad (-1.52)$

1981Q3 to 1990Q4 $\bar{R}^2 = 0.51$

$\hat{p}_t = -8.87 + 0.21p_{t-1} + 0.09p_{t-2} + 0.20p_{t-3} + 0.20r_{t-1} + 0.10c_{t-1} + 0.10m_{t-1} - 0.15y_{t-1}$
$\quad\ (-1.54)\ (1.16)\qquad (0.53)\qquad (1.15)\qquad (1.07)\qquad (1.68)\qquad (1.02)\qquad (-0.21)$

1952Q2 to 1990Q4 $\bar{R}^2 = 0.59$

$\hat{p}_t = -3.84 + 0.30p_{t-1} + 0.23p_{t-2} + 0.22p_{t-3} + 0.005r_{t-1} + 0.05c_{t-1} + 0.17m_{t-1} - 0.22y_{t-1}$
$\quad\ (-1.42)\ (6.38)\qquad (2.89)\qquad (2.71)\qquad (0.07)\qquad (1.54)\qquad (2.95)\qquad (-0.59)$

Note: Coefficients were estimated by ordinary least squares; t-statistics are in parentheses.

applying the Lucas critique to VAR models implies that a VAR model's coefficients would be expected to change in response to a discrete policy change.

Several researchers who have recognized these deficiencies but were unwilling to give up the simplicity of the VAR approach have turned to structural VARs, or SVARs.[10] These models attempt to apply both economic theory that is often loosely specified and statistical assumptions to a VAR in order to interpret the residuals and conduct meaningful policy analysis. In many studies key statistical assumptions are that the economy is accurately described by a small number of equations containing stochastic error terms, and that these structural errors are uncorrelated across equations. The economic restrictions vary considerably from model to model; the common feature is that just enough restrictions are introduced so that the reduced-form errors, such as in equations (7′) and (8′) above, can be used to estimate the structural errors. For example, two of the restrictions used in a widely cited paper by Blanchard (1989) were (1) that reduced-form GDP errors were equal to structural aggregate demand errors, and (2) that reduced-form unemployment errors, given output, were equal to structural supply errors. After presenting those and other restrictions, the author noted "There is an obvious arbitrariness to any set of identification restrictions, and the discussion above is no exception" (p. 1150).

---

[10] A clear exposition of the SVAR approach is given by Sarte (1999).

It is often the case that a reader will find an identifying assumption of an SVAR somewhat questionable. A major difficulty of the SVAR approach is that there is no empirical method for testing a restriction. Moreover, if different models give different results, there are no accepted performance measures that can be used to identify superior performance. Since there are millions of possible SVARS that could be based on the last half century of U.S. macroeconomic data, their results will not be persuasive to a wide audience until a method is found to separate the best models from the rest.[11]

4.   FINAL THOUGHTS ON CONDITIONAL FORECASTING

This article has discussed two approaches to macroeconomic forecasting. Both approaches have produced econometric models that fit observed data reasonably well, and both have produced fairly accurate unconditional forecasts. The VAR approach was found unsuitable for conditional forecasting and policy analysis. There is a wide division within the economics profession on the usefulness of large Keynesian models for policy analysis. At one extreme are those who accept the Lucas critique as a fatal blow and accordingly see little value in using Keynesian models for policy analysis. At the other extreme are analysts who are comfortable with traditional Keynesian models. In the middle are many economists with some degree of discomfort at using the existing Keynesian models, in part due to the features that allow the models to fit the historical data well but may not remain valid in the event of a significant policy change. But policy analysis will continue, formally or informally, regardless of economists' comfort with the models and with the strategies for using them. Decisions on the setting of policy instruments will continue to be made and will be based on some type of analysis.

One possibility is that policy analysis and economic forecasting will be seen as two different problems requiring two different types of models. Economists have constructed a large number of small structural models that can be quantified and used for policy analysis. A large number of statistical approaches to forecasting are available as well. It is not necessary that the same model be used for both.

Keynesian models, though, are still widely used for policy analysis, and there are actions that model builders could take to enhance the persuasiveness of their results. One would be to publish two forecasts on a routine basis—the usual forecast with add factors incorporating the modelers' judgment and a mechanical forecast with no add factors. In that way a user could easily distinguish the judgmental content from the pure model forecast. For example,

---

[11] Other authors have argued that SVAR results are not robust, including Cooley and Dwyer (1998) and Cecchetti and Rich (1999).

if one wanted to determine the possible impact of a tax cut on consumption, one would want to consider whether large add factors in a consumption equation such as equation (1) above were needed to achieve satisfactory results.

It would also be helpful for forecast consumers to know how much a model's specification has changed over time. Of course one hopes that new developments in economics are incorporated into models and that poorly performing specifications are discarded. As a result, some specification changes are to be expected. But if one saw that the consumption equation of a large model had been in a state of flux for several years, the numerous changes could signify that the model's analysis of a tax cut's effect on consumption was based on an unstable foundation.

In addition, it would be helpful to see more analysis of forecast errors. At a minimum, each forecast should be accompanied by confidence intervals for the most important variables stating the likely range of results. As the ex post errors indicate in Tables 1 and 2, these confidence intervals could be quite wide. For example, real GDP growth has averaged 2.8 percent over the last 30 years. In Table 2, the RMSE for four-quarter predictions of real growth from the best commercial forecasting service was 2.2 percent. Thus if a model predicted real growth to be the 2.8 percent average, and one used that RMSE as an approximate standard deviation of future forecast errors, then one would expect actual outcomes to be outside of a wide 0.6 to 5.0 percent range about 30 percent of the time. Now suppose that an exercise in policy analysis with that model revealed a difference of 1.0 percent for real GDP growth over the next year; a user might not consider that difference very meaningful, given the relatively large imprecision of the model's GDP forecast.

Finally, it would be especially helpful to have a detailed analysis of errors in a manner relevant for policy analysis. For example, continuing with the predicted effect of a tax cut, the model's predictions could be stated in the form of a multiplier that related the tax cut to a predicted change in real growth. That multiplier would be a random variable that could be statistically analyzed in the context of the whole model, and the user could be told the sampling distribution of that statistic. Also, one would want data on how well the model predicted the effects of tax cuts that had actually occurred in the past.

The unifying theme of these recommendations is for model builders to open the black box that generates forecasts. Until this supplementary information routinely accompanies the output of large forecasting models, many will see an exercise in policy evaluation as having unknowable properties and value it accordingly.

REFERENCES

Blanchard, Olivier Jean. "A Traditional Interpretation of Macroeconomic Fluctuations," *American Economic Review*, vol. 79 (December 1989), pp. 1146–64.

Brayton, Flint, Eileen Mauskopf, David Reifschneider, Peter Tinsley, and John Williams. "The Role of Expectations in the FRB/US Macroeconomic Model," *Federal Reserve Bulletin*, vol. 83 (April 1997), pp. 227–45.

Cecchetti, Stephen G., and Robert W. Rich. "Structural Estimates of the U.S. Sacrifice Ratio," Federal Reserve Bank of New York, 1999.

Christ, Carl. "Discussion," in *The Econometrics of Price Determination*. Washington: Board of Governors of the Federal Reserve System, 1972, pp. 330–31.

Cooley, Thomas F., and Mark Dwyer. "Business Cycle Analysis without Much Theory: A Look at Structural VARs," *Journal of Econometrics*, vol. 83 (March–April 1998), pp. 57–88.

_____, and Stephan F. LeRoy. "Atheoretical Macroeconomics: A Critique," *Journal of Monetary Economics,* vol. 16 (November 1985), pp. 283–308.

Fair, Ray. *A Short-Run Forecasting Model of the United States Economy.* Lexington, Mass.: D.C. Heath, 1971.

Friedman, Milton. "The Role of Monetary Policy," *American Economic Review*, vol. 58 (March 1968), pp. 1–17.

Klein, Lawrence. *Economic Fluctuations in the United States 1921–1941.* New York: Wiley, 1950.

Litterman, Robert. "Forecasting with Bayesian Vector Autoregressions—Five Years of Experience," *Journal of Business and Economic Statistics,* vol. 4 (January 1986), pp. 25–38.

Lucas, Robert. "Econometric Policy Evaluation: A Critique," in Karl Brunner and Allan H. Meltzer, eds., *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, vol. 1 (1976), pp. 19–46.

_____. "Econometric Testing of the Natural Rate Hypothesis," in *The Econometrics of Price Determination*. Washington: Board of Governors of the Federal Reserve System, 1972, pp. 50–59.

_____, and Thomas Sargent. "After Keynesian Macroeconomics," Federal Reserve Bank of Minneapolis *Quarterly Review*, vol. 3 (Spring 1979), pp. 1–16.

Lupoletti, William, and Roy Webb. "Defining and Improving the Accuracy of Macroeconomic Forecasts: Contributions from a VAR Model," *Journal of Business,* vol. 59 (April 1986), pp. 263–85.

McNees, Stephen. "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts," *Journal of Business and Economic Statistics,* vol. 4 (January 1986), pp. 5–16.

Sarte, Pierre-Daniel G. "An Empirical Investigation of Fluctuations in Manu-facturing Sales and Inventory within a Sticky-Price Framework," Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 85 (Summer 1999), pp. 61–84.

Sims, Christopher A. "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity,* 1:1982, pp. 107–165.

——————. "Macroeconomics and Reality," *Econometrica,* vol. 48 (January 1980), pp. 1–48.

Tinbergen, J. *Business Cycles in the United States: 1912–1932.* League of Nations, 1939.

Tobin, James. "Comments," in J. Kmenta and J. B. Ramsey, eds., *Large-Scale Macro-Econometric Models.* New York: North-Holland, 1981, pp. 391–92.

Webb, Roy. "Forecasts of Inflation from VAR Models," *Journal of Forecasting,* vol. 14 (May 1995), pp. 267–85.