Bad News From a Forecasting Model of the U.S. Economy

David E. Runkle*
Senior Economist
Research Department
Federal Reserve Bank of Minneapolis
and Associate Director
Institute for Empirical Macroeconomics

In 1990, after eight years of expansion, U.S. economic activity slowed down while inflation sped up. This bad economic news is quite different from a forecast made a year ago by a model that researchers at the Federal Reserve Bank of Minneapolis have developed and used for some time. In the fall of 1989, this model predicted strong economic growth and moderate inflation for 1990 (Runkle 1989). The model's 1990 errors may make you more skeptical than usual about its general ability to forecast—but they shouldn't. Historical data show that errors as large as these are not unusual among economic forecasters; they're simply the result of bad luck. Despite its 1990 performance, that is, the model is sound, and its current predictions are as worthy of notice as ever. Unfortunately, those predictions are not good news. The model now says that a recession (as typically defined) is likely in 1991 and neither growth nor inflation will improve through the end of 1992.

A Bad Year for the Model . . .

The forecasting model used by Minneapolis Fed researchers is a Bayesian vector autoregression (BVAR) model. Unlike many others, this type of forecasting model relies exclusively on statistical methods to extrapolate the historical relationships among economic variables into the future. (Human judgment about what might happen in the economy does not influence its forecast.)[1]

Unfortunately, as I said, the BVAR model's last published forecast turned out to be pretty bad. Data for the first three quarters of 1990 (the most recent available as I write) show both slower growth and faster inflation than the model predicted. They also show that the model's errors were larger than those of many other economic forecasters. Chart 1 displays how poorly the model did compared to U.S. business economists on average, represented by the median predictions of forecasters surveyed by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER).[2]

The model's errors in predicting 1990 growth in the inflation-adjusted gross national product (real GNP) are primarily in two components: personal consumption of goods and services and private investment in residential structures. The model predicted that consumer spending would grow at an annual rate of 4.2 percent during the first three quarters of 1990; it actually grew only 1.5 percent. The model expected residential investment to increase at an annual rate of 5.7 percent during that period; instead it decreased 6.2 percent.

The model's errors in predicting inflation, as mea-

*Also Adjunct Associate Professor of Finance, University of Minnesota.

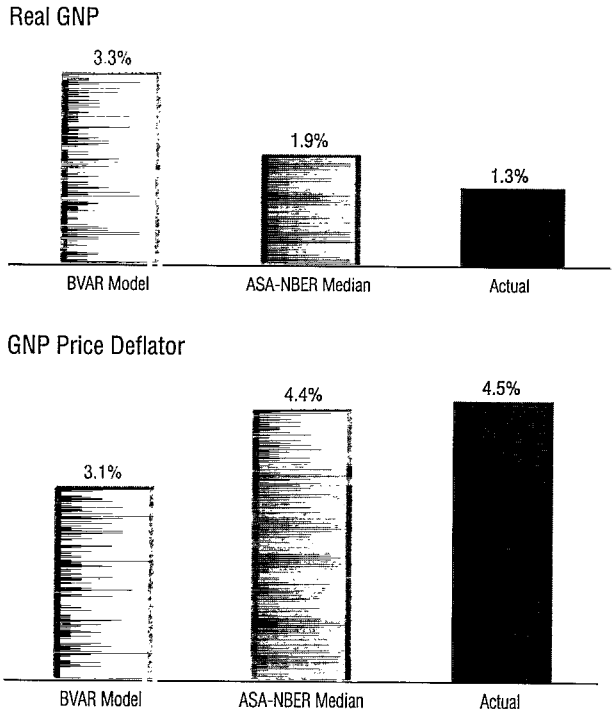
[1]For background on BVAR models like this one, see Litterman 1984 and Todd 1984.

[2]For background on this survey, see Zarnowitz 1969 and Keane and Runkle 1989.

Chart 1

How Bad Was It?

Predicted and Actual Annual Rates of Change in the First Three Quarters of 1990*



*These rates of change are for only the first three quarters because actual fourth quarter data are not yet available (on November 30, 1990). These rates do not match those in Runkle 1989 or on the table below because those are for the full year.

Sources: BVAR model using data available on November 30, 1989; NBER 1989; U.S. Department of Commerce

sured by changes in the GNP price deflator, are not so easily isolated. We can, however, make a good guess that energy price increases had a lot to do with them. Even before the Persian Gulf crisis, energy prices were rising rapidly. Yet, when the model made its 1990 forecast, it predicted that energy prices, as measured by the energy component of the consumer price index, would fall at an annual rate of 1.9 percent during the first three quarters. This measure of energy prices actually rose at an annual rate of 23.5 percent during that period.

. . . That Won't Necessarily Be Repeated

Certainly it is regrettable that the BVAR model made large errors in predicting 1990's real growth and inflation. However, we need not assume that those errors mean the model is seriously flawed. The errors could be no larger than are reasonable for any forecaster, given the intrinsic uncertainty that exists in predicting economic activity.

Three Tests

There are three standard ways to check whether a forecasting model is sound:

- Compare the model's recent errors to its past errors. Are its recent errors unusually large?
- Add some variables to the model. Does that greatly reduce its errors?
- Compare the model's recent errors to the past errors of other forecasters. Are the model's errors unusually large?

If the answer to any of these questions is yes, then probably something is wrong with the model that should be fixed before it forecasts again. But if the answer to all three is no, then we can be fairly confident that recently the model just had some normal bad luck.

□ *The Model's Uncertainty*

The BVAR model can objectively quantify the amount of uncertainty in its own forecast using the history of its errors. As a result, we can examine whether the errors that the model made in 1990 are really unusual for it.

They aren't. The model estimates that at least 30 percent of the time it has made an error as large as it did in predicting the 1990 rate of growth in real GNP. And at least 20 percent of the time it has made an error as large as it did in predicting the 1990 change in the GNP price deflator.

Even the model's worst errors are not unusual by this standard. Errors as large as the one for consumption growth have occurred about 20 percent of the time, while errors as large as the one for residential investment have occurred almost half the time.

□ *An Obvious Omission*

Even though the model's 1990 errors are not unusual for it, they are disturbing because they are larger than the errors made by many other forecasters. Why should that be? Maybe the model does not include a crucial variable that those forecasters used in making their predictions. If so, the model could still be structurally flawed. Such a flaw would be serious because it might imply that the model would continue to make errors larger than those of other forecasters. Thus, we should

try to see if adding some obviously missing variable to the model could make it more accurate.

The most obvious, potentially serious omission in the model is a measure of the market value of the nation's housing stock. Since housing prices have declined recently in many parts of the country, that omission may explain some of the model's error in predicting both consumption growth and residential investment growth. An unexpected reduction in the value of any asset owned by consumers would make them reduce their spending. Houses are the most valuable asset most consumers own, so declining housing prices could cause substantial cutbacks in consumer spending. Declining housing prices would also make investment in housing less attractive, which would reduce residential investment.

Since all forms of wealth should affect consumption growth, we would ideally include in the BVAR model some measure of the market value of all assets owned by consumers. Currently, the only measure of wealth it has is a measure of the value of the stock market. Therefore, adding some measure of the market value of the nation's housing stock should improve the model's forecast of consumption growth.

But adding the best available measure doesn't. That measure is the U.S. Department of Commerce's constant-quality index of new house prices. A statistical test shows that adding this measure to the model does not help it predict growth in real consumption or residential investment.[3] One possible reason for this failure is that the new house price index is a very inaccurate measure of the total market value of the nation's housing stock. Another possible explanation is that lower housing prices affect consumption and residential investment so quickly that even a perfect measure of housing prices would not improve the model's accuracy because the effect of changes in housing prices would be captured directly by other variables.

□ *Other Forecasters' Uncertainty*

One test for a structural flaw remains: Have other economic forecasters been consistently more accurate than the BVAR model? Answering that question is more complicated than you might think. But the answer appears to be no.

One way of comparing accuracy is fairly simple: Compare the model's 1990 errors to errors made by the average forecaster over the last 21 years. According to this comparison, errors as large as the model's in 1990 are common. Between 1969 and 1989, when predicting three quarters ahead, the median ASA-NBER forecaster made larger errors in predicting real GNP changes 38 percent of the time and larger errors in predicting deflator changes 52 percent of the time.

Unfortunately, comparing the BVAR model's errors in 1990 to those of other forecasters over a much longer period does not necessarily mean the model is sound. Even if the model's 1990 errors were small compared to other forecasters' previous errors, the model's errors in previous years could have been larger than those of the other forecasters. Then the model would not be very useful. Therefore, we need to compare the model's long-term accuracy with that of other forecasters.

Before doing that, we should define more explicitly what we mean by *accuracy*. One way to define it is to say that one forecaster is *more accurate* than another if, on average, the error the forecaster makes is closer to zero. We could see which of two forecasters is more accurate according to this definition by plotting how often each of the forecasters make different-sized errors. Chart 2 shows how the errors of two forecasters could be compared using this method. With average error as our guide, forecaster *A* is more accurate than forecaster *B* because *A*'s errors are closer to zero, on average.

But average error should not be the only component in the definition of accuracy. Even if two forecasters have exactly the same average error, one forecaster could be more accurate than the other by making fewer large errors. We could see this, too, by plotting how often different-sized prediction errors occur for each of two forecasters. Chart 3 shows that forecaster *A* and forecaster *B* have the same average forecast error, but here *B* is more accurate than *A* because *B* has fewer large errors.

We hope that, according to both of these definitions, the BVAR model is at least as accurate as the median ASA-NBER forecaster. And it is. In the Appendix, I describe a method of testing simultaneously that each of two sets of forecasts has an average error close to zero and that the difference in the number of large errors in the two sets of forecasts is not statistically significant. I used this method to test whether the three-quarter-ahead forecasts of real GNP and deflator changes that the model would have made at the end of each quarter during the past 21 years are as accurate as the actual

[3] I did Granger-causality tests using quarterly data from 1963:1 to 1990:3 to see whether the log of the constant-quality index of new house prices (deflated by the consumer price index, exclusive of shelter) helped to predict either the log of real consumption or the log of residential investment. It did not.

Charts 2 and 3
Two Views of Accuracy

Chart 2 Size of Average Error
(The Smaller, the More Accurate)

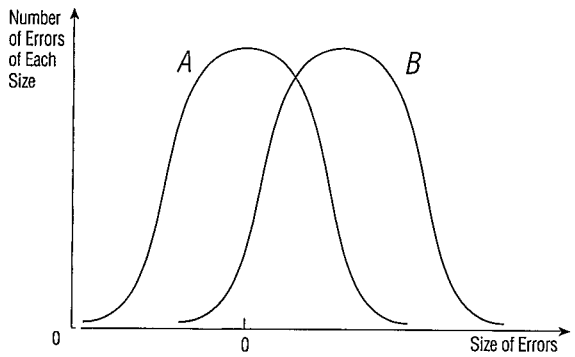
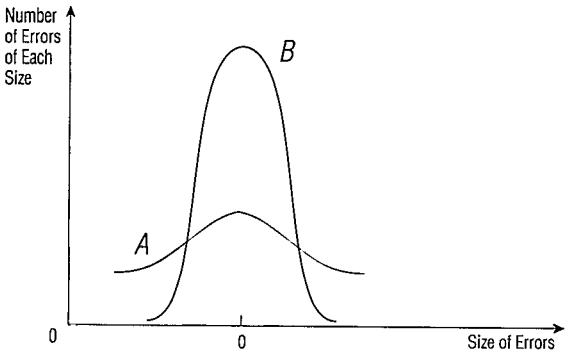


Chart 3 Number of Large Errors
(The Fewer, the More Accurate)



three-quarter-ahead forecasts of those variables made by the median ASA-NBER forecaster.[4] The results of those tests: both the model and the median ASA-NBER forecaster made predictions which have an average error not significantly different from zero, and the number of large errors the two forecasters made is not significantly different. This suggests that the BVAR model's predictions are not systematically less accurate than those of other forecasters.[5]

Two Extra Uses

Thus, the BVAR model has passed all three tests of soundness. Even though the model's 1990 errors are large, they are not unusual compared to the model's own past errors or to the past errors of other forecasters. Furthermore, the errors do not appear to have been caused by the exclusion of a measure of housing prices. This suggests that the model has no serious structural flaws and so can be trusted (as much as ever) to predict beyond 1990. We can also, by the way, use it to perform two tasks that conventional forecasters cannot.

First, as I said, the model can objectively quantify the amount of uncertainty in its forecasts. By using the history of its own errors, the model can simulate the likely range of its errors in the future. Since the model can do this, it can also compute the probability that a certain economic event will happen. For example, the model can compute the probability that during 1991 the U.S. economy will fall into recession (defined in some precise way) and inflation will be greater than 5 percent (or any other number we choose).

Second, the model can answer useful hypothetical questions. For example, as I write, real GNP growth during 1990's fourth quarter is still uncertain. I could ask the model how much more likely a recession would be during the next year if, during the fourth quarter of this year, real GNP contracted at an annual rate of 3 percent instead of growing at an annual rate of 1 percent. Other forecasters cannot answer such questions in an objective, statistical manner.

A Bad Outlook for the Economy

Since Iraq's invasion of Kuwait in August 1990, U.S. economic conditions have deteriorated rapidly. Among the falling indicators are measures of employment, industrial output, and both consumer confidence and spending. These weak economic conditions lead the BVAR model to now predict fairly weak conditions through 1992.

1991: Recession

Near the end of November 1990, the head of the Federal Reserve Board, Alan Greenspan, described the U.S. economy as in a "meaningful downturn" (Rosen-

[4]These model forecasts use revised, not preliminary data, so their errors may be different than those the model actually would have made at the time. However, the model has been used for only six years, so it does not have enough of a forecast history to meaningfully compare its actual forecasts to those of other forecasters. Using revised data is thus the best method of comparing the accuracy of the model's forecasts to that of the median ASA-NBER forecaster.

[5]The model's own past errors also suggest that even the large errors in predicting 1990 consumption and residential investment growth are not unusual. Errors as large as the one for consumption growth occur about 15 percent of the time, while errors as large as the one for residential investment growth occur almost half the time.

baum 1990). But how likely is a *recession,* as usually defined? That is, how likely is it that real GNP will decline at least two quarters in a row? The model predicts that real GNP will decline at an annual rate of 1.5 percent during the fourth quarter of 1990. It says the odds of at least two consecutive quarters of negative real growth occurring between then and the fourth quarter of 1991 are 50-50. But the model also says that a recession starting in those five quarters would not last long. The chances that real growth will be negative in four of the five quarters, it says, are only 12 percent.

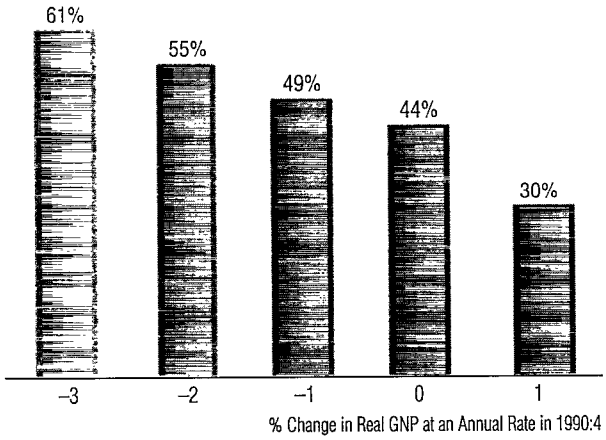
As I said, today there is still considerable uncertainty about how much real GNP will actually grow during the fourth quarter of 1990. But the model can estimate how the amount of real growth during this quarter will affect its estimate of the probability of recession by the end of 1991. Chart 4 shows these estimates. If real GNP grows at an annual rate of 1 percent during the fourth quarter, the model estimates the probability of recession as only 30 percent. But if real GNP falls at an annual rate of 3 percent during the quarter, it estimates that probability as 61 percent.

Restricting *recession* to its usual definition may be too restrictive, however. The NBER, which many economists view as the official arbiter of recessions, does not always use the narrow definition. For example, the NBER decided a recession had occurred in 1980 even though real GNP had declined only during the second quarter.

Chart 4

How the Chances of a 1991 Recession* Depend on Growth in 1990's Fourth Quarter

Estimated by the BVAR Model



*A *recession* here is defined as two or more consecutive quarters of negative real growth in GNP.

Currently, the BVAR model predicts that employment will decline from July 1990 until June 1991 and that industrial production will fall from October 1990 until September 1991. If employment and industrial output fall that long, the NBER will probably declare that the economy is in a recession even if real GNP has not fallen for two consecutive quarters. In fact, since employment has declined in every month since July, the NBER may have already declared a recession by the time you read this. Thus, the model suggests that the probability of a 1991 recession, defined by some broader-than-usual measure, is very high.

1991–92: Weak Growth and Moderate Inflation

As the accompanying table shows, the BVAR model's longer-term outlook is not much better.

The model predicts that real GNP growth will remain weak over the next two years because domestic, private sector demand will remain sluggish. Real GNP is predicted to grow only 1.5 percent between the fourth quarter of 1990 and the fourth quarter of 1991—less than half as much as the average growth during the past 40 years. In fact, growth was that low in only 11 of those 40 years. Growth is expected to pick up only slightly in 1992.

The table also shows what's behind this growth prediction. Over the next two years, the model predicts, consumption and investment spending growth will remain at levels usually seen in recessions. The model expects export growth to be the only source of strength in these years.

Consumer spending is expected to grow at an annual rate of only 1.1 percent during both 1991 and 1992. To put that prediction in perspective, in only two of the seven recessions since 1950 did consumption grow slower than 1.7 percent during a year. Thus, the forecast calls for significantly weaker consumption growth than usually occurs in recessions.

Weak consumption growth may stem from uncertainty about future employment. The model predicts that total employment will grow only 0.4 percent in 1991 and will remain below its July 1990 level until January 1992. The model also predicts that unemployment will rise to 6.5 percent by the end of 1991 and creep even higher by the end of 1992.

Investment spending is expected to decline in each of the next two years. The slump in residential construction

The BVAR Model's Forecast for the U.S. Economy in 1991–92*

Indicator	Actual** 1990	Model Forecast 1991	Model Forecast 1992	1948–89 Average
Annual Growth Rates (4th Qtr. % Changes From Year Earlier)				
Real Gross National Product (GNP)	.6 %	1.5 %	2.1 %	3.3 %
Consumer Spending	.9	1.1	1.1	3.4
Durable Goods	.5	–3.0	–3.5	5.0
Nondurable Goods and Services	1.0	1.8	2.0	3.2
Investment	–2.4	–4.6	–2.0	4.2
Business Fixed	.7	–4.6	.3	3.8
Residential	–8.3	–5.6	–9.5	3.3
Government Purchases	1.7	.1	2.3	3.9
GNP Price Deflator	4.7	4.3	4.3	4.2
4th Quarter Levels				
Change in Business Inventories (1982 $)	13.1 bil.	14.1 bil.	14.6 bil.	13.4 bil.
Net Exports (1982 $) (Exports less Imports)	–45.0 bil.	18.6 bil.	71.0 bil.	–19.4 bil.
Civilian Unemployment Rate (Unemployment as a % of the Civilian Labor Force)	5.8 %	6.5 %	6.7 %	5.7 %

*This is the forecast of a Bayesian vector autoregression model using data available on November 30, 1990.

**Actual numbers are based on data for the first three quarters of 1990 and the BVAR model's forecast for the fourth quarter of 1990 (using data available on November 30, 1990).

Sources of actual data: U.S. Departments of Commerce and Labor

is expected to continue over this time. This suggests that, though the model does not include an explicit measure of housing prices, it has picked up their decline indirectly, in other data that are influenced by them. Business fixed investment is also predicted to remain weak for the next two years.

Both the consumption and the investment forecasts bear patterns typical of recessions. The outlook for consumption includes cuts in spending on durable goods accompanied by below-average growth in spending on nondurables and services. And again, the outlook for investment includes weakness in both the residential and business sectors. These patterns have occurred during each of the seven recessions since 1950.

The fact that the model predicts weakness for both consumption and residential investment over the next two years also suggests that it has learned from its recent errors. BVAR models use their own past errors to adapt to changing patterns among economic variables. And this BVAR model appears to have adjusted quickly to the errors that it made in overpredicting 1990 consumption and residential investment growth.

The model's 1991–92 consumption and investment forecasts are so weak that the only reason the model does not predict an extended U.S. recession is that it predicts a simultaneous large improvement in net exports. It expects the difference between the values of exports and imports to turn around from a large negative at the end of 1990 to a large positive by the end of 1991—and then to nearly quadruple a year later. This prediction comes from a forecast of slower import growth and continued fast export growth.

Is that reasonable? Slower import growth certainly seems so, in the context of the model's anemic forecast for consumption and investment growth. It is also consistent with the historical observation that import demand grows slowly when the economy is weak.

The model's prediction of fast export growth might also seem reasonable, since the value of the U.S. dollar has fallen recently, making U.S. products less expensive

to buyers in other countries. But the prediction of strong export growth will be correct only if the economies of other countries remain strong. If other countries also experience weak consumption and investment demand, then U.S. export growth is likely to be slower than the model predicts.

Slower export growth would, of course, make the U.S. outlook for real growth even worse. For example, if net exports were to improve in both of the next two years by the amount the model now predicts for 1991 ($20 billion) and the rest of the model's forecast was correct, real GNP would grow by less than half a percentage point in both of the next two years.

Compared to the growth forecast, the inflation forecast is optimistic. According to the model, the current spate of inflation, induced by rising oil prices, will not continue into the next two years. The model predicts that the GNP deflator will rise at an annual rate of 5.4 percent in the fourth quarter of 1990, pulling up the 1990 increase to 4.7 percent. But the model also predicts that the increase in the deflator will slip to 4.3 percent in both 1991 and 1992.

A quick look at what other forecasters are predicting for 1991 supports at least that part of the model's forecast. Early in November, the consensus of private economists surveyed by the *Blue Chip Indicators* was for a 1991 increase of 1.2 percent in real GNP and 4.3 percent in the deflator (Blue Chip 1990). This is essentially the same forecast as the BVAR model's.

Beware of Good News

Despite 1990's poor showing, the BVAR model makes good forecasts, on average. Still, more than the average amount of uncertainty remains about the current outlook. One reason for that is political: the current instability in the Mideast. Another is statistical. History shows that during periods of economic contraction, the early estimates of data can be particularly misleading. Even if real GNP actually declines in the fourth quarter, there is a substantial probability that the initial estimates of GNP growth will not show that. According to the data, real GNP has actually declined in 21 quarters since 1968. But in 9 of those 21 quarters, the initial estimates of growth were positive, not negative. Thus, we must be especially careful interpreting the early data for the fourth quarter of 1990. Even if good news is initially reported, the actual situation may be quite different.

Appendix
Comparing the Accuracy of Two Sets of Forecasts

This appendix describes a simple method—used in the preceding paper—of testing whether two sets of forecasts are equally accurate.

Suppose that two different forecasters, i and j, predict the k-step-ahead growth rate of inflation-adjusted gross national product (real GNP) in each period. For example, at time t, forecaster i makes the prediction $F^i_{t,k}$. At time $t + k$, the error in that prediction becomes known, when the actual growth rate, A_{t+k}, becomes known. The error, $E^i_{t,k}$, is equal to $A_{t+k} - F^i_{t,k}$. Suppose that the two forecasters each have a set of forecast errors, $E^i_{t,k}$ and $E^j_{t,k}$ where $t = 1, 2, 3, ..., T$. One condition that seems reasonable to impose is that a set of forecasts will be called *accurate* only if, on average, its errors are close to zero; that is, the forecast is *unbiased.* Two sets of unbiased forecasts might be called *equally accurate* if the average difference of their squared forecast errors is small.

Diebold and Runkle (1990) derive a simple generalized method-of-moments (GMM) specification test for the unbiasedness and equality of average squared errors in two sets of k-step-ahead forecasts.[1] Every GMM specification test requires a set of *orthogonality conditions,* that is, conditions which should have an expected value of zero in the population. The simplest version of Diebold and Runkle's test has three such conditions:

- The expected value of the forecast error from forecaster i is equal to zero.
- The expected value of the forecast error from forecaster j is equal to zero.
- The expected value of the difference in the squared forecast errors is equal to zero.

If the sample averages of these orthogonality conditions are denoted as g and the inverse of the covariance matrix of the orthogonality conditions is denoted as W, then $Tg'Wg$ will be distributed asymptotically as a chi-squared random variable with three degrees of freedom.[2] If the value of that test statistic were larger than 7.81, then the null hypothesis of unbiasedness and equality of average squared errors could be rejected at the 95 percent confidence level.

For tests of the forecasts in the preceding paper, I used 84 quarterly observations on the errors in predicting the annualized rate of change in real GNP and the GNP price deflator. The forecasters were the two in the paper: the Bayesian vector autoregression (BVAR) model developed by researchers at the Minneapolis Fed and the median business economists surveyed by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER). The time period was from the fourth quarter of 1968 to the third quarter of 1989.[3]

The results are clear. The statistic from the chi-squared test comparing the errors that the two forecasters made for real GNP growth was 2.16, while that for the deflator was 4.64. Both of these test statistics are far below the critical value of 7.81. Thus, the test does not reject the hypothesis that for the changes in both real GNP and the deflator, the BVAR model and the median ASA-NBER forecaster made unbiased forecasts with equal average squared errors. In this sense, that is, the two sets of forecasts are equally accurate.

[1]Newey (1985) developed the first GMM specification tests.

[2]For further details of this test, including how to construct W, see Diebold and Runkle 1990.

[3]The model's forecasts are not real-time; they are based on revised data.

References

Blue Chip. 1990. *Blue Chip Economic Indicators* 15 (November 10): 5.

Diebold, Francis X., and Runkle, David E. 1990. Tests for comparing the accuracy of forecasts. Manuscript. Federal Reserve Bank of Minneapolis.

Keane, Michael P., and Runkle, David E. 1989. Are economic forecasts rational? *Federal Reserve Bank of Minneapolis Quarterly Review* 13 (Spring): 26–33.

Litterman, Robert B. 1984. Above-average national growth in 1985 and 1986. *Federal Reserve Bank of Minneapolis Quarterly Review* 8 (Fall): 3–7.

National Bureau of Economic Research (NBER). 1989. *NBER Reporter* (Fall).

Newey, Whitney K. 1985. Maximum likelihood specification testing and conditional moment tests. *Econometrica* 53 (September): 1047–70.

Rosenbaum, David E. 1990. Greenspan says nation's economy is in a downturn. *New York Times* (November 29): 1.

Runkle, David E. 1989. The U.S. economy in 1990 and 1991: Continued expansion likely. *Federal Reserve Bank of Minneapolis Quarterly Review* 13 (Fall): 19–26.

Todd, Richard M. 1984. Improving economic forecasting with Bayesian vector autoregression. *Federal Reserve Bank of Minneapolis Quarterly Review* 8 (Fall): 18–29.

Zarnowitz, Victor. 1969. The new ASA-NBER survey of forecasts by economic statisticians. *American Statistician* 23 (February): 12–16.

The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.