Using Monthly Data to Predict Quarterly Output

Robert Ingenito
and Bharat Trehan

Fiscal and Policy Analyst, California Legislature, and Research Officer, Federal Reserve Bank of San Francisco. Robert Ingenito was a Senior Research Associate at the FRBSF at the time this was written. The authors are grateful to John Judd, Joe Mattey, Brian Motley, and Glenn Rudebusch for their helpful comments.

Some time ago, the Commerce Department changed the way it calculates real Gross Domestic Product. In response to that change, this paper presents an update of a simple model that is used to predict the growth rate of current quarter real output based on available monthly data. After searching over a set containing more than 30 different variables, we find that a model that utilizes monthly data on consumption and nonfarm payroll employment to predict contemporaneous real GDP does best.

Although monetary policy actions are usually undertaken with a view to affecting the economy sometime in the future, policymakers are also interested in the current state of the economy. One reason is that estimates of the current state of the economy constitute the starting point for predictions of the future state of the economy. In addition, these estimates can also be used as an input for policy rules whose prescriptions are based on the current state of the economy.[1] Towards this end, a small model to predict current quarter real GDP growth was developed at the Federal Reserve Bank of San Francisco about ten years ago. This model has done reasonably well over this period. For instance, in Trehan (1992) it was shown that real time forecasts from this model outperformed the Blue Chip average forecast (though the sample period available for comparison was relatively short). In fact, the model has been incorporated into the forecasting process of one member of the panel of Blue Chip forecasters.[2]

In late 1995, the Commerce Department changed the methodology they use to calculate GDP, moving from the use of fixed weights to chain weights.[3] In this paper we discuss how the model—known as the monthly indicators model—has been modified in response to this change.

I. The Original Specification

The specification search for the original model was guided by the following considerations: We wanted a method to predict real GDP that did not involve any judgmental adjustments to the forecast; we also wanted the forecasts to be available relatively early in the quarter. The result of our search was a model that predicted current quarter real GDP based on knowledge of nonfarm payroll employment, industrial production, and real retail sales. These series have the virtue of being available monthly; an additional advantage is that all the data we need for a given month become available by the middle of the following month.

1. The rule recommended by Taylor (1993) is a well-known example.

2. See Laurence H. Meyer and Associates (1994).

3. See Motley (1992) for a discussion of some of the issues involved in the change.

Is there any reason to change the specification in response to the change in how real GDP is measured? At first glance, the answer appears to be no. When the original specification is estimated over the 1968–1995 period[4] using the new chain-weighted GDP we obtain the following equation:

$$\underset{}{RGDP_t} = \underset{(4.4)}{1.24} + \underset{(5.6)}{1.01}\,EMP_t + \underset{(3.5)}{.17}\,IP_t + \underset{(5.5)}{.19}\,SAL_t$$
$$- \underset{(-2.7)}{.16}\,RGDP_{t-1} - \underset{(-2.6)}{.15}\,RGDP_{t-2} - \underset{(-3.3)}{.18}\,RGDP_{t-3}$$

where the adjusted R^2 = .74, SEE = 1.98, and the *t*-statistics are shown in parentheses. *RGDP* is chain-weighted real GDP measured in 1992 dollars, *EMP* is nonfarm payroll employment, *IP* is industrial production, *SAL* is real retail sales; all variables are entered in growth rates. While these estimates are not too different from prior estimates where GDP was measured in constant 1987 dollars,[5] an examination of the forecasting performance of this equation over the last 10 years shows that it does not do particularly well over this period. Estimating the equation over the last ten years of the sample (actually over the period from 1985.Q1 to 1995.Q3) shows why:

$$RGDP_t = \underset{(2.3)}{1.08} + \underset{(3.1)}{1.04}\,EMP_t - \underset{(-0.1)}{.01}\,IP_t + \underset{(2.8)}{.15}\,SAL_t$$
$$- \underset{(-0.5)}{.07}\,RGDP_{t-1} + \underset{(0.1)}{.01}\,RGDP_{t-2} - \underset{(-1.5)}{.25}\,RGDP_{t-3}$$

where the adjusted R^2 = .52, SEE = 1.47. The *IP* variable is no longer significant over this period; the coefficients on the GDP lags are somewhat different as well, suggesting that the dynamics of the process may have changed. While the smaller sample can be expected to lead to larger standard errors, the change in the *IP* coefficient is harder to attribute to the small sample. To establish that this change was the result of the new GDP data, we estimated this equation over the same sample period (1985–1995) using GDP measured in 1987 dollars. We then obtained:

$$GDP87_t = \underset{(2.7)}{1.06} + \underset{(2.6)}{.70}\,EMP_t + \underset{(2.5)}{.22}\,IP_t + \underset{(2.9)}{.13}\,SAL_t$$
$$- \underset{(-1.6)}{.22}\,GDP87_{t-1} + \underset{(0.8)}{.09}\,GDP87_{t-2}$$
$$- \underset{(-0.7)}{.09}\,GDP87_{t-3}$$

where the adjusted R^2 = 0.63, SEE = 1.24. As can be seen, *IP* helps predict *GDP*87 over this sample (as it does when the equation is estimated over the entire 1968–1995 sample period).

These results suggest that we would be better off respecifying the monthly indicators model in response to the change in the GDP data. Our goals are the same as before: We would like a small model to forecast real GDP that does not involve judgmental adjustments. It would also be useful to obtain forecasts relatively early in the quarter.

II. SELECTION STRATEGY

One way to select the variables that will be used to forecast GDP is to rely on measures of in-sample performance. For instance, one could select the set of variables that maximizes R^2 in an equation that predicts real GDP or select those variables that have *t*-statistics above a certain value. However, specifications obtained in this way generally do not lead to good forecasts, since attempts to explain in-sample variation often lead to over-fitting. In other words, while the movements of a particular series within any given sample often can be explained by adding additional variables to the regression, relationships "discovered" in this way can sometimes be spurious and fail to hold up outside the sample under study. To minimize the possibility of such an outcome, we will use a strategy based on the results from two different search procedures. First, we use a procedure that selects a set of variables based on within-sample performance. Specifically, we use what Maddala (1977) calls the "Stepwise Regression Procedure" to determine an initial set of variables to be included in the model. Second, we select variables by looking at how well they help forecast real GDP out of sample.[6] In our final specification we place more weight on the second criterion, especially with regard to the number of variables included in the model.

Another set of issues involves the date at which we would like to make a forecast. The underlying issue is a familiar one: A forecast that is available relatively early in the quarter is likely to be less accurate than a forecast that is available later; yet it is possible to wait too long in an effort to get the most accurate forecast. As a practical matter, the date at which we would like to make a forecast will determine the set of variables that will be considered potential candidates for our model. All the variables included in our original specification were available by the 15th of

4. The start date is dictated by the availability of the retail sales data.

5. For instance, based on data through 1991, the coefficients in Trehan (1992) are: $GDP87_t = 1.1 + 0.96\,EMP_t + 0.20\,IP_t + 0.16\,SAL_t - 0.20\,GDP87_{t-1} - 0.10\,GDP87_{t-2} - 0.26\,GDP87_{t-3}$.

6. Here, too, one has to be careful to do more than simply choose the specification with the smallest prediction error; we provide the details of our procedure below.

the following month; for example, December data for all three variables in the original model are usually available by January 15. We have relaxed this constraint somewhat this time and will consider variables that are available by the end of the following month. This has the advantage of introducing variables such as personal income and consumption in the set of candidates that we consider.[7] However, it still excludes potentially important variables, such as inventory accumulation, which become available only about six to eight weeks after the end of the month in question.

III. Implementing the Strategy

We used the stepwise procedure to determine which of the 34 variables shown in Table 1 could be included in an equation that "explains" contemporaneous real GDP growth.[8] This list contains a representative of just about any kind of variable for which data become available within 30 days of the end of the relevant month. (For instance, while we did not consider every interest rate series available, we did make sure that we had both long and short maturities, as well as rates on private and government instruments, etc.) Our sample covers the 1967–1995 period, where the starting date is determined by the availability of the retail sales data. The procedure adds variables to the regression one at a time, choosing the one that has the highest partial correlation with output.[9] Only those variables whose *t*-statistic had a marginal significance level below 0.05 were included; further, if the introduction of a new variable caused a variable that was already in the regression to become insignificant at the 5% level, then the insignificant variable was dropped.

This procedure led to including the following variables in the equation: nonfarm payroll employment, average weekly hours, the number of passenger automobiles sold (the dollar value of which is included in retail sales), personal

TABLE 1

LIST OF VARIABLES CONSIDERED FOR INCLUSION IN MONTHLY INDICATORS MODEL

1. Federal Funds Rate
2. 3-Month Treasury Bill Rate
3. 6-Month Commercial Paper Rate
4. 1-Year Treasury Bond Rate
5. 10-Year Treasury Bond Rate
6. Moody's AAA Corporate Bond Rate
7. M2
8. Standard & Poors 500 Composite Stock Price Index
9 Loans and Leases at Commercial Banks
10. Index of Consumer Confidence (University of Michigan)
11. Index of Consumer Confidence (Conference Board)
12. New Privately Owned Housing Units Started, Total
13. The Consumer Price Index
14. Commodity Research Bureau Spot Market Index—All Commodities
15. Retail Sales deflated by the Consumer Price Index
16. National Association of Purchasing Managers' Index
17. New Orders for Durable Goods
18. Total New Passenger Cars Sold
19. Index of Industrial Production (Factories, Mines & Utilities)
20. Capacity Utilization, Manufacturing Sector
21. Real Personal Income
22. Real Consumption
23. Index of Leading Economic Indicators
24. Civilian Unemployment Rate
25. Total Employment (Household Survey)
26. Total Workers on Non-agricultural Payrolls (Establishment Survey)
27. Workers on Manufacturing Payrolls
28. Total Non-farm Payrolls Less Manufacturing Payrolls
29. Average Weekly Hours of Production Workers on Total, Private, Non-farm Payrolls
30. Index of Aggregate Weekly Hours, Production Workers on Non-farm Payrolls
31. Average Weekly Initial Claims for Unemployment Insurance
32. Diffusion Index: Percent of Firms Adding to Non-farm Payrolls (1-Month Span)
33. Gross Average Hourly Earnings, Constant Dollars
34. Gross Average Weekly Earnings

---

7. Note that for variables such as consumption the third month of data for any quarter will generally become available after the preliminary estimate of GDP is released. However, as our results below will demonstrate, the third month of data do not have a large effect on the accuracy of the current quarter forecast. More specifically, the model attains its lowest root mean square error before the preliminary GDP data are released.

8. All variables listed in Table 1 are available on the Citibase data tape.

9. An alternative strategy is to include all the variables we have in the regression and keep dropping variables that are insignificant; this approach is reminiscent of the "general to specific" approach recommended by Hendry. (See Hendry and Mizon 1978, for example.) However, following this procedure leads to including an extremely large set of variables in the model. We chose to follow a more conservative strategy here, for reasons we discuss below.

income, and consumption. Three lags of real GDP and a constant also were included in the equation.

Next, we used forecast performance over the 1985.Q1–1995.Q3 period to choose among alternative specifications. The procedure we used was as follows: For each forecast, the GDP equation is estimated up to the prior quarter, and the values of the indicator variables are used to predict the current quarter's output. For example, for the first forecast the equation is estimated through 1984.Q4 and used to forecast real GDP for 1985.Q1 using the contemporaneous values of the indicator variables. Next, the estimates are updated through 1985.Q1 and the equation is used to forecast 1985.Q2. The best specification is defined to be the one that leads to forecasts with the lowest root mean square error (RMSE) over the 1985.Q1–1995.Q3 period.[10] We carried out the search in two steps. We first searched for the set of variables that was the best at predicting real output, conditional on including a given number of variables in the set. We then varied the number of variables in the set, going from two to four.

In Table 2 we show the forecast error statistics for the best set of variables for the three different set sizes. For the two-variable case, the combination of total nonfarm payroll employment and real personal consumption leads to the smallest RMSE. The error falls slightly (from 1.40 to 1.31) when we move to the three-variable specification. The best specification here includes the two variables in the first set plus weekly hours. It turns out, however, that the third variable in the set does not matter very much; any of about a dozen variables when added to the first two lead to about the same size RMSE. Perhaps more to the point, only a dozen of the three-variable specifications actually perform better than the best two-variable case. Given that roughly 6,000 combinations were considered, it seems reasonable to attribute the slightly superior performance of 12 of these to chance. Hence, our conclusion is that the three-variable model does no better than the two-variable version.

We reach the same conclusion regarding the four-variable specification. The best specification there leads to a RMSE of 1.28 over this period; the variables included are real consumption, manufacturing payroll employment, nonmanufacturing payroll employment,[11] and a measure of commodity prices. Given that we looked at more than 46,000 combinations to find the lowest error, the small improvement we obtain does not appear to warrant rejecting the two-variable specification in favor of the four-variable one.

10. This means that the specification we choose could have a nonzero average forecast error.

11. These are the two components of nonfarm payroll employment, which is the variable that is selected in the first two specifications.

TABLE 2

PREDICTING CONTEMPORANEOUS OUTPUT: FORECAST PERFORMANCE OF ALTERNATIVE SPECIFICATIONS
SAMPLE PERIOD: 1985.Q1–1995.Q3

Best Specification	Mean Error	Mean Absolute Error	Root Mean Square Error
2 variables	–0.01	1.17	1.40
3 variables	–0.43	1.13	1.31
4 variables	–0.15	1.07	1.28

Of course, the same logic also can be used to question whether the two-variable specification is really any better than a specification that uses a single variable to forecast output. It turns out that among all the specifications that use only one indicator variable to predict output, the one that contains nonfarm payroll employment alone has the smallest RMSE: 1.67 percent. Thus, adding consumption to the equation that contains payroll employment leads to a reduction of about 0.3 percentage points in the RMSE.

IV. THE FINAL SPECIFICATION

Our preferred model is one that contains only two variables: nonfarm payroll employment and real consumption. The estimated equation is:

$$\begin{aligned} RGDP_t = {} & \underset{(0.1)}{0.05} + \underset{(10.7)}{1.41}\,EMP_t + \underset{(7.3)}{.51}\,CONS_t \\ & - \underset{(-3.1)}{.19}\,RGDP_{t-1} - \underset{(-3.4)}{.19}\,RGDP_{t-2} \\ & - \underset{(-4.2)}{.23}\,RGDP_{t-3}\,. \end{aligned}$$

RGDP is real GDP measured in chain-weighted 1992 dollars, *EMP* denotes nonfarm payroll employment, and *CONS* denotes real personal consumption; all variables are in growth rates. The equation is estimated over the 1968.Q2–1995.Q3 period. The adjusted R^2 is 0.71, and Godfrey's (1978) test reveals no evidence of either first or fourth order autocorrelation in the residuals.

The two indicator variables in our model are included in the set of variables selected by the stepwise procedure; they are also usually in the set of variables selected on the basis of our minimum RMSE criterion. A natural question here is: Are two variables enough to forecast contemporaneous output, or should we include additional variables? For instance, as discussed above, the stepwise regression procedure leads to the inclusion of a number of other variables in the set of variables used to forecast real output.

However, it seems to us that significant *t*-statistics alone are not sufficient to include a given variable in the model. This is especially the case because we have not selected the set of variables that we have searched over on any *a priori* basis, but have simply searched over (the relatively large set of) all available variables. Thus, there is a good chance that we will find variables that have large *t*-statistics but that are not really useful in predicting real output. In view of this, it seems desirable to opt for a relatively conservative specification.

V. PREDICTING THE INDICATOR VARIABLES

So far we have focused on how to predict output when we have all the monthly data we require available to us. However, most of the time, the model will be used to predict GDP when we have only partial data for the quarter. For instance, forecasts of Q4 real GDP made in late November or early December will be based on only one month of data on consumption and payroll employment and will require that we forecast how these variables will evolve over the following two months. In other words, in order to produce a model that predicts real GDP we need to produce an auxiliary model that generates forecasts of the indicator variables themselves.[12]

We begin by presenting the forecast errors that result from univariate models of both nonfarm payroll employment (which we denote by *EMP*) and real personal consumption (denoted by *CONS*). We regressed each variable on a constant and lags of itself, and then generated forecasts for the one-month-ahead to three-months-ahead horizon over the period from January 1985 to September 1995 (129 monthly forecasts). Once again, the forecast from each model was generated after estimating the model through the month prior to the (first) month being forecast; e.g., the forecast for 1988:01 was made after estimating the model from 1967:01–1987:12, while the forecast for 1992:7 was done after estimation through 1992:6. We tested several alternative specifications (by allowing the number of lags to vary) and concluded that using six lags results in the smallest forecast errors for both *EMP* and *CONS*. The forecast error statistics are presented in Table 3; *EMP* errors are in thousands of jobs per month, and *CONS* errors are in billions of dollars.

The first class of alternative specifications we examined combined the two variables into a vector autoregression (VAR). A VAR system models each variable (*EMP* and *CONS*) as a function of a constant and lags of both variables, where the univariate model employs its own lags only. Intuition suggests, for example, that previous changes in nonfarm payrolls contain information that might improve forecasts of consumption. However, forecasts from this specification were consistently worse than its univariate counterparts. For instance, the RMSEs for both variables when using the VAR were higher at all three forecasting horizons. Changing the lag lengths of the two specifications did not change this result.

As a robustness check, we augmented the bivariate VAR system by including real retail sales, various short and long-term interest rates, industrial production and the National Association of Purchasing Managers index. Not a single VAR specification including these variables (or any sub-group) improved upon the forecasting performance of the autoregressive specifications described above.

Our finding that the VARs do not forecast very well has been known for a while. Litterman (1986) suggested that the way to overcome this problem was to impose "Bayesian priors" on the VAR. The priors recommended by Litterman push each equation in the VAR towards a random walk. Specifically, in each equation, the coefficient on the first lag of the dependent variable is pushed towards one while all other lags are pushed to zero. How tightly these priors are imposed depends upon the forecasting performance of the model.[13] We tested several different Bayesian VARs (BVARs), including the bivariate case and several three- and four-variable systems. Imposing a Bayesian prior on the bivariate system produced slightly more accurate (lower RMSE) forecasts of both indicator variables than the univariate regimes, and it was also the most promising of all the BVAR specifications we tried.

Our final step was to see if the forecasts of the indicator variables could be improved by including contemporaneous values of those monthly series that are released before the indicator variables themselves. This is not a significant issue for the employment data, since that is one of the first releases that becomes available to us. However, consump-

12. We are following a two-stage strategy here: First, we search over the set of variables that leads to the best forecasts given that we have all the data we need for the quarter. Second, we try to find the best model to predict the indicator variables themselves. An alternative strategy is to integrate the two stages. This would allow us to compare the forecasting performance of alternative models at different points in the quarter (i.e., when we have partial data for the quarter we are trying to forecast). This latter approach was followed when the model was first estimated; the results obtained were not sufficiently different to justify the effort of the extensive search that would be required. Note that Figure 2 below does provide one comparison of this kind; it also provides a hint of why the more extensive search may not be very useful, since it shows that the model that does best based on three months of data also does best at every other point in the quarter.

13. For a more detailed discussion of how such a prior is imposed, see Todd (1984) or Litterman (1986).

tion data come out rather late, and it is natural to ask if consumption forecasts can be improved by taking account of other data already available to us. After some searching, we found that retail sales data—which are released roughly ten days to two weeks prior to the consumption data—are extremely useful in predicting contemporaneous consumption. The specification we finally settled on contains employment, consumption, and retail sales. The retail sales equation contains contemporaneous employment data, while the consumption equation contains contemporaneous employment as well as retail sales data; the inclusion of contemporaneous values reflects the order in which the data are released. Six lags of each variable are also included in each equation. Once again we have placed Bayesian priors on this system; the only exceptions are the contemporaneous terms, which have been left unrestricted.

The results from this exercise are contained in Table 4. In the first panel we show the forecast errors from an exercise where we assume we have no contemporaneous information when making our forecasts. By contrast, in the second panel we assume that we know the values of employment and retail sales during the first month that we are forecasting. Incorporating this information into the model cuts the RMSE of the *CONS* forecast for the first month by about a half; subsequent months are not affected as much, however. Note also that while the errors for employment in the second and third month seem to have declined noticeably, that is because they really represent one- and two-month-ahead forecasts.

Based on these results, our preferred specification for forecasting the indicator variables is the three-variable system

TABLE 3

Forecasts of *EMP* and *CONS* from Univariate Autoregressive Models (85:01–95:09)

	EMP			*CONS*		
Months Ahead	Mean Error	Mean Absolute Error	Root Mean Square Error	Mean Error	Mean Absolute Error	Root Mean Square Error
1	8.1	90.0	116.4	0.6	12.2	17.0
2	18.7	137.7	178.2	1.1	14.6	19.9
3	31.0	194.9	252.6	1.7	17.2	22.5

TABLE 4

Forecasts of *EMP* and *CONS* from a 3-variable System

	EMP			*CONS*		
Months Ahead	Mean Error	Mean Absolute Error	Root Mean Square Error	Mean Error	Mean Absolute Error	Root Mean Square Error
			Assuming no contemporaneous information			
1	–1.9	89.8	115.4	–1.2	12.1	17.0
2	–3.9	134.2	175.3	–2.0	14.2	19.4
3	–7.7	187.9	246.7	–3.2	15.7	21.7
			Assuming one month of information on *EMP* and retail sales			
1	NA	NA	NA	–3.6	7.6	9.4
2	–2.5	90.2	115.7	–3.9	13.9	18.5
3	–4.5	134.9	176.0	–4.9	15.3	21.5

where we include contemporaneous values of employment and retail sales in the equation for predicting consumption.

VI. MID-QUARTER OUTPUT FORECASTS

We are now in a position to analyze how the performance of the monthly indicator model would change as more and more information became available over the course of the quarter. It is easiest to understand how this works by means of a concrete example. Assume that we are in the second week of November and wish to generate an estimate of Q4 GDP. At this point we are likely to have employment data through October, but no Q4 data for either consumption or sales. Thus, we will use the monthly equations for predicting employment, consumption, and sales to fill out the remainder of the quarter. The quarterly averages of these (actual and forecasted) values can then be incorporated into the real output equation to estimate Q4 output. We can then repeat this exercise for every quarter of our forecast sample (1985–1995 again) and obtain a set of forecasts based on the same amount of information each quarter. Error statistics based on this set of forecasts are plotted as the point E1 in Figure 1. We show the mean absolute error and the root mean square error; the mean error stays close to zero throughout and therefore is not shown.

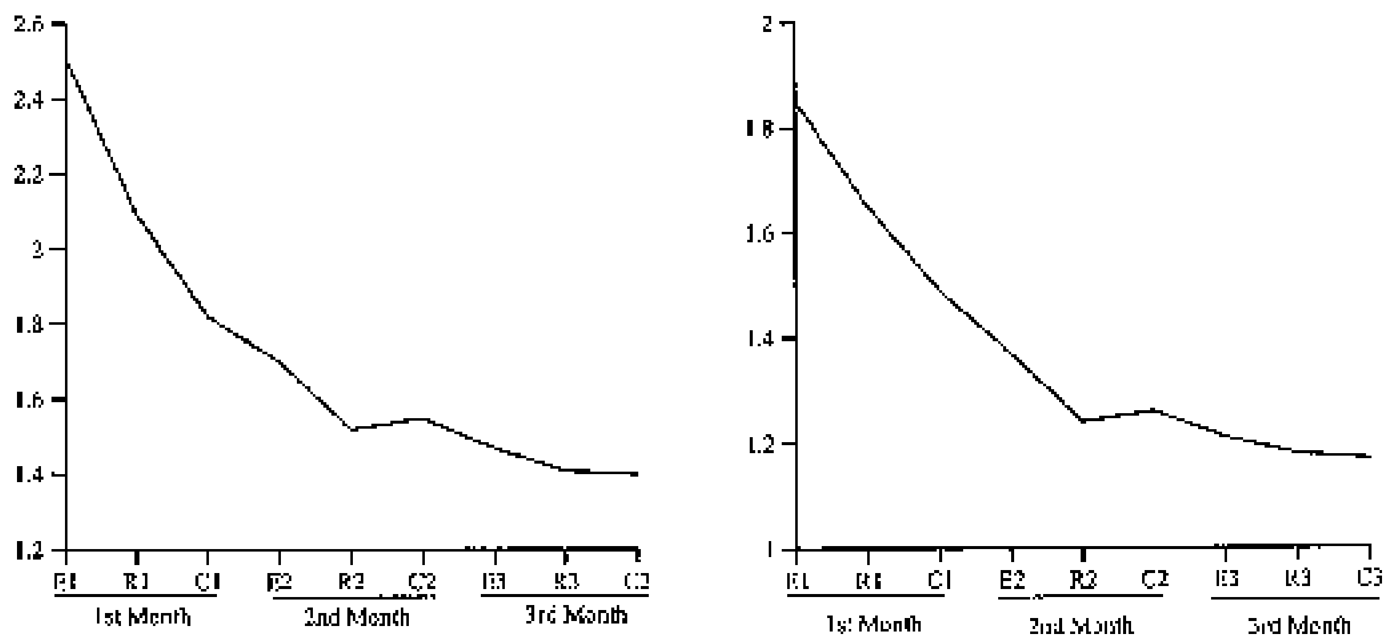
Successive points on the figure show how the performance of the model changes as more data become available. Thus, the point labeled R1 shows the forecasting performance of the model once the first month of retail sales data become available (that is, based on one month of employment and retail sales data), while the point labeled C2 shows the performance of the model once consumption data for the second month become available. The RMSE is 2.5% when employment data for the first month are received; it falls below 1.8% when we receive consumption data for the first month and is 1.5% based on complete data for the second month. The RMSE hits its minimum when we obtain retail sales data for the third month of the quarter.

Another issue has to do with the timeliness of the forecast. Use of consumption data in this version of the monthly indicators model means that a forecast based on complete data for the month will be available relatively late; for instance, if we had used retail sales we would have had a comparable forecast available about two weeks earlier. It is natural to wonder whether the new specification means that we will be worse off during the period between the re-

FIGURE 1

ERRORS IN PREDICTING REAL GDP AT VARIOUS DATES IN THE QUARTER

ROOT MEAN SQUARE ERROR | MEAN ABSOLUTE ERROR



Note: E1 is the date at which employment data for the first month become available; R2 is the date at which retail sales for the second month is released; and C3 is the date at which consumption data for the third month are published. Errors are measured in annualized growth rates.

ceipt of the retail sales data and the consumption data. Figure 2 provides an answer to this question. It compares the forecast errors from this specification to a specification where we use employment and retail sales to forecast output.[14] The figure shows that the errors from the specification that uses consumption to forecast output are never greater than those from the specification that uses retail sales. (Of course, both specifications also use employment.) The only time the RMSEs are close, for instance, is upon receipt of the first month of data on retail sales. Thus, this exercise does not suggest that the use of consumption instead of retail sales in the equation to predict real GDP leads to a less accurate forecast during the period in which we have retail sales data but do not have consumption data.

VII. SOME RELATED ISSUES

It is worth discussing two other issues before concluding this paper. The first one has to do with the use of initial versus revised data. All the results we have presented here have been based upon data as it existed at the time this project was first started (in early 1996).[15] It is not likely that we would obtain the same results using data that would actually be available to us in real time. Unfortunately, since the required data are not available to us, it is not possible to determine how the model would perform under these circumstances. (Recall also that the chain-weighted GDP data are new.) However, it is possible to get some sense of how the error statistics might change with data revisions by looking at the historical performance of the *original* model (with GDP measured in 1987 dollars). Over this small sample of 16 forecasts, we obtain a RMSE of 1.1% when values of the monthly variables as they exist today are used to forecast GDP87 data as they exist today. The real time forecast error is the same, that is, when the forecasts that the model actually made over this period are

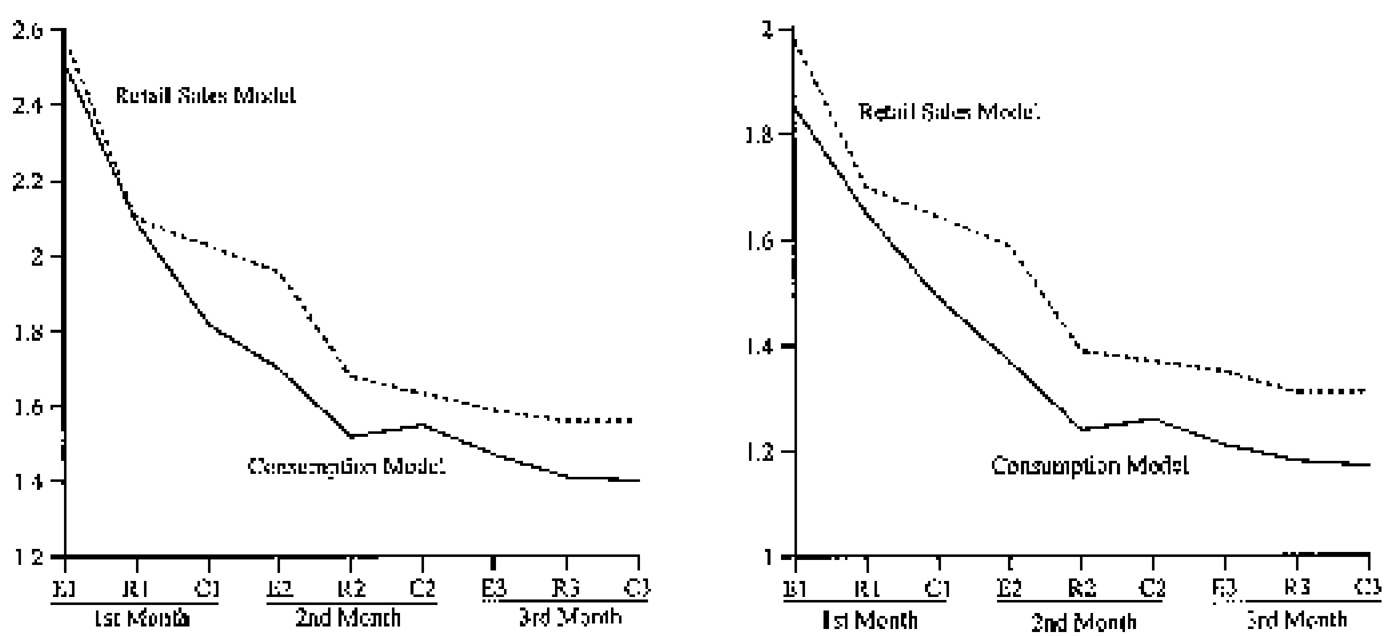
14. Note that this requires monthly forecasts of retail sales; these forecasts are obtained from the same 3-variable system that is used to forecast employment and consumption.

15. The monthly consumption data are current as of June 1996.

FIGURE 2

COMPARISON OF FORECAST ERRORS

ROOT MEAN SQUARE ERROR | MEAN ABSOLUTE ERROR



Note: See Figure 1.

compared to the initial estimates of GDP87 the RMSE is 1.1% as well. However, when the model's historical forecasts are compared to currently available GDP87 data, the RMSE is 1.6%.

A final issue has to do with the stability of the estimated equation. It is well known that estimated macroeconomic relations shift over time, and it is quite possible that the coefficients of our estimated equation will change as well. This suggests that it might be better to forecast using a specification based on time-varying coefficients. We tried a number of such specifications, including several that assume that the coefficients follow a random walk and others that assume that the coefficients can move around but tend to return to some fixed value. We were unable to come up with a specification that outperformed the fixed coefficient version.[16] In fact, the only time we were able to get RMSEs smaller than those from the base version (which involves Kalman filtering) was when we set the coefficients equal to their final period value at the beginning of the forecast period and held them there throughout.

VIII. SUMMARY AND CONCLUSIONS

In this paper we have presented a revised version of a small model that is used to forecast current quarter GDP. We have shown that a specification based on two indicator variables does about as well at forecasting GDP as specifications that contain three or four variables. In addition, we have searched over a larger set of indicator variables this time, allowing for variables that are available up to one month after the month to which the data pertain. As a result, we found that monthly consumption data provide key information about contemporaneous output. There is a potential trade-off here: While forecasts based on the consumption data are more accurate, we have to wait longer to get the relevant consumption data. So there is a period of time when a model based on consumption could make forecasts that are worse than a model that does not contain consumption (because the latter model will have more current information over this period). It seems that we do not have to pay such a price, because retail sales data help forecast consumption.

16. This finding is consistent with what Stock and Watson (1996) find. They analyze bivariate regressions based on a data set of 76 monthly series (5,700 relationships) and conclude that "...in over half the pairs, random walk TVP models or rolling regressions perform better than fixed coefficient or recursive least squares, although the gains typically are small." In other words, time-varying parameter models do no better than fixed coefficient models about half the time.

REFERENCES

Godfrey, L.G. 1978. "Testing against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables." *Econometrica* 4, pp. 1293–1302.

Hendry, D.F., and G. E. Mizon. 1978. "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study by the Bank of England." *Economic Journal* 88, pp. 549–563.

Laurence H. Meyer & Associates. 1994. *The U.S. Economic Outlook.* September 7.

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

Maddala, G.S. 1977. *Econometrics.* New York: McGraw Hill.

Motley, Brian. 1992. "Index Numbers and the Measurement of Real GDP." Federal Reserve Bank of San Francisco *Economic Review* 1, pp. 3–14.

Stock, James H., and Mark W. Watson. 1996. "Evidence on Structural Instability in Macro Time Series Relations." *Journal of Business and Economic Statistics* (January) pp. 11–30.

Taylor, John B. 1993. "Discretion Versus Policy Rules in Practice." *Carnegie–Rochester Conference Series on Public Policy* 39, pp. 195–214.

Todd, Richard M. 1984. "Improving Economic Forecasting with Bayesian Vector Autoregression." Federal Reserve Bank of Minneapolis *Quarterly Review* (Fall) pp. 18–29.

Trehan, Bharat. 1992. "Predicting Contemporaneous Output." Federal Reserve Bank of San Francisco *Economic Review* 2, pp. 3–12.