

# Real-Time Forecasting in Practice

THE U.S. TREASURY STAFF'S REAL-TIME GDP FORECAST SYSTEM

By John Kitchen and Ralph Monaco



*John Kitchen is the Chief Economist of the House Budget Committee. Previously, he has held positions in the U.S. Department of the Treasury, the Council of Economic Advisers, the U.S. Department of Agriculture, and Washington and Jefferson College. Kitchen received his*

*Ph.D. and MA in economics from the University of Pittsburgh and his AB in economics and history from the College of William and Mary.*



*Ralph Monaco is an economist with the Office of Economic Policy, U.S. Department of the Treasury. Previously, he was President of the I.E.R.F., Inc., a not-for-profit economic research corporation at the University of Maryland in College Park. He has been a contributor to the Blue Chip*

*Economic Indicators panel and has held positions in the Department of Agriculture, the Central Intelligence Agency, and the Council of Economic Advisers. Dr. Monaco received his Ph.D. from the University of Maryland in 1984.*

*This paper outlines a method for making effective use of monthly indicators to develop a current-quarter GDP forecast. Estimates and projections of real GDP growth are usually used to describe how the economy is doing. But estimates of GDP are only available quarterly, and the first GDP estimate for a quarter is released late in the month following the end of the quarter. The lack of a timely, comprehensive economic picture may mean that policymakers and business planners may be as much as four months behind in recognizing a significant slowdown or acceleration in the economy. This problem is especially important around business cycle peaks or troughs, where there may be some evidence that the economy is changing direction.*

*There are many less-comprehensive, but higher-frequency data series about the economy, however. The chief difficulty with using the multiple indicators is that different indicators can give different signals, and there is no agreed-upon way for aggregating the statistics to give a single-valued answer.*

*In this paper, we describe the approach we have adopted at the Treasury Department to use a broad variety of high-frequency incoming data to construct “real-time” estimates of quarterly real GDP growth. We draw on the recent work by Stock and Watson and others and describe the indicators, the techniques, and the recent performance of the system.*

This paper won the Edmund A. Mennis Contributed Paper Award, presented at the NABE Annual Meeting, September 15, 2003.

The views expressed in this paper are the authors' and do not necessarily reflect those of the U.S. Department of the Treasury or the House Budget Committee.

**P**olicymakers and economists often turn to real GDP growth to assess how the economy is performing since GDP is a well-known, comprehensive measure that covers the economy as a whole, rather than a single sector or market (however broad) like manufacturing or employment. Further, GDP growth is a key variable in many policy or strategy analyses. For example, at the Treasury Department, projections of nominal GDP serve as the basis for projections of taxable incomes and the resulting projections of government tax receipts. The GDP metric enjoys wide recognition and usage as the nation's economic barometer.

In practice, however, a significant problem arises because GDP estimates are not timely enough for many needs. GDP estimates are only available quarterly, and the Bureau of Economic Analysis' (BEA) first estimate of GDP for a quarter is released late in the month following the end of the quarter. For policymakers as well as business planners, the lack of a timely, comprehensive picture of the economy can present a critical problem: they may be as much as four months behind in recognizing a significant slowdown or acceleration in the economy. Such a lag in timely information is an important part of the "recognition lag" that economists have identified as a major impediment to the successful implementation of discretionary counter-cyclical policies. This problem is especially important around business cycle peaks or troughs, where there may be some evidence that the economy is changing direction.<sup>1</sup>

At the other extreme, high-frequency data about specific industries and markets abound and these can be—and are—often used to inform judgments about the economy's current performance or where it may be headed. A wide variety of monthly, weekly, daily—even intraday—data are available to alert policymakers and analysts to changes in the course of the economy. Various problems arise in attempting to use such high-frequency data, however. Such data are inherently "noisy," and it is often difficult to identify the underlying "signal" information. In addition, the very multiplicity of the data—although providing additional potential sources of information—presents a problem for analysts. The chief difficulty with using multiple indicators is that they can, and usually do, provide conflicting signals; and there is no agreed-upon way for aggregating the statistics to give a single-valued answer. For example, it is difficult to decide

how to "add up" the Bureau of the Census' housing starts and the Institute of Supply Management Purchasing Managers Index (PMI) to give a single-valued answer. Without some way to aggregate these pieces into a consistent picture, it is often difficult to separate the signal in the statistics from the very short-run noise.

In this paper, we describe the approach we have adopted at the U.S. Treasury to use the broad variety of incoming data to construct "real-time" estimates of quarterly real GDP growth. For us, "real time" refers to the effort to conduct continuous, contemporaneous analyses of incoming information that allow forecasters to make continual and instantaneous updates to their forecasts as new data become available. The real time forecasting system (RTFS) is the result of our efforts to produce a fluid, data-based forecast of contemporaneous real GDP growth that is subject to continual updating the instant new data become available.

An alternative strand of the "real-time" data and analysis literature has received much attention in recent years in economics research. This research has focused on the important issue of how using contemporaneous "vintage" data for historical sample periods in empirical analysis can yield different estimation results than the "last-available vintage" data that are typically used (see, for example, Croushore and Stark (1999, 2001) and Orphanides (2001)).<sup>2</sup> The sensitivity of observed historical relationships to the vintage of the data used potentially can affect the ability to forecast (e.g., Dynan and Elmendorf (2001), Koenig, Dolmas, and Piger (2001)). While we believe this is an important issue, in this paper we focus on the rationale for a coherent forecasting system and our efforts to construct one. Now that the system is largely in place, we look forward to conducting future research that examines the sensitivity of near-term projections to the use of alternative data vintages for estimating the system.

### **Approaches to Forecasting Current-Quarter GDP Growth**

In practice, several general approaches typically have been used to predict current-quarter economic performance and real GDP growth in particular. In the main, these approaches tend to follow the design of the GDP accounts themselves. That is, they use available indicators to develop forecasts of GDP components, and then aggregate the components to make a guess at a quarter's GDP. These are: (1) monthly GDP and real activity measures, (2) quar-

<sup>1</sup>Dynan and Elmendorf (2002) address the issue of the accuracy of initial GDP estimates at turning points in the business cycle and the potential role for additional data to improve the ability to forecast turning points.

<sup>2</sup>Data "vintage" reflects the data and information set that was available at a particular point in time. Later data vintages have revised data for that period estimated at later points in time.

terly GDP accounting, and (3) model-based estimates.<sup>3</sup>

*Monthly GDP and real activity measures.* Some analysts construct monthly analogs of GDP, tracking the available monthly data that enter into the calculation of quarterly GDP or are historically related to it, and then simply reporting the resulting monthly series. Of course, some data that go into GDP are not available on a timely basis—inventories and foreign trade statistics are leading examples—and some parts of GDP are imputed by the BEA. As a result, complete monthly GDP estimates for any given quarter will lag as much as two months and will then only imperfectly capture the BEA's GDP estimate for a given quarter. The monthly data for GDP components, when aggregated to a quarterly frequency, provide, in a sense, a *forecast* of the official quarterly GDP figures.<sup>4</sup>

Other monthly measures of general economic activity already in use—aggregates of a variety of indicators—are essentially one step removed from a forecast of quarterly GDP. For example, business cycle indicators such as the Conference Board's indexes of leading, coincident, and lagging indicators are essentially weighted summations of key monthly indicators.<sup>5</sup> The four components of the index of coincident indicators, for example, are used by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER) to help determine the month when a business cycle turning point occurs. Another example of a monthly indicator of the contemporaneous performance of the economy is the Federal Reserve Bank of Chicago's National Activity Index, a weighted average of 85 indicators.<sup>6</sup> These monthly indicators are designed to illustrate the relative performance of the economy—whether it is expanding or contracting on a monthly basis. They are, however, typically published once a month, and because of the data series included, often with a lag of a month or with extrapolated or estimated components, potentially are subject to considerable revision. To use them to create an explicit quarterly real GDP growth forecast, the analyst needs to take the final step of relating movements in these monthly indicators to movements in quarterly GDP growth.

<sup>3</sup>In practice, most forecasters use all three approaches to one degree or another. Treasury staff use a wide variety of approaches to estimate current-quarter real GDP growth. The RTFS is one of several approaches used.

<sup>4</sup>Recently, the Business Cycle Dating Committee of the National Bureau of Economic Research cited the use of monthly real GDP estimates prepared by Macroeconomic Advisers as an information variable for their efforts to identify turning points in economic activity.

<sup>5</sup>For the Conference Board business cycle indicators, web address: <http://www.globalindicators.org>

<sup>6</sup>Web address: <http://www.chicagofed.org/economicresearchanddata/data/index.cfm>

*Quarterly GDP accounting.* A related, and often labor-intensive, approach tracks the data that go into the GDP calculation as they appear, and then estimates the expected quarterly growth in real GDP by filling in assumptions about what will happen to these data in the remaining months of the quarter. For example, data on retail sales are released in the middle of the month. The retail sales “control” component goes directly into the calculation of GDP in a quarter (and affects real GDP after suitable deflation). If only one month of data is available, it often is assumed that the remaining two months of data for the quarter will be unchanged or continuing on a recent trend. Using the emerging data on major components and similar assumptions, analysts can develop a moving forecast of the current quarter's GDP growth. Presumably, as more data for the quarter are released and the importance of assumed data diminishes, the forecast of GDP growth becomes closer to the figure that is subsequently released by BEA.

*Model-based estimates.* Quarterly macroeconomic models provide another approach to developing a view of current-quarter real GDP growth. In general, as data for the quarter are released, model users revise their “add factors” for the various components of GDP. For example, as data on industrial production and manufacturers' shipments are released, it may become apparent that the model's equation for business equipment spending is predicting too much activity relative to the industrial production indicator. In this case, model users would usually reduce their estimate of business equipment spending for the quarter—usually by adjusting the add-factor (in this case a negative adjustment) to their estimate of current quarter business investment growth. The model would then combine this new estimate of business investment spending with forecasts of other components to arrive at a view of current-quarter GDP. It is useful to think about the macroeconomic model as a machine-based generalization of the forecasting approach in which estimated economic relationships—like the response of consumer spending to interest rates—are used to help hone the forecasts for the implied missing months of data.

## Regression-Based Approaches

In contrast to the GDP-accounts approaches, another approach is to forecast GDP growth directly from available indicators by using historical statistical relationships, usually with regression equations of the form:

$$(1) y_t = \alpha + \beta(L)x_t = e_t$$

where  $y_t$  is the percentage change in real GDP at an annu-

al rate for quarter  $t$ ,  $x_t$  is an indicator variable that is related to GDP growth,  $\beta(L)$  is a set of coefficients for current and lagged values of the indicator variable, and  $e_t$  is an error term with the assumed typical properties.

One well-known example of this approach uses the monthly Bureau of Labor Statistics series on aggregate hours of nonfarm production workers to predict GDP growth. In the simplest version of this relationship, a regression like equation (1) is run relating GDP growth to a constant term and the contemporaneous quarterly movement in aggregate hours (e.g., for  $\beta(L)$ ,  $L = 0$ ). For the sample period 1964.2 to 2002.2, results for the simple regression are:

$$(2) \quad \% \Delta \text{GDP} = 1.76 + 0.81 \% \Delta \text{Hours}$$

(0.21) (0.06)

DW = 2.02; Adj. R<sup>2</sup> = 0.575; SE = 2.36.

where the standard errors of the coefficient estimates are reported in parentheses beneath the coefficients. As it turns out, the regression coefficient on aggregate hours growth is typically somewhat less than unity, at least partially because GDP includes activity in farm business, households and institutions, and governments; and these are likely to be less volatile than nonfarm business labor input on a quarter-to-quarter basis. A slightly more sophisticated equation would recognize the importance of the changing behavior of underlying labor productivity growth, as illustrated by the results from a regression that splits the constant coefficient term into three alternative historical periods:

$$(3) \quad \% \Delta \text{GDP} = 2.21(\text{Pre-1974}) + 1.50(1974-1995) + 2.03(\text{Post-1995}) + 0.81 \% \Delta \text{Hours}$$

(0.40) (0.27) (0.47) (0.56)

DW = 2.06; Adj. R<sup>2</sup> = 0.578; SE = 2.36.

The results show the oft-cited shifts in productivity growth, with higher productivity growth prior to 1974, lower productivity growth in the post-1973 period until 1995, and a return to higher productivity growth in the post-1995 period.<sup>7</sup>

To use the relationship, the analysts would plug in their view of quarterly hours growth based on the BLS data releases, and then use the estimated regression to predict GDP growth. For example, if nonfarm hours were known

or were expected to have risen by 1-1/2 percent at an annual rate in the current quarter, equation (3) would suggest that an estimate for current-quarter real GDP growth would be equal to: 2.03 + 0.81\*1.5% = 3.25 percent.

Alternatively, a regression approach that uses multiple indicators as explanatory variables also could be considered:

$$(4) \quad y_t = \alpha + \beta_1(L)x_{1,t} + \beta_2(L)x_{2,t} \dots + \beta_n(L)x_{n,t} = e_t$$

where the  $x_{1,t}$ ,  $x_{2,t}$  ... ,  $x_{n,t}$  variables are indicators chosen for their predictive power for explaining current-quarter real GDP growth. A good example of this approach is Ingenito and Trehan (1996), who develop a GDP forecasting equation that ultimately relies on monthly employment and consumption data. This particular equation was the result of a winnowing process that began from 34 candidate variables. Ingenito and Trehan then tested alternative combinations of variables in forecasting regressions, choosing the one that minimized the root mean square error of a series of one-step ahead forecasts made over the 1985.Q1 through 1995.Q3 period. Having chosen a single equation with two independent variables, it remained to create forecasts of the quarterly values for the independent variables during the quarter for which a forecast of GDP growth was required. (It is worth noting here that consumption for the third month of the quarter is usually released after the first estimate of GDP for the quarter.) This was done by using other monthly data to predict the indicator variables via another set of regression equations. In essence, Ingenito and Trehan follow a two-step approach, which is likely followed by many other analysts: (1) find a parsimonious regression equation relating GDP growth to a small set of available monthly indicators, and (2) find a forecasting relationship to forecast the chosen monthly indicator variables when their full-quarter values are not yet known.

Ingenito and Trehan attempt to deal with the problems inherent in the use of equation (4) in a GDP forecasting context. First, there are a number of candidate variables, and this number is potentially large relative to the sample size of GDP growth rates. For example, many analysts have their "favorite" variable that they use to forecast real GDP growth. Competitor variables include the Conference Board's leading index, their coincident index, interest rate spreads, real M2 growth, commodity price growth, a variety of employment indicators, industrial production, and retail trade indicators. Including too many variables in an equation like (4) results in over-fitting and poor forecasting performance. Yet, at the same time, each probably does contain some unique information about GDP growth,

<sup>7</sup>The breakpoints used were 1973.4 and 1995.4. Note that an alternative formulation could be based on breakpoints at business cycle peaks.

which is likely what recommended that variable to the analyst in the first place.

Second, as with any regression, there are two main sources of error. One is the usual error in the regression itself, because the relationship between GDP growth and hours in equations, such as equations (1) through (3), is estimated and is not an identity. A second is the error that arises when we have to substitute a forecast of quarterly hours growth for actual quarterly hours growth. This must be done in the two months before we have complete hours growth for a quarter. Whatever error is embodied in our function for making up the missing hours data is translated into an error in the GDP growth forecast. Ingenito and Trehan make forecasts of the indicator variables using “second stage” estimates relating other monthly series to the indicator variables chosen for the GDP regression.

The work of Ingenito and Trehan highlights the issues that arise in using regression-based methods for forecasting current quarter GDP, but the work is focused on a single forecasting equation. Klein, with various coauthors, has laid out a comprehensive, systematic view of forecasting current quarter GDP using regression methods (the Current Quarter Model). Klein and Sojo (1989) describe a regression-based current quarter GDP forecasting system in which GDP components are modeled individually. The “bridge” equations relate quarterly values of closely relevant monthly series to the component of GDP being modeled—for example national accounts consumer spending on gasoline and oil are regressed on Census retail trade spending on gasoline service stations—and these components are then aggregated to form a forecast of GDP. Forecasts of each monthly indicator’s data using time-series techniques fill out the quarter’s data when no actual data for the monthly indicator is available. To continue the example above, time series forecasts of the retail trade series spending on gasoline service stations for say, February and March, would be combined with actual data for January to create a first quarter observation, which would itself be plugged into the quarterly regression equation to forecast national accounts consumer spending on gasoline and oil. The procedure is carried out separately for data on the expenditure and income sides of the national accounts. Current quarter forecasts are reported for real GDP, nominal GDP, and the GDP deflator, as well as for selected sub-aggregates of the expenditure side of the national accounts.

In addition to the work by more detailed component, Klein’s Current Quarter Model uses single-equation regression equations to predict real GDP, nominal GDP, and the GDP deflator. The independent variable in the three regressions is the first principal component of about

25 monthly indicators, aggregated to a quarterly frequency. As described above, forecasts of monthly observations for which actual data are not yet available are developed from time series methods.<sup>8</sup>

### A Generalization of the Indicator Approach

The approach that we developed to generate “real-time” current-quarter GDP growth forecasts is a generalization of the indicator approach and is conceptually similar to Klein’s principal components approach. Like the principal components approach, rather than try to find the “best” set of a small number of regressors in a single equation (as in equation (4) above) or relying on a single indicator, like aggregate worker hours as in equations (1) through (3) above, we use a wide variety of indicators and estimated relationships to predict GDP growth in a quarter. Each of these relationships produces an individual-indicator GDP forecast. We then aggregate these individual-indicator forecasts to give us an overall forecast of GDP growth. We have experimented and continue to experiment with different aggregation schemes to produce the overall GDP forecast. This approach is “real-time” in the sense that the estimate for current-quarter GDP growth can be updated instantaneously as new data in the broad set of indicators become available individually.

RTFS first attempts to quantify the past relationships between each monthly indicator and GDP growth for a quarter. The general form of the specifications is:

$$(5) y_t = \alpha + \beta(L)x_t + \lambda(L)y_{t-1} + e_t$$

In practice we have found that the lagged GDP growth does not help to forecast current quarter GDP growth when another variable is present so this term is dropped from (5), although, in practice, it need not be. Independent variables could include many lags of the indicator variable. In practice, we are currently allowing each regression to contain either no lags or four lags, and we use the Schwarz criteria to determine which results to use in the RTFS.

Because of the timing of release of monthly data and their relationship to the quarterly GDP variable, the ongoing process of estimating and updating the forecast system is intuitively simple but somewhat complex in application and in the computer code used to implement it. Each indicator has three different equations that relate it to GDP. These correspond to periods in each quarter for which

<sup>8</sup>Up-to-date reports from the Current Quarter Model approach are available weekly from the Project Link website. Use <http://www.chass.utoronto.ca/link/uscqm/uscqm.htm> to link to the current CQM report.

one, two, or three months of data for the indicator are available. As the forecast quarter unfolds and data on the monthly indicators are announced, the RTFS selects the appropriate forecasting equation for each indicator based on how many months of data are available and uses it to generate the indicator-specific forecasts for GDP growth:

$$(6) \begin{aligned} \hat{y}_{1,t} &= \hat{\alpha}_1 + \hat{\beta}_1(L)x_{1,t} + \hat{\lambda}_1 y_{1,t-1} \\ \hat{y}_{2,t} &= \hat{\alpha}_2 + \hat{\beta}_2(L)x_{2,t} + \hat{\lambda}_2 y_{2,t-1} \\ &\vdots \\ \hat{y}_{n,t} &= \hat{\alpha}_n + \hat{\beta}_n(L)x_{n,t} + \hat{\lambda}_n y_{n,t-1} \end{aligned}$$

for indicators  $x_1, x_2, \dots, x_n$ . The RTFS then combines the results across indicators, weighting the individual forecasts by the strength of the estimated past relationship to GDP, specifically, the estimated  $R^2$  values of the fitted historical equations:

$$(7) \hat{y}_{s,t} = \sum \omega_{i,t} \hat{y}_{i,t} \text{ where } \omega_{i,t} = \frac{R_{1,t}^2}{\sum R_{j,t}^2}$$

where the subscript S represents a value for the “system” and the summations across the  $i$  and  $j$  in each case are across all values for indicator variables  $x_1, x_2, \dots, x_n$ . The system estimate  $\hat{y}_{s,t}$  gives a projected value of GDP growth for the quarter that is updated as new data on the indicators  $x_1, x_2, \dots, x_n$ , are released. Note that the weights  $\omega_{i,t}$  are time period dependent, evolving each period as the data used to estimate the forecasting equations evolve. In fact, the weights even evolve intra-period, as new data on indicators become available within the monthly time period.

The general approach used here is similar to that adopted by Stock and Watson (1999, pp. 314ff) in their examination of alternative estimation procedures for producing forecasts of inflation. Stock and Watson considered alternative combination procedures for combining individual bivariate forecasts of inflation, including a simple mean, a simple median, and ridge regression techniques. In our approach, we have chosen to use a weighting scheme based on relative in-sample explanatory power.

### Implementing Treasury’s Real-Time Forecasting System

An explicit discussion of the process underlying the RTFS will help illustrate its operation. In the first month of a quarter, the only “real time” data that are available are the daily observations on financial market variables such as interest rates, interest rate spreads, and stock prices and index levels.<sup>9</sup> At the end of the first month, the first full month’s values for those variables become available. The system then could be run to generate a forecast of the cur-

rent-quarter real GDP growth based on the financial market variables alone. The system would take the first month’s values for the financial variables and, for each indicator, select the historical bivariate estimation equation that estimates the relationship between real GDP growth and the historical data for the explanatory indicator for the first month of the quarter. As subsequent indicator variable data become available for the first month—e.g., the PMI, inventory indexes, payroll jobs, etc.—the system can then be run to get updated estimates for the quarter based on a broader coverage of indicators. After the second month evolves, a mix of data across indicators by month will become available. For example, some indicators will have two months of data available while others have only one month available. As the quarter evolves, data on explanatory indicators become available at different times; and the RTFS always (1) chooses the most recent data available by indicator, (2) estimates the proper historical equation based on the months of data available by indicator, (3) produces forecasts by indicator, and then (4) combines the individual forecasts according to their historical explanatory power to generate a single GDP forecast. A key difference between the RTFS approach and other approaches is that the regressions used in the RTFS do not use forecasts of the independent variables. Rather the RTFS equations are designed explicitly to capture the statistical relationships when different numbers of months of actual data are available for a quarter.

In practice, we are using a data horizon that extends over the 80 quarters previous to the quarter to be forecast. The RTFS is implemented through code written for use in EViews econometric software with online access to comprehensive data bases from Haver Analytics. At this time, the RTFS uses 30 indicators; it could easily be expanded to include a broader array of indicators if we chose to do so—and that may yet happen.

### Testing the Real Time Forecasting System

To test the potential efficacy of the RTFS, we performed a sequence of one-step-ahead forecasts starting in the first quarter of 1995 and running through the first quarter of 2003. To be specific, we made three forecasting “rounds” to make a forecast of the first quarter of 1995: we call the three rounds “early,” “middle,” and “late,” reflecting the monthly data pattern during the quarter. In our “early” estimates for the quarter, we forecasted with regressions that had been estimated using current quarter

<sup>9</sup>The system could be run with preliminary, partial values of the financial market variables, e.g., average values to date during the month, but typically, the system is run when the full month’s values are available, e.g., consistent with the observations in the historical series.

estimates of the monthly indicator, which used only one month of data and used only January 1995 data to predict the contemporaneous current quarter GDP. In our “middle” and “late” estimates, we forecasted with regressions using two and three months, respectively, of data to predict current quarter GDP. The dataset included current vintage versions of both the monthly indicators and real GDP growth, and equations were estimated using data beginning in

Table 1 shows several interesting features. First, there is a small, but noticeable improvement in all of the average error measures as we move through the quarter and include more monthly information as well as use regressions that have been estimated using more monthly data for the contemporaneous quarter. The RMSQEs fall from 1.9 percentage points on GDP growth using only early estimates to 1.7 percentage points using late estimates (three months of monthly data available in the quarter). Second, there seems to be little difference between the different aggregation schemes in terms of average forecasting performance. Further, the simple mean, median, and weighted average estimates for early, middle, and late estimates are quite close. There seems to be no tendency for the any of the estimates to systematically over- or under-predict; the median errors are quite close to zero, especially for estimates based on two and three months of data. The distribution of errors seems to be somewhat asymmetric (skewed slightly toward the positive side) because third quartile errors are larger in absolute value than first quartile errors.

As a comparison, we developed a series of one-step ahead forecasts based on four lags of real GDP growth. For example, starting in 1995Q1, we predicted quarterly GDP growth using a regression with data from 1982Q1 through 1994Q4 and four lags of GDP growth. We then added 1995Q1 to the historical data set and forecast 1995Q2, and so on up to the first quarter of 2003. The table shows that the lagged GDP growth

model has a higher RMSQE—by about 0.5 percentage points—than any of the real-time system aggregates.

It is worth noting that even these early, middle, and late breakdowns are a gross simplification of the data stream. As we use the system on a daily basis, the composition of the GDP prediction is actually an amalgam of early, middle, and late projections. One of the advantages of the RTFS approach is that it can capture the continuous flow of information accounting not only for the updated data, but also by using a regression that has been fit using actual data from previous periods with the same relative information content.

Table 2 shows RMSQEs by indicator for the early, middle, and late quarter estimates. Although all of the RMSQEs are quite close, the regression using the growth in the Conference Board’s Leading index has the lowest

TABLE 1				
ERROR MEASURES FOR ONE-STEP AHEAD REAL GDP GROWTH PREDICTIONS				
1995Q1-2003Q1, percentage points				
	Root Mean Squared Error	First Quartile	Median	Third Quartile
<b>One Month of Quarter's Data</b>				
Mean	1.88	-1.30	-0.27	1.78
Median	1.86	-1.46	-0.16	1.73
RBarSq-weighted average	1.85	-1.18	-0.35	1.68
<b>Two Months of Quarter's Data</b>				
Mean	1.78	-1.20	-0.10	1.64
Median	1.79	-1.41	0.06	1.57
RBarSq-weighted average	1.75	-1.15	-0.15	1.69
<b>Three Months of Quarter's Data</b>				
Mean	1.73	-1.25	-0.11	1.62
Median	1.73	-1.43	-0.05	1.51
RBarSq-weighted average	1.69	-1.24	-0.08	1.66
<b>Lagged GDP Growth Model</b>	2.46	-1.99	-0.12	2.22

Column 1 shows the root mean squared error of the one-step-ahead predictions; other entries show summary statistics of the error series by number of months of data included and type of aggregation.

1982Q1. We forecasted real GDP three times for each variable (early, middle, late) for each of the 30 variables for each quarter from 1995.Q1 through 2003.Q1. Then we examined the forecasting performance of the early, middle, and late month projections. Those results are shown in Table 1.

We also examined three aggregates of the 30 indicators. Our favored aggregate is the R2 weighted aggregation, but we also report the simple mean of the forecasts and the median. For each of these summaries, we report the root-mean-squared error, and a three-number summary of the error distribution.

Column 1 shows the root mean squared error of the one-step-ahead predictions; other entries show summary statistics of the error series by number of months of data included and type of aggregation.

RMSQE of all the indicators. The regression based on unemployment insurance claims shows the most improvement as data are added for a quarter, declining from 1.98 percent to 1.67 percent. It is worth emphasizing that these regressions are current-vintage regressions and thus do not account for the “real time” aspect of the effect of data revisions. In practice, this probably means that tabulated RMSQEs are too small. Research has shown that taking account of vintages of data usually widens the error bands. Sometimes these bands can widen sufficiently to render the forecasting equation meaningless, as in, for example Diebold and Rudebusch (1991). The equations using data that are not revised—equity prices, interest rates, survey data—are probably least affected by this problem. On the other hand, RMSQEs from data that can be revised substantially—the leading index, export growth, retail sales, etc.—are likely to be understated relative to the performance that could be expected in “real time.”

Table 2 also shows the RMSQE’s for four major aggregations of like data. Of those aggregations, financial data have the highest RMSQE, while production and sales data have the lowest.

### The Recent Track Record for the RTFS

The RTFS is largely still in its infancy: we began using this kind of system in December 2001. Despite the short track record, our experience with the “on the ground” application of real time forecasting has been educational. For example, initially we reported results based on only the production and employment indicators. These indicators, at the time, were producing GDP estimates for the fourth quarter of 2001 that were largely in line with private estimates—looking for roughly a two percent decline in GDP at an annual rate. However, the broader system prediction was predicting a number much closer to zero, indicating the economy was not as weak as suggested by the production and employment indicators alone. By the middle of January 2002, most private forecasters had revised up their forecasts to pre-

TABLE 2

#### ROOT-MEAN-SQUARED ERRORS BY INDICATOR By number of months of available data

Number	Indicator and Transformation	Months of data for quarter included		
		1-Month	2-Month	3-Month
1	Consumer Confidence, Levels	2.19	2.11	2.07
2	Consumer Sentiment, Levels	2.17	1.98	1.91
3	Total Light Vehicle Sales, Growth	2.31	2.15	2.10
4	Real Retail Sales, Growth	2.03	1.94	1.92
5	Real Personal Income, Growth	2.13	2.03	2.00
6	Real PCE, Growth	1.94	1.92	1.87
7	Aggregate Hours, Growth	2.08	1.75	1.77
8	Payroll jobs, Growth	2.04	1.96	1.92
9	Unemployment rate, Levels	2.33	2.27	2.23
10	Un. Insurance Claims, Levels	1.98	1.83	1.67
11	Help Wanted Index, Levels	2.45	2.34	2.17
12	Job Diffusion Index 1Mo., Levels	2.05	2.02	2.00
13	S&P 500, Growth	2.02	2.10	2.07
14	Dow Jones 30, Growth	2.09	2.10	2.13
15	Nominal Goods Exports, Growth	2.07	2.08	1.99
16	Real 10 Yr Govt Yield, Levels	2.15	2.14	2.22
17	Baa-10 Yr Rate, Levels	2.72	2.62	2.54
18	ISM Manuf. PMI, Levels	2.06	1.97	1.90
19	Industrial Production, Growth	1.76	1.74	1.69
20	IP, Computers and Off. Equip., Growth	2.17	2.22	2.23
21	Total Capacity Utilization, Levels	2.13	1.95	1.91
22	Business Week Prod. Index, Growth	2.24	2.10	2.09
23	Real Durable Gds Orders, Growth	2.11	2.06	1.85
24	Housing Starts, Growth	2.10	2.07	2.03
25	Construction, Growth	2.18	2.05	2.06
26	NFIB Optimism Index, Levels	2.19	2.18	2.01
27	Conf. Brd. Leading Index, Growth	1.71	1.61	1.57
28	ECRI Wkly Leading Index, Growth	2.09	2.05	2.04
29	ISM Inventory Index, Levels	2.18	2.10	2.07
30	Phil Fed Index, Levels	1.96	1.91	1.86
<b>Sub-Aggregates</b>				
	Employment	2.04	1.90	1.82
	Financial	2.48	2.34	2.31
	Survey	1.98	1.91	1.83
	Production and sales	1.81	1.70	1.64

*Employment* sub-aggregate consists of RBarSq weighted predictions of rows 7, 8, 9, 10, 11, 12  
*Financial* sub-aggregate consists of RBarSq weighted predictions of rows 13, 14, 16, 17  
*Survey* sub-aggregate consists of RBarSq weighted predictions of rows 1, 2, 18, 26, 29, 30  
*Production and sales* sub-aggregate consists of RBarSq weighted predictions of rows 3, 4, 5, 6, 15, 19, 20, 21, 22, 23, 24, 25  
 ISM: Institute for Supply Management  
 IP: Industrial Production  
 NFIB: National Federation of Independent Business  
 ECRI: Economic Cycle Research Institute

dict something close to zero for the quarter. The final prediction for the quarter was a slightly positive number. The advance estimate was somewhat larger, at 0.8 percent. This early experience suggested that the information in the broader set of indicators should not be ignored—reinforcing the fundamental methodology on which the system was based.

Experience in subsequent quarters has provided additional insights and has been somewhat more directly revealing of potential shortcomings. During the first and second quarters of 2002, the performance of the system relative to actual results highlighted the importance of good forecasts of inventories and imports, neither of which is tracked well by the system. As we watched these results evolve, we began using the system to predict final sales growth as well as final sales to domestic purchasers—key measures of aggregate demand in the economy. After two difficult quarters, the system nearly hit the GDP advance for the third quarter of 2002 exactly.

Although the sample is short, it is useful to compare the RTFS predictions with the NABE macroeconomic outlook for the five quarters from 2002Q1 through 2003Q1. The NABE outlook is usually conducted late in the first month or early in the second month of a calendar quarter. The September outlook is an exception to this general rule. That outlook is presented at the annual NABE meeting and the survey is conducted up through roughly the middle of September (the third month of the third quarter).

As a result of this schedule, the information available to NABE forecasters is dramatically different among the surveys. For all surveys but September, there are very few monthly indicators available for the current quarter. Among those that might be available to forecasters include the Michigan consumer sentiment figure, ISM manufacturing indexes, the Philadelphia Fed's business outlook survey, and, potentially, the first month of the quarter's employment report. For all but the September survey, the NABE outlooks mostly are a true one-quarter ahead forecast.

For the September survey, all monthly indicators would have at least one month's worth of data on the quarter, and many would have two month's worth (employment and the ISM indexes, for example). Financial data (money supplies,

TABLE 3

COMPARING NABE OUTLOOK AND RTFS FORECASTS 2002-2003

Forecast for Quarter	NABE Outlook	RTFS Forecasts Made at Different Points in the Quarter			Actual GDP Growth	
		Early	Middle	Late	Advance	Current
2002Q1	Date: 1st-2nd wk of Feb Forecast: 1.3	Feb 14 1.0	Mar 15 2.3	Apr 16 2.6	5.8	5.0
2002Q2	Date: Last wk Apr, 1st wk May Forecast: 3.0	May 1 3.1	Jun 7 3.3	Jul 5 3.2	1.1	1.3
2002Q3	Date: Sep 3-19 Forecast: 3.1	Aug 2 3.2	Sep 13 3.2	Oct 11 3.1	3.1	4.0
2002Q4	Date: Nov 6-14 Forecast: 1.4	Nov 8 2.7	Dec 6 2.3	Jan 10 2.2	0.7	1.4
2003Q1	Date: Jan 27-Feb 4 Forecast: 2.7	Feb 7 2.6	Mar 7 2.7	Apr 11 2.3	1.6	1.4
2003Q2	Date: May 2-16 Forecast: 1.8	May 8 1.5	June 6 2.0	Jul 11 1.9	2.4	3.1

interest rates, equity prices, etc.) would be available for two full months of the quarter and for some of the third month. Table 3 shows the RTFS forecast that was done at a roughly equivalent time in the quarter as the forecasts contained in the NABE survey. For the surveys other than September, the RTFS prediction is using only very partial information (about half of the 30 indicators have no data).

The table shows that the behaviors of the NABE and RTFS forecasts are quite similar. Both forecasts had considerable difficulty picking up the surge in GDP in the first quarter of 2002, and both over-predicted growth in the 2002Q2. Both forecasts did well for the third quarter (with much more data about the quarter in hand). The NABE forecasters correctly saw the slowdown in the fourth quarter of 2002, which the RTFS missed by a considerable margin. NABE forecasters and the RTFS made similar over-predictions for the first quarter of 2003, and also had similar outlooks for the second quarter in the early part of May of 2003.

General Observations

Initially, we focused on using the RTFS to predict the specific quarterly values of real GDP growth. However, we have observed that the RTFS forecasts may actually be doing a better job of tracking the “underlying” real GDP growth performance, not necessarily the actual quarterly values. In other words, the RTFS may be extracting the “growth trends” from the GDP growth series that exhibit substantial quarterly volatility. In practice, that may be a good result. For example, suppose the underlying growth rate of real GDP across a number of quarters is actually 3-1/2 percent, but the pattern of the actual quarterly growth rates is randomly dispersed across a range of one to six percent. Policymakers would be better informed by a system that properly tracked the underlying growth rate rather than the abnormally (and presumably transitory) low or high quarterly movements. Even at a turning point for the econ-

omy, a forecast that properly captures the underlying signal and not the volatile movements might still be preferred, as long as it somehow indicated that GDP was “slowing significantly.” The RTFS has not been in place long enough to establish a track record on this score. However, the one observation we had—in the fourth quarter of 2001, suggests that the broad system was initially successful in recognizing that the economy was not as weak as private forecasters initially had thought. In retrospect, this conforms to the recent determination by the NBER Business Cycle Dating Committee’s determination that the recession ended in November 2001.

We continue to experiment with and to further develop the RTFS. We are already applying it beyond GDP to final sales and domestic final sales, and we also have been experimenting with it to forecast the growth in real equipment and software investment and corporate economic profits. We are looking at several extensions, including perhaps extending the coverage of indicators beyond GDP and its closely-related aggregates. In addition, we are looking at several technical issues, including fitting the models on quarterly growth in real GDP, rather than annualized quarterly growth (which magnifies the actual movements in GDP); using a criterion other than least squares to fit the regressions; and, as stated earlier, investigating the issue of data vintages.

So far we have found that there are several advantages in the general layout of the system. They are: (1) it is relatively easy to maintain, (2) it is easy to understand, (3) it can incorporate incoming data in a systematic way nearly instantaneously, and (4) it is a competitive predictor of quarterly GDP growth. The system has a potentially broad application—the approach can be used for virtually any indicator or goal variable of interest as discussed above.

On the whole, the RTFS system has shown much promise in its early operation and applications, and it represents a potentially valuable approach for providing timely information to policymakers and business analysts. ■

## ACKNOWLEDGEMENTS

Much of the initial research and the work on earlier versions of this paper were conducted while Kitchen was an economist with the Department of the Treasury. The authors thank Rich Clarida, who described the original idea to us during the early days of his tenure as Assistant Secretary for Economic Policy. Thanks are also extended to Karen Hendershot and Robert Stein for their many useful comments on the system as it has evolved.

## REFERENCES

- Croushore, Dean and Tom Stark. 1999. “A Real-Time Data Set for Macroeconomists.” Federal Reserve Bank of Philadelphia. Working Paper. Pp. 99-4. June.
- . 2001. “A Real-Time Data Set for Macroeconomists.” *Journal of Econometrics*. 105. Pp. 111-130. November.
- Diebold, Francis X., and Glenn D. Rudebusch. 1991. “Forecasting Output With the Composite Leading Index: A Real-Time Analysis.” *Journal of the American Statistical Association*. 86. Pp. 603-10. September.
- Dynan, Karen E. and Douglas W. Elmendorf. 2001. “Do Provisional Estimates of Output Miss Economic Turning Points?” Federal Reserve Board of Governors. Working paper.
- Ingenito, Robert and Bharat Trehan. 1996. “Using Monthly Data to Predict Quarterly Output.” *Federal Reserve Bank of San Francisco Economic Review*. No. 3. Pp. 3-11.
- Klein, L. R. and E. Sojo. “Combinations of High and Low Frequency Data in Macroeconometric Models.” In Klein, L.R. and J. Marquez (eds). 1989. *Economics in Theory and Practice: An Eclectic Approach*. Kluwer Academic Publishers. Pp 3-16.
- Koenig, Evan, Sheila Dolmas, and Jeremy Piger. 2001. “The Use and Abuse of ‘Real-Time’ Data in Economic Forecasting.” Federal Reserve Bank of Dallas. Working paper.
- Orphanides, Athanasios. 2001. “Monetary Rules Based on Real-Time Data.” *American Economic Review*. Volume 91. Number 4. Pp. 964-985.
- Robertson, John C., and Ellis W. Tallman. “Data Vintages and Measuring Forecast Model Performance.” *Federal Reserve Bank of Atlanta Economic Review*. Fourth Quarter 1998b. Pp. 4-20.
- Stock, James H. and Mark W. Watson. 1999. “Forecasting Inflation.” *Journal of Monetary Economics*. Volume 44. Number 2. Pp. 293-335.
- Stock, James H. and Mark W. Watson. 2001. “Forecasting Output and Inflation: The Role of Asset Prices.” NBER Working Paper 8180. March.
- Watson, Mark W. 2001. “Macroeconomic Forecasting Using Many Predictors.” Working Paper. July 2000. Revised December.