Federal Reserve Bank of Minneapolis
Research Department

ALGORITHMS FOR EXPLAINING FORECAST REVISIONS

Richard M. Todd*

Working Paper 459

May 1990

ABSTRACT

Forecasts are routinely revised, and these revisions are often the subject of informal analysis and discussion. This paper argues 1) that forecast revisions are analyzed because they help forecasters and forecast users to evaluate forecasts and forecasting procedures, and 2) that these analyses can be sharpened by using the forecasting model to systematically express its forecast revision as the sum of components identified with specific subsets of new information, such as data revisions and forecast errors. An algorithm for this purpose is explained and illustrated.

KEYWORDS: Forecasting Forecast revisions Data revisions Innovation accounting

*Federal Reserve Bank of Minneapolis. I have benefited from conversations with Larry Christiano, Preston Miller, David Runkle, John Rust, Christopher Sims, and Warren Weber. Peter Young and a referee gave me valulable ideas for improving the exposition of my ideas. Thanks also to Kathy Rolfe for suggesting formats for my tables and to Pat Haswell for editorial assistance. I remain responsible for any faults.

Forecasters in economics and other disciplines frequently forecast the same event more than once, as time passes and information relevant to the event accumulates. Since the newly accumulated information — consisting mainly of revisions of old data and releases of new data — often changes the forecast, the forecaster frequently generates a sequence of forecast revisions. The revisions, in turn, raise questions about why the forecast changed, or, more specifically, about which pieces of new information were primarily responsible for particular changes in the forecast. Attempts to answer those questions are, in the broadest sense, forecast revision analysis.

The purpose of forecast revision analysis is to improve our understanding of forecasts, forecasting procedures, and the actual systems that are being forecasted. It does this by helping to reveal the properties of the forecasting procedure, especially its dynamic properties.

Although forecast revision analysis is a useful and fairly common activity, published accounts of a systematic method for providing explanations of forecast revisions appear to be nonexistent. This paper proposes, discusses, and illustrates an accounting framework that should be a useful and practical first step toward a complete methodology for forecast revision analysis. The core of the procedure is fairly simple and proceeds roughly as follows: At time t+k, the set of information accumulated since the time–t forecast is partitioned into subsets. Then a series of data sets is constructed, each one of which augments the information set used in the time–t forecast by some subset (or union of subsets) from the partition of new information. The model is then used to compute a forecast for each of these artificial data sets, and the resulting forecasts are examined to see which subsets of information had the biggest impact in changing the forecast of each variable at each horizon of interest.

This kind of an accounting scheme can provide information useful to both forecasters and forecast users. The first section of the paper briefly discusses why forecasts are revised and why these revisions are analyzed. The second section uses an economic forecasting model to illustrate how a specific accounting procedure has already been used to provide useful and sometimes surprising information. The third section describes how the accounting scheme can be generalized for use in a wide variety of forecasting contexts.

THE ROLE OF A FORECAST REVISION ALGORITHM

Forecast revision analysis is a common activity whose purpose is to give insights into forecasts, into the properties of forecasting procedures, and possibly, into the properties of the forecasted systems. However, existing examples of forecast revision analysis appear to be ad hoc and informal. An algorithm that would automatically decompose a model's forecast revision into components identified with the direct and interactive effects of the members of a detailed partition of the set of new information would make forecast revision analysis both easier and more informative.

Forecasts are frequently revised

Forecasts often change as the information set they are based on is revised or enlarged over time. To put this in formal terms, assume that an analyst forecasts a vector Y_t. That is, at each date t, the analyst assembles an information set I_t and uses it to estimate some parameters of the probability distribution of Y_{t+k}, $k = 1,2,\ldots,K$. For each k, let ${}_tX_{t+k}$ be an $n\times1$ vector containing estimates based on I_t of n parameters of the probability distribution of Y_{t+k}. (Typically X_{t+k} will include at least a measure of central tendency, such as the mean or mode, of each component of Y_{t+k}, and not uncommonly some second moment information is also included.)

After h periods, $1\leq h<K$, an updated information set I_{t+h} is compiled and a new forecast, ${}_{t+h}X_{t+k}$, is computed for all k such that $h<k\leq h+K$. This means that for all k such that $h<k\leq K$, one can compute the forecast revisions $({}_{t+h}X_{t+k}-{}_tX_{t+k})$. When the forecasts are taken directly from a formal statistical model and the set of new information (I_{t+h}/I_t) contains anything that affects the model, it is almost certain that the revision to at least some component at some horizon is nonzero. When the forecasts

are constructed by an analyst who subjectively adjusts the forecasts of a formal model, however, revisions may occur less frequently. (See Caton 1987.)

Analyzing the revisions is useful

It is clear that forecasts are revised fairly frequently. It is less clear why we want explanations of forecast revisions. If we have an optimal forecast in our hands today, one that we wholeheartedly believe to contain a true and complete description of the conditional (on I_t) probability distribution of all future events that interest us, why should we care about how it came to be? Today's optimal forecast is sufficient for making today's decisions; yesterday's forecast should be of concern only to historians. In fact, however, forecasters, journalists, and forecast users all pay attention to forecast revisions. (See Todd (1989) for documentation of this phenomenon for economic forecasts.)

We reason that if forecast users believe that the forecast they receive accurately describes the conditional probability distribution of future events, then they will not care about forecast revision analysis. We observe, by contrast, that forecast revision analysis occurs. An obvious conclusion, and probably the main reason why forecast revisions are analyzed, is that forecast users don't regard the forecasts they receive as true and complete descriptions of the conditional probability distribution of future events. Instead, forecast users perceive a variety of suboptimal forecasts competing for their attention. They must decide what weight to attach to each one. Forecast revision analysis is one natural way to satisfy forecast users' needs for information about a forecast and the procedure used to create it.

Although the above discussion of the role of forecast revision analysis could be extended (see Todd 1988 and 1989), it is adequate to justify attention to methods for providing forecast revision analyses. There appear to be rational reasons for forecast users to demand and forecasters to provide explanations of forecast revisions. We observe that, in fact, forecast revisions are analyzed. We will now see how forecast revision analysis can be performed by an automatic, comprehensive, model–based algorithm.

AN EXAMPLE OF AN AUTOMATIC FORECAST REVISION ANALYSIS

To illustrate the results of a forecast revision algorithm, consider a vector autoregressive model used at the Federal Reserve Bank of Minneapolis. It forecasts monthly values of seven key U.S. macroeconomic variables—the Dallas Federal Reserve Bank's index of the foreign exchange value of the U.S. dollar (DOLLAR), Standard and Poor's index of 500 stock prices (SP500), the interest rate on three–month U.S. treasury bills (TBILL), real gross national product (GNP), the GNP deflator (DEFL), the change in business inventories (CBI), and the Federal Reserve Board's measure of the monetary base (MB). (Monthly values for some of these series are created by interpolation. See Amirizadeh (1985) for how this is done and Todd (1989) for complications this causes in forecast revision analysis.) The model can stand alone but also acts as the core sector in a large macroeconomic model. (See Litterman (1984) for a description of both the core model and the larger model.)

When a revised forecast is computed from this model, a forecast revision algorithm is used to partition the set of new information (accumulated since the previous forecast) into 15 subsets. Seven subsets contain information on data revisions (revised estimates, available only after the previous forecast was made, of the values of data used in making the previous forecast). Seven other subsets contain information on newly released data. The 15th subset is a catchall for a variety of miscellaneous factors affecting the forecast.

Table 1 shows the standard format used for reporting the algorithm's decompositions of the model's forecast revisions. The initial forecast in this example was made in September 1987, based on data through July 1987 for GNP, DEFL, and CBI and through August 1987 for the other variables. The subsequent forecast was made in March 1988, based on revised and new data through January 1988 for GNP, DEFL, and CBI and through February 1988 for the others. The table analyzes percentage point changes in forecasts of the levels of the variables in December 1988, except for TBILL (basis point units) and CBI ($billion units).

To see the decomposition of the revision in a particular variable's forecast (for example, GNP), find the column corresponding to that variable (column four) and read down. The first row shows the variable's total forecast revision for December 1988 (–1.09 percentage points). The next three rows give a gross decomposition of the total revision into effects attributed to data revisions (+1.70 percentage points), newly released data (–2.77 percentage points), and miscellaneous factors (–0.02 percentage points).

The next block of rows further decomposes the total data revision effect into effects attributed to each variable's data revision. The first row of column four in this block, for example, attributes –0.02 percentage points out of the total +1.70 percentage point effect of data revisions on the GNP forecast, to revisions in the data on the exchange value of the U.S. dollar.

The final block decomposes the total effect of the newly released data into the effects attributed to each variable. The third row of column four in this block, for example, says that fundamental forces affecting the GNP deflator contributed +0.29 of the total –2.77 percentage point effect of new data on the GNP forecast for December 1988.

In many ways the results in Table 1 are typical for this model's forecast revisions. Surprises in the new data are usually more important than data revisions, but data revision effects are far from trivial. The information summarized by the miscellaneous factors is trivial. In the lower blocks, diagonal elements generally dominate. This reflects the fact that in this model a variable's own revisions and the part of its forecast errors attributed to itself usually account for most of its forecast revisions.

Table 1 shows some less common results too. Most prominent are some strong cross–variable effects attributed to new data on SP500. Considering that the October 1987 stock market crash intervened between the initial and subsequent forecasts, these strong effects are understandable. The algorithm says that the stock market crash itself cut the forecasts of December 1988 GNP, MB, and TBILL by 2.75 percent, 0.57 percent, and 57.16 basis points, respectively.

Table 1 also illustrates some of the benefits of forecast revision analysis. The large effect attributed to GNP revisions highlights the importance of data revisions in general. In a concerted effort by several economists at the Minneapolis Federal Reserve Bank to unravel why the forecast of GNP had not been depressed more by the stock market crash, this factor had not been considered until the algorithm was used to produce Table 1. The table also helped rank competing explanations. For example, the fall in interest rates in late 1987 had been proposed as an explanation for why the model's GNP forecast had changed by only –1.09 percent. However, the table shows that errors in forecasting late 1987 interest rates had a trivial effect (+0.04 percent) on the revision of the GNP forecast for December 1988.

THE SPECIFIC ALGORITHM BEHIND THE EXAMPLE

Table 1 was produced by a computer-based algorithm that contains the forecasting model and requires as input mainly the old and new data sets and the dates at which the old and new forecasts were computed.

The algorithm constructs the subsets of data revisions used in Table 1 in a very simple manner. There is one of these subsets for each variable in the model, and each one contains all of the revisions to the recent values of a particular variable. (The cutoff between "recent" and older revisions is made by examining the revisions to each series and subjectively selecting the most recent date before which the revisions all appear to be trivial. If a conservative approach is desired and speed of the algorithm is not an issue, the cutoff can be pushed back in time arbitrarily far.) For example, the subset on GNP data revisions contains all the most up-to-date revisions to the GNP data used in the previous forecast, and nothing else. (This scheme for handling revisions implicitly assumes that a) revisions are independent across variables and b) it is not interesting to distinguish among revisions at different lags for a given variable. If these assumptions are not reasonable in a particular forecast revision analysis, they could be changed, as discussed in Section III.)

The algorithm constructs the subsets of newly released data by a more elaborate procedure. First, to isolate what is really new, the new data are converted into the form of forecast errors. Second, to assign these forecast errors to a well–defined underlying cause, auxiliary assumptions are built into the algorithm to unravel the correlations among the components of the forecast error vectors.

The process of converting the new data into forecast error vectors is only a bit complicated, primarily because the algorithm requires a sequence of one–step–ahead error vectors rather than a mixture of one– to j–step–ahead ($j>1$) error vectors. Because the effects of data revisions have already been assigned to other subsets, the computation of the sequence of one–step–ahead error vectors uses only the revised values of all data. It begins by reestimating the model's coefficients using the revised data through t, the time of the previous forecast. These coefficients and the revised data through t are used to forecast X_{t+1}, and then the one–step–ahead error vector for t+1 is computed. Next the coefficients are reestimated through t+1 and a forecast and forecast error vector are computed for t+2. This process continues iteratively until a one–step–ahead forecast error vector for time t+h has been computed.

One way to proceed at this point would be as in the case of the data revisions: Assume that each variable's one–step–ahead forecast errors are independent of the one–step–ahead errors made in forecasting the others and simply assign each variable's one–step–ahead forecast errors to itself. This is not likely to be reasonable, however. The point of having a multivariate model, after all, is that the analyst believes that the variables affect one another. If so, a one–step–ahead error in forecasting one variable may contribute to one–step–ahead errors in forecasting all the others, and vice versa. Some scheme is needed to unravel this mutual interdependence of the forecast errors.

There is no unique way to devise the necessary unraveling, however. It can only be done by adding auxiliary assumptions to the model. (In economics, this unraveling is referred to as identifying the model and is quite controversial. I shall avoid the term "identifying" here, though, because of its different meanings in economics, time series analysis, and engineering.)

In the forecast revision algorithm used at the Minneapolis Federal Reserve Bank, the unraveling assumptions take the form of a causal chain imposed upon the components of the one-step-ahead forecast error vectors. (Analysts who prefer other sets of assumptions may wish to change this part of the algorithm for their own applications, but the causal chain suffices for illustrating the procedure.) Briefly, a causal order running from DOLLAR to SP500 to TBILL to MB to DEFL to CBI to GNP is imposed on the one-step-ahead errors. Then, the covariance matrix Σ of the components of the one-step-ahead errors, computed from the recursive residual vectors generated when the model is estimated via the Kalman filter, is used to decompose the series of one-step-ahead forecast error vectors.

The unraveling begins, in each forecast error vector, by attributing 100 percent of the component corresponding to DOLLAR (the first variable in the causal chain) to fundamental forces that determine exchange rates. In addition, these exchange rate determinants are held responsible for a part of each other component. Those additional parts are computed by projecting each of the other components onto the DOLLAR component, using a population regression coefficient computed from Σ. At this stage, the DOLLAR component of each one-step-ahead forecast error has been unraveled, but only the dollar's own forecast error has been completely assigned to an underlying cause.

In each one-step-ahead forecast error vector, the remaining unexplained portion of the next variable in the chain—SP500—is then assigned to fundamental forces affecting stock prices. In addition, parts of the forecast error components of variables lower in the chain are also assigned to the SP500 forces, again via projections based on population regression coefficients computed for Σ. (See Hakkio and Morris 1984 or Doan 1988, pp. 8–9, for details.)

This process continues down the chain until all the components of the one–step–ahead error vectors have been decomposed into one or more parts, each assigned to underlying factors determining one of the model's variables. This completes the construction of the seven subsets of information contained in newly released data.

The 15th subset is a collection of miscellaneous information. It consists primarily of distant data revisions (mainly induced by small changes in the seasonal adjustment of distant time periods) and slight data–induced changes in the "priors" of the Bayesian procedure used to estimate the model (see Doan, Litterman, and Sims 1984).

Given these 15 subsets of new information, the forecast revision algorithm next computes each one's incremental contribution to the forecast revision. The data revision subsets are ordered DOLLAR, SP500, TBILL, MB, DEFL, CBI, and GNP. So are the seven subsets of the information contained in newly released data. (These orderings could also be different from each other or from the ordering used in the causal chain discussed above. In linear models with fairly stable coefficients, such as the model used here, the ordering of the subsets at this stage of the algorithm doesn't affect the results much.)

The algorithm first proceeds through the data revision subsets. It compares the original forecast with a forecast based on the original data set modified by revisions in the DOLLAR data. (To do this, it reestimates the model's coefficients and then reforecasts.) The difference between the two forecasts—for every horizon and in each variable—is attributed to revisions in the data on DOLLAR. Next, the algorithm adds revisions to both SP500 and DOLLAR to the original data set. The model is reestimated, a new forecast is computed, and the differences between this forecast and the one based on the addition of just the DOLLAR revisions are attributed to SP500. This continues through the subset of data revisions attributed to GNP. At that point, the total effect of data revisions has been decomposed into seven parts, one for each variable.

The algorithm must also proceed through the seven subsets of information from newly released data. Recall that the new data cover the period from t through t+h, the date of the revised forecast. The algorithm first computes a forecast of t+h+1 to t+K based on the revised data through t plus the forecasted "data" for t+1 to t+h. This forecasted data is obtained by reestimating the model, with fully revised data, through t and then forecasting t+1 through t+h. In principle, the forecast of t+h+1 to t+K based on this combination of revised and forecasted data through t+h should be the same as the final forecast of t+h+1 to t+K made in the part of the algorithm that analyzed data revisions. This is because the forecasted data for t+1 through t+h contain no surprises that would cause the Kalman filter to revise the model's coefficients. In practice, minor discrepancies arise, due to miscellaneous factors (distant data revisions ignored by the data revision analysis, data–based revisions to "priors" induced by reestimating through t+h instead of t). In this way, processing of the subset of miscellaneous information is interposed between processing of subsets on data revisions and subsets on newly released data.

From the baseline just introduced—a model estimated through t+h on revised data through t plus forecasted data through t+h—analysis of the newly released data proceeds. First the parts of the t+1 to t+h forecast errors attributed to DOLLAR are added to the data base used in the baseline. (Recall that the components assigned to DOLLAR include parts of the one–step–ahead errors in forecasting the other six variables, as well as all of the one–step–ahead errors in forecasting DOLLAR. So the data on all seven variables are changed in this step.) The model is reestimated through t+h and a new forecast of t+h+1 to t+K is computed. The difference between this forecast and the baseline forecast is attributed to DOLLAR.

Next the data set is further augmented by adding the parts of the t+1 to t+h one–step–ahead forecast errors attributed to SP500. Again the model is reestimated and reforecasted, and the incremental changes are assigned to SP500. This process repeats through GNP, when the data set is identical to the full new set of data, the forecast equals the fully revised forecast, and the decomposition of forecast errors is complete.

A GENERAL ALGORITHM FOR FORECAST REVISION ANALYSIS

In principle, algorithms like the one used in the previous section can be developed to automatically and comprehensively analyze most forecast revisions, or at least those based, exactly or approximately, on formal forecasting models. The informativeness of the algorithm is limited mainly by the richness of the assumptions—such as the causal chain in the examples above—used to partition the set of new information into subsets. These assumptions vary from one forecasting model or procedure to another, so I present a description of how to construct an algorithm for a generic, model–based forecast and an arbitrary partition of the set of new information.

A general algorithm

As in the examples of Section II, the first stage in the algorithm is to partition the set of new information, I_{t+1}/I_t, into disjoint subsets w^i, $i = 1,2,...,m$. (The w^i may be scalars or vectors, depending on how finely the analyst wishes to conduct the analysis.) Then, the forecasting model is used to compute the effects of each subset in the partition of new information. Consider the forecast of the i^{th} component of X_{t+k} as an exact function of the information it is based on, so that the revision between the forecasts at times t and t+h can be expressed as an exact function of the new and old information, or

$$_{t+h}X_{i,t+k} - {_t}X_{i,t+k} = g_{hik}(w^1;w^2;...;w^m;I_t).$$

The general algorithm proceeds, in the spirit of a Taylor series expansion, by approximating g by the sum of higher and higher orders of interactive effects among the subsets, up to a finite limit.

To compute a first order analysis of the effect of each subset of new information on the forecast revision, let d_j be (I_{t+h}/I_t) if j=0 and be the null set otherwise. Then let

$$
{}_{t+h}X_{i,t+k} - {}_{t}X_{i,t+k} = \sum_{s=1}^{m}\{g_{hik}(w^{1}\cap d_{1-s};\ldots;w^{m}\cap d_{m-s};I_{t})
$$

$$
- g_{hik}(\emptyset;\emptyset;\ldots;\emptyset;I_{t})\} + R(1).
$$

In this equation, each term in brackets shows how the forecast of the i^{th} component of X at horizon t+k would have changed if the old information set I_t had been augmented by just the s^{th} subset of new information, $s = 1,2,\ldots,m$. Note that this can be computed without deriving the function g_{hik}. All that is required is to run the forecasting procedure twice, once with the original dataset and once with a dataset augmented by the s^{th} subset of new information. The sum in the expression above is the part of the total revision explained by first order effects, with R(1) as the first order residual.

A second order analysis could be computed by decomposing R(1) into a sum of second order effects and a new residual, R(2). The individual second order effects would be computed by forming all pairs (w^i,w^j), $i\neq j$. For each of these pairs, subtract a forecast based on I_t from a forecast based on I_t augmented by the union of w^i and w^j. From this difference, subtract the first order effects associated with w^i and w^j. The remainder is the second order effect attributed to the interaction of w^i and w^j. The total change would then be expressed as the sum of the first order effects plus the sum of the second order effects plus R(2).

The analysis could be extended to third order by examining all triples, net of their first and second order effects, and so on for higher orders up to m, for which all new information is used as a single subset and the residual is identically zero. Alternatively, the analysis might be truncated short of order m, under the assumption that the residual was small and further efforts to decompose it were not worthwhile. Whatever the chosen order, the analysis would also be done, typically simultaneously, for all components of X and all forecast horizons of interest. The result would be a decomposition of each revision in the forecast into components attributed to the direct and interactive effects of the subsets of new information. This completes the description of the general forecast revision analysis algorithm.

Some comments on partitioning the set of new information

Since the subsets of new information are the building blocks of the forecast revision algorithm, the method of partitioning the set of new information strongly influences the amount and quality of information that the forecast revision algorithm provides. Unfortunately, it is difficult to give general advice on how to construct the partition. The optimal partition will depend on both the nature of the forecasting procedure and the interests and beliefs of the forecaster and forecast users.

To illustrate the partitioning concept, consider partitioning the set of data revisions. Let ${}_tZ_{t-j}$ be a p×1 data vector containing estimates made at time t of the values that the p variables took on at time t–j, j = 0,1,2,.... In particular, let ${}_tZ_{t-j}$ contain the most recent estimates available from official statistical agencies at time t. Suppose that I_t contains ${}_tZ_t$, ${}_tZ_{t-1}$, ... ${}_tZ_{t_0}$, where data before t_0 either are unavailable or are not used in preparing the forecasts. (Many U.S. macroeconomic forecasters seem to use $t_0 \geq 1948$, for example.) Then at time t+h the set of new information will contain a subset of data revisions, namely $R \equiv \{r_{i,t-j}$, for $i = 1,2,...,p$ and $j = 0,1,2,...,t_0-t\}$, where $r_{i,t-j} = ({}_{t+h}Z_{i,t-j} - {}_tZ_{i,t-j})$ is the revision of the i^{th} component of Z for time period t–j.

The question of how to treat the elements of R in the forecast revision analysis will depend in part on what the forecast users want from the analysis. If, for example, the forecast users don't care about the details of how data revisions affected the forecast, then it might be adequate to treat all of R as a single one of the w^k. That is, all data revisions would be treated collectively, as one indivisible piece of new information. If the users are interested in the details of how data revisions affected the forecast, then one could assume that each element of R is a separate, independent piece of new information. This means that each $r_{i,t-j}$ could be set equal to an w^k. Alternatively, the $r_{i,t-j}$ elements could be grouped into nonoverlapping unions, such as by taking unions over the i components and treating as one indivisible piece of new information all the revisions to all the components of Z at time t–j, for each j.

The treatment of R may also depend on what information the forecaster can supply about how to partition R. For example, the forecaster may believe that the elements of R are correlated, being themselves the effects of a smaller set of fundamental data measurement corrections, which we can call $R^* \equiv \{r_1^*, r_2^*, \ldots, r_K^*\}$. Then each r_i^* could be set equal to an w^k. To do this, the forecaster would have to explain both how to recover R^* from R and how to functionally relate the elements of R to the elements of R^*. In every case, the approach depends on auxiliary assumptions used to establish mappings between a set of observed changes and an underlying set of primitive, independent causes.

As noted previously, economists refer to such assumptions as identifying assumptions. Because this term has other meanings in other fields, the term unraveling assumptions has been used in this paper.

Because the forecast revision algorithm treats each subset of the partition of new information as an independent cause of the revision, unraveling assumptions are essential for interpreting the results of the algorithm. Much of the economics literature on this topic focuses on relating the forecast errors to an underlying set of theoretically meaningful disturbances. These techniques could be used to apply this paper's algorithm to an econometric model. As the above example shows, however, for forecast revision analysis it is also useful to apply unraveling assumptions to other components of the set of new information, such as data revisions, changes in exogenous variables, and changes in subjective adjustments. When this is not possible, analysis of the part of the forecast revision attributed to these components may not be very informative.

Although the forecast revision algorithm requires some unraveling assumptions to partition the set of new information, it does not depend on any particular type. That is important given the controversies, at least in economics, about which assumptions are appropriate. Two forecasters who disagree about unraveling assumptions will probably also disagree about forecast revision analyses, because their differing assumptions will result in different partitions. Given their respective partitions, however, they can both use the general algorithm outlined above.

Other practical considerations

A wide variety of model-based forecasting procedures are in use. The convenience and usefulness of the general algorithm is likely to vary among the various types of forecasting procedures. Some of the complications are forecasts made by combining a pure model forecast with subjective adjustments made by the human analyst, forecasts based on models with exogenous variables (whose own forecasts are produced outside the model), and forecasts revised after a change in the units in which data are measured. Todd (1989) discusses how to adapt the general method presented here to address those complications.

It is also worth noting that the convenience of the forecast revision algorithm will depend on the complexity of the forecasting model. The algorithm requires the whole forecasting procedure to be repeated many times on a variety of datasets. If the forecasting model was reestimated between the old and new forecasts, then the algorithm requires reestimating it for each of those datasets as well. This can be quite demanding for large models with complicated estimation procedures. Furthermore, the order of the expansion required to make the residual a relatively small component may be large for models with high order nonlinearities.

By contrast, a linear model like the one used in Section II can be reestimated quickly (e.g., via a Kalman filter) and encounters nonlinearities only to the extent that new data are multiplied by revised coefficients. A first order expansion is thus easier to compute and more likely to have a small residual for a linear model.

A careful reader may have noticed, in fact, that the forecast revision analyses described in Section II do not exactly conform to the general algorithm. The Section II algorithm did not even compute the first-order expansion of the general algorithm. Instead of computing each subset's individual effect relative to a baseline forecast computed from just the old database, the Section II algorithm ordered the subsets and computed each one's incremental effect relative to a forecast computed from the old dataset augmented by all previous subsets in the order. (See Todd 1988 for more details.) Because the models are linear with coefficients that have been changing only very slowly over time, the mapping from the set of new plus old information to the forecast revision is nearly linear. As a result, the modified algorithm, which is somewhat simpler to program, closely approximates the first order terms of the general algorithm. Use of the general algorithm would change the results shown in the text tables by no more than a few tenths of a percentage point.

CONCLUDING REMARKS

An algorithm for organizing the analysis of forecast revisions has been abstractly defined and concretely illustrated. Although use of the algorithm requires that an analyst use statistical reasoning and modeling theory to partition the set of new information, the algorithm itself is essentially an accounting framework that adapts to a wide variety of models and theories. The algorithm can be programmed to conveniently and cheaply provide detailed information above forecast revisions, and a version of the algorithm is in use at the Federal Reserve Bank of Minneapolis. This algorithm, or something like it, could be used routinely to help organize the inevitable discussions of forecasts and forecast revisions.

Table 1. An explanation of the revised forecast for December 1988*

	Forecast Variables						
	DOLLAR	SP 500	DEFL	GNP	MB	TBILL (basis pts.)	CBI ($ bils.)
Change in Forecast (%, except as noted)	–4.57	–30.04	–2.01	–1.09	–1.24	–101.88	–9.61
Breakdown by Reason for Change (% pts., except as noted)							
Revisions in Old Data	–0.55	0.27	–0.49	1.70	–0.59	2.86	–1.97
Structural Shocks	–3.81	–29.98	–1.50	–2.77	–0.58	–105.72	–7.33
Miscellaneous Factors	–0.21	–0.33	–0.02	–0.02	–0.07	0.98	–0.31
Breakdown by Variable Changed for Each Reason (% pts., except as noted)							
Revisions in Old Data							
DOLLAR	–0.65	–0.09	0.01	–0.02	0.03	4.19	0.06
SP 500	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DEFL	0.03	0.27	–0.37	0.05	–0.01	–1.58	0.11
GNP	–0.02	–0.13	–0.01	1.71	0.08	3.69	0.49
MB	0.09	0.25	–0.12	–0.04	–0.68	–3.64	–0.88
TBILL	0.00	0.00	0.00	0.00	0.00	0.02	0.00
CBI	0.00	–0.02	0.00	0.00	–0.01	0.18	–1.76
Structural Shocks							
DOLLAR	–3.61	0.87	–0.04	0.01	0.12	3.16	0.45
SP500	–0.30	–32.16	0.12	–2.75	–0.57	–57.16	–5.68
DEFL	0.07	0.80	–1.50	0.29	0.00	–6.85	0.67
GNP	0.00	0.02	0.00	–0.15	–0.01	–0.42	–0.04
MB	0.00	0.00	0.00	–0.01	–0.06	–0.01	–0.06
TBILL	0.02	0.48	–0.08	0.04	–0.05	–44.08	–0.24
CBI	0.00	0.00	0.00	–0.19	0.00	–0.37	–2.43

*The following order of variables was used to identify the model: Dollar, SP500, TBILL, MB, DEFL, CBI, and GNP.

Table 2

An Alternative Explanation of the
Monthly Core Model's Revised Forecast for December 1988*

	Forecast Variables						
	Exchange Index	S&P 500	Deflator	GNP	Base	T-Bill (basis pts.)	Change in Invents. ($ bils.)
Change in Forecast (%, except as noted)	-4.57	-30.04	-2.01	-1.09	-1.24	-101.88	-9.61
Breakdown by Reason for Change (% pts., except as noted)							
Revisions in Old Data	-0.55	0.27	-0.49	1.70	-0.59	2.86	-1.97
Structural Shocks	-3.81	-29.98	-1.50	-2.77	-0.58	-105.72	-7.33
Miscellaneous Factors	-0.21	-0.33	-0.02	-0.02	-0.07	0.98	-0.31
Breakdown by Variable Changed for Each Reason (% pts., except as noted)							
Revision in Old Data							
Dallas Exchange Index	-3.44	0.39	0.02	-0.02	0.14	22.44	0.50
Standard & Poor's Index	-0.31	-31.66	0.09	-2.67	-0.31	-101.64	-5.45
GNP Deflator	0.08	0.82	-1.51	0.16	-0.01	-7.18	0.11
GNP	0.00	0.01	0.01	-0.22	0.08	3.69	0.49
Monetary Base	0.09	0.25	-0.12	-0.04	-0.68	-3.64	-0.88
3-Month T-Bill Rate	0.00	0.00	0.00	0.00	0.00	0.02	0.00
Monthly Change in Business Inventories	0.00	-0.02	0.00	0.00	-0.01	0.18	-1.76
Structural Shocks							
Dallas Exchange Index	-3.44	0.39	0.02	-0.02	0.14	22.44	0.50
Standard & Poor's Index	-0.31	-31.66	0.09	-2.67	-0.31	-101.64	-5.45
GNP Deflator	0.08	0.82	-1.51	0.16	-0.01	-7.18	0.61
GNP	0.00	0.01	0.01	-0.22	-0.01	-0.43	-0.23
Monetary Base	0.00	0.06	-0.08	-0.06	-0.40	-3.93	-0.59
3-Month T-Bill Rate	0.14	0.39	-0.03	0.04	0.01	-14.95	0.01
Monthly Change in Business Inventories	0.00	0.00	0.00	0.00	0.00	-0.03	-2.18

*The following order of variables was used to identify the model: Dollar, S&P 500, T-BILL, MB, DEFL, CBI, and GNP.

REFERENCES

Amirizadeh, H., 'Technical appendix for the Bayesian vector autoregression model of the U.S. economy'. Unpublished manuscript, Federal Reserve Bank of Minneapolis, Minneapolis, MN, 1985.

Caton, C. W., 'Forecast revision, or what to do when the future is no longer what it used to be'. Paper presented to the Seventh International Symposium on Forecasting, Boston, MA, May, 1987.

Doan, T. A., 1988. *User's Manual: RATS Version 3.00*, Evanston, IL: VAR Econometrics, 1988.

Doan, T. A., Litterman, R. B. and Sims, C. A., 'Forecasting and conditional projection using realistic prior distributions', *Economic Reviews*, 3 (1) (1984), 1–100.

Hakkio, C. S. and Morris, C. S., 'Vector autoregressions: A user's guide', Working Paper 84–10, Federal Reserve Bank of Kansas City, 1984.

Litterman, R. B., 'Above–average national growth in 1985 and 1986', *Federal Reserve Bank of Minneapolis Quarterly Review*, 8 (4), (1984), 3–7.

———. 'How monetary policy in 1985 affects the outlook', *Federal Reserve Bank of Minneapolis Quarterly Review*, 9 (4), (1985), 2–13.

Todd, R. M., 'Explaining forecast revisions', Working Paper 355, Federal Reserve Bank of Minneapolis, 1988.

———. 'Algorithms for explaining forecast revisions', Working Paper 459 (D), Federal Reserve Bank of Minneapolis, 1989.

Author's biography: **Richard Todd** is Senior Economist, Federal Reserve Bank of Minneapolis and Associate Director of the Institute for Empirical Macroeconomics, Minneapolis, Minnesota.