

# *Improving Economic Forecasting With Bayesian Vector Autoregression*

Richard M. Todd

*Economist*

Research Department

Federal Reserve Bank of Minneapolis

Economic forecasting is often referred to as an art, perhaps because it involves not only data and groups of equations, or statistical *models*, but also the forecaster's personal beliefs about how the economy behaves and where it is heading at any moment. Artistry is an appropriate metaphor for what economic forecasters commonly do: blend data and personal beliefs according to a subjective, undocumented procedure that other forecasters cannot duplicate.<sup>1</sup> This is not the only way to achieve that blend, however. The Bayesian approach to statistics, a general method for combining beliefs with data, suggests an objective procedure for blending beliefs and data in economic forecasting models. This procedure provides a framework that forecasters can use to document and discuss their beliefs, which can help make economic forecasting more of a science and less of an art.

Today's most widely used economic forecasting models are not usually discussed in terms of the Bayesian approach, but they can be. When viewed this way, most of them seem too rigid, allowing their human managers to express beliefs within the models only in forms that are often too vague or too precise to accurately represent the managers' true beliefs. This may partly explain why these models' forecasts often seem implausible to the models' managers and thus why managers often subjectively adjust those forecasts before presenting them to the models' users.

One type of economic forecasting model, known as the *Bayesian vector autoregression* (BVAR) model, has been developed explicitly along Bayesian lines and seems to be an improvement over other types of forecasting

models. BVAR procedures give modelers more flexibility in expressing the true nature of their beliefs as well as an objective way to combine those beliefs with the historical record. A specific version of the procedures, developed by Minnesota researchers, has been used to build models whose unadjusted forecasts seem to be as accurate as the subjectively adjusted forecasts of other common models. The fact that BVAR models forecast well without subjective adjustment also allows them to objectively estimate answers to questions about future events and probabilities that subjectively adjusted models cannot as plausibly address.

## **A Bayesian View of Economic Forecasting**

All statistical forecasting models combine, in some way, information from historical data with information supplied by the builders of the model. Because modelers must supply at least some information before they examine data and because modelers presumably supply information they think will improve the model's forecasting ability, the information they supply is known as their *prior beliefs* (or simply their *priors*) about the best way to forecast. Forecasting techniques differ in how they represent prior beliefs and how much weight they place on them. BVAR models have been developed to let modelers represent their beliefs more accurately and to combine those beliefs with the information in historical data according to a standard, objective procedure.

---

<sup>1</sup>This practice is well known among economists, but is not often discussed in print. See Litterman and Supel 1983 and Litterman 1984 for some critical comments.

### Beliefs and Forecasts

At an elementary level, all forecasting procedures clearly rely on at least a minimal level of prior belief. For example, to forecast any given variable a modeler must at least suggest which currently known variables are related to it and might therefore be useful in forecasting it. The final forecasting model might not use all of the suggested variables, and data may be used to eliminate some variables from the model. But in picking the candidate variables for the model, the modeler must rely on beliefs derived from prior knowledge of the theory and practice of economics and statistics. The same is true for the types of algebraic formulas that express the relationships among the variables in the model (whether they are linear or quadratic, for example); data may be used to pick the final types of formulas, but only from among a set of candidates previously picked by the modeler. The common transformations of economic data that are used in forecasting models—such as detrending, deseasonalizing, interpolating, and linearizing by means of ratios or logarithms—are also chosen in this way.

Beyond this minimal level of prior belief, the modeler might also have views on the candidate variables or formulas most likely to produce good forecasts or even on the exact numerical relationships between the variables in the model and the variables to be forecasted. One major statistical theory, *Bayesian decision theory*, holds that modelers can be thought of as having prior beliefs in the form of probabilities about which of the possible models will forecast best. Bayesian decision theory explains the best way that data can then be used to revise these prior probabilities. Essentially all forecasting procedures, even supposedly non-Bayesian procedures in which the role of prior beliefs is not at first apparent, can be interpreted as at least approximating a Bayesian procedure in which data are used to revise certain implicit prior beliefs.

### Complete Ignorance

A forecasting model incorporating only minimal prior beliefs can be constructed by selecting a group, or *vector*, of variables to forecast, allowing all the variables to interact linearly with their own and each other's current and past values, and using historical data to determine the quantitative impact that each variable has on its own future values and the future values of the other variables. Because such a model relates future values of a vector of variables to past values of that vector, it is known as a *vector autoregression* (VAR) model. Because it is only

minimally restricted by the modeler's prior beliefs, it is more particularly known as an *unrestricted* vector autoregression (UVAR) model. Economists rarely have enough data to construct UVAR models, especially ones with more than a few variables, that forecast as well as models that supplement the data with more informative prior beliefs.

From a Bayesian point of view, UVAR models more or less let the data speak for themselves. To see this more clearly, consider a simple model built to forecast just the money supply and real output. The UVAR modeler might use economic and statistical knowledge to pick linear formulas for the current and two previous quarterly values of some measure of the money supply (*MONEY*) and real (inflation-adjusted) gross national product (*RGNP*) as well as some constant terms ( $k_m$  and  $k_r$ ). Beyond these minimal and relatively uninformative priors, however, the UVAR modeler would essentially claim to be ignorant. Forecasts of each variable in the model would be based on current and past values of all variables in the model, or

$$(1) \quad \begin{aligned} \text{MONEY}_{t+1} = & k_m + a_0\text{MONEY}_t + a_1\text{MONEY}_{t-1} \\ & + a_2\text{MONEY}_{t-2} + b_0\text{RGNP}_t \\ & + b_1\text{RGNP}_{t-1} + b_2\text{RGNP}_{t-2} + m_{t+1} \end{aligned}$$

$$(2) \quad \begin{aligned} \text{RGNP}_{t+1} = & k_r + c_0\text{MONEY}_t + c_1\text{MONEY}_{t-1} \\ & + c_2\text{MONEY}_{t-2} + d_0\text{RGNP}_t \\ & + d_1\text{RGNP}_{t-1} + d_2\text{RGNP}_{t-2} + r_{t+1} \end{aligned}$$

(where  $m_{t+1}$  and  $r_{t+1}$  represent the errors that will occur when  $\text{MONEY}_{t+1}$  and  $\text{RGNP}_{t+1}$ , respectively, are predicted from a constant term and the three most recent values of both). Furthermore, from a Bayesian point of view, all possible values of the UVAR model's coefficients [the  $k$ 's,  $a$ 's,  $b$ 's,  $c$ 's, and  $d$ 's in equations (1) and (2)], which determine the quantitative impact each variable has on the model's forecasts, would be treated as equally likely. This frequently makes the prior beliefs used in this procedure vaguer than the modeler's true beliefs; most modelers probably don't really believe that

all possible values are equally likely even though the UVAR model is estimated as though they are. For example, a modeler might believe that positive values of  $b_0$ ,  $b_1$ , and  $b_2$  are more likely than negative values to lead to good forecasts. Nonetheless, the final values of these coefficients would be chosen without any restrictions to represent these prior beliefs. A simple statistical procedure (ordinary least squares, or OLS, regression) would pick the coefficient values that best explain the historically observed patterns of interaction among the chosen variables.

Small UVAR models such as the one described above sometimes forecast fairly well, but economists have long recognized that UVAR models with more than a few variables generally do not. Since forecasters are often called upon to forecast large groups of variables and the relationships between them, this is a serious limitation of UVAR models.

The forecasting problems of large UVAR models stem from the fact that economists often have too little data to isolate in a model's coefficients only the stable and dependable relationships among its variables. The statistical procedure used to estimate the coefficients (OLS regression) picks values which best explain the available data, data in which the stable relationships among variables have been obscured by numerous random effects. Furthermore, because current and past values of each variable appear in every equation of a UVAR model, the number of coefficients to be estimated is large compared to the number of observations on the variables. With so many coefficients available to explain so few observations, the estimated coefficients can explain the data very well—too well, in fact. The coefficients are subject to overfitting, a sort of red herring effect. The statistical procedure chooses the coefficients to explain, or *fit*, not only the most salient features of the historical data, which are often the stable, enduring relationships between variables that are most useful for forecasting. The coefficients are so numerous that the statistical procedure can choose them to also fit many of the less important features of the historical data, features which often reflect merely accidental or random relationships that will not recur and are of no use in forecasting. *Overfitting* refers to this incorporation of useless or misleading relationships in the coefficients of a model. It tends to make large UVAR model forecasts inaccurate and overly sensitive to changes in economic variables.

#### *Complete Certainty or Complete Ignorance*

The traditional solution to the overfitting problem of large UVAR models has been to use prior beliefs to reduce the number of coefficients to be estimated. In the *structural econometric* models that are widely used for economic forecasting, this is done by including in each equation of the model only a few variables (or lags of variables) that economic theory suggests are most directly related to the variable that the equation forecasts. (This is an attempt to use theory and statistics to recreate the basic structure of the economy; hence the models' name.) Economic theory is thus the main source of priors in structural models, and these priors are built into the model by excluding most variables from most equations.

Note that excluding variables from an equation amounts to certainty that their coefficients are zero. Certainty is an absolute belief, not subject to revision by any amount of historical evidence. So such *exclusion restrictions* also amount to assigning coefficients of zero to the variables regardless of historical evidence. Although these restrictions can prevent overfitting in a structural model, they are often too rigid to accurately express the modeler's true beliefs and tend to cause useful information in the historical data to be ignored.

A given economic theory may imply that some economic variables are related strongly and others are related weakly or hardly at all. To see how these theoretical implications can be used to overcome overfitting, consider how a modeler using the structural econometric approach might modify the simple UVAR model described above. For example, suppose economic theory and experience led the modeler to believe that next period's real gross national product is mainly affected by that variable's current value and last period's value of the money supply, while monetary policy generally insulates the money supply from changes in real gross national product. These beliefs would lead the structural modeler to modify the UVAR model [equations (1) and (2)] by excluding all values of *RGNP* from the *MONY* equation and excluding all past values of *RGNP* and all values except the previous period's value of *MONY* from the *RGNP* equation. This would be the resulting structural model:

$$(3) \quad \text{MONY}_{t+1} = K_m + A_0 \text{MONY}_t + A_1 \text{MONY}_{t-1} \\ + A_2 \text{MONY}_{t-2} + M_{t+1}$$

$$(4) \quad RGNP_{t+1} = K_r + C_1 MONY_{t-1} + D_0 RGNP_t + R_{t+1}.$$

Even in so small a model as equations (1) and (2), where overfitting is not nearly as serious as in the large models often used by forecasters, exclusion restrictions cut in half the number of coefficients to be estimated. Because of their fewer number, the coefficients of structural econometric models are less likely than UVAR model coefficients to reflect the useless random relationships among variables. This helps explain why structural econometric models have been the dominant form of large forecasting model for many years.

Although the structural econometric approach does reduce the effects of overfitting and has been widely used for forecasting, many modelers are likely to find the exclusion restrictions it is based on too extreme and inflexible. Exclusion restrictions probably either overstate or understate the modeler's prior beliefs about the best forecasting model.

On the one hand, by specifying ahead of time that coefficients on excluded variables will be zero no matter what the historical data suggest, the exclusion restrictions imply that the modeler is absolutely confident that the best coefficients for the variables excluded from the equation are zero. This may often exaggerate the modeler's true confidence. The structural modeler above, for example, might have to choose among competing theories about how real gross national product affects the money supply. Even the modeler's preferred theory is unlikely to suggest such sharp distinctions between the effects of current and past values of money and real output on real output as that model's exclusion restrictions imply. Nonetheless, to reduce overfitting the modeler pretends to know for sure that real output doesn't affect money and that current money and past real output don't affect next period's real output. The data are not allowed to revise these pretended beliefs.

On the other hand, by letting the data completely dictate the coefficients of the included variables, reliance on just exclusion restrictions implies that for those coefficients the structural modeler is as ignorant as a UVAR modeler pretends to be. And here, too, this absolute ignorance about which coefficients lead to good forecasts, or the belief that all of their possible values are equally likely, may often understate the modeler's actual

belief. A structural modeler might observe, for example, that in most countries real gross national product rarely changes more than a few percent per period, even when the money supply has changed a fair amount recently. In the structural model above, these observations might suggest that equation (4)'s coefficient on current *RGNP* ( $D_0$ ) should be close to one and its coefficient on past *MONY* ( $C_1$ ) should be close to zero. Instead, since the modeler's prior beliefs are only used to restrict the coefficients of the excluded variables (to be exactly zero no matter what), the statistical procedures used to estimate the coefficients of the included variables treat values such as 10 or 10 million to be just as likely as one or zero for these variables.

These statistical procedures are generally similar to OLS regression and also basically pick coefficients, at least for the included variables, so as to best explain the historical data. As a result, serious overfitting of structural model coefficients can still sometimes occur. To avoid this possibility, modelers often estimate several versions of a model (for example, by experimenting with different transformations of the data or different sets of exclusion restrictions) and then pick the one whose combination of plausible coefficients and ability to fit historical data seems best. However, it is hard to document, let alone to evaluate, this highly subjective adaptation of the structural econometric approach.

Besides often distorting the modeler's actual prior beliefs, exclusion restrictions also prevent the modeler from even seeing the historical evidence on the relationship between a forecasted variable and a variable excluded from its equation. In the structural model above, for example, suppose that contrary to the structural modeler's prior belief the historical data would strongly suggest that real gross national product is very useful in forecasting the money supply. If the modeler's belief is expressed as a restriction that excludes *RGNP* from the equation for *MONY*, the modeler's chance to be surprised by and learn from the data's unexpected information is sharply limited. (This information will show up only indirectly, if at all, in systematic errors in the model's forecasts or in statistics suggesting that the model has not extracted all the useful information in the historical data.)

#### *A More Flexible Expression of Beliefs*

Many structural models forecast more accurately than large UVARS, but they still don't satisfy many economists. (See, for example, Sims 1980, Lucas and Sargent

1981.) This has led some to suspect that exclusion restrictions derived from imperfect economic theories may be a barrier to improved forecasting and that explicitly Bayesian models with the flexibility to more accurately represent prior statistical and economic beliefs might give better forecasts. Bayesian vector autoregression (BVAR) forecasting models have been developed to test this possibility.

At first glance a simple BVAR model might seem to be no different than a simple UVAR model. In both types of models each variable is allowed to depend on the current and past values of all the variables that the modeler has included in the model. The equations of a BVAR model of the money supply and real output, for example, would have the same form as equations (1) and (2).

Although a BVAR model resembles a UVAR model in the form of its equations, it also differs from a UVAR model—and resembles a structural econometric model—by making heavy use of prior beliefs to reduce overfitting. The sources of the prior beliefs and the ways they are used are generally different in a BVAR model than in a structural model, however. Whereas economic theory is the main source of prior beliefs in structural models, it is often secondary to statistical theory and observations in BVAR models. Furthermore, whereas in structural modeling (at least in its textbook form) each coefficient either is set to zero no matter what the data show or is determined solely by the data no matter what the modeler believes, BVAR modeling doesn't make these sharp distinctions. Instead the BVAR modeler uses prior statistical and economic knowledge to guess which values of all the coefficients will lead to the best forecasts and to specify an extensive system of confidences in each "guesstimated" coefficient. The modeler then uses a statistical procedure to revise these prior beliefs in light of the evidence in the data and thus to override each of the guesses. The extent to which the data are allowed to revise the modeler's guess about a particular coefficient depends on the modeler's initial confidence in the guess: the more confidence, the less weight given to the patterns in the data and vice versa.

Thus, unlike structural modelers, BVAR modelers do not try to avoid overfitting by reducing the number of coefficients. Instead, they try to reach that goal by allowing lots of coefficients but reducing the data's influence on them. As long as the BVAR modeler expresses enough confidence in enough coefficient guesses to significantly limit the revisions that accident-

al patterns in the data can produce, overfitting will generally be limited as well. This is accomplished, however, without preventing important but unexpected historical relationships between variables from being discovered, as when the data strongly override a modeler's guess.<sup>2</sup>

### **A Close Look at a BVAR Procedure: The Minnesota Prior**

Prior beliefs about the coefficients of a forecasting model can come from many sources and take many forms. I have already discussed the sources and forms that are implicitly used in the UVAR and structural modeling approaches. When an explicitly Bayesian approach is used, it becomes clear that the modeler is free to take prior beliefs from many sources and cast them into many forms. Therefore it is impossible to speak of a unique Bayesian or even BVAR approach to forecasting.

In recent years, however, at least one BVAR approach has evolved as an alternative to structural econometric modeling for a variety of economic forecasting applications. This approach is more objective and reproducible than most procedures for combining beliefs and data in forecasting, and it has been described, documented, and evaluated by several researchers. (For example, see Litterman 1980; Kinal and Ratner 1983; Hoehn, Gruben, and Fomby 1984; Doan, Litterman, and Sims 1984; and the technical appendix to the outlook paper by Litterman in this issue, available on request to the Research Department, Federal Reserve Bank of Minneapolis.) Although this approach was developed with the overfitting problems of large forecasting models in mind, its essential elements can be illustrated with a BVAR version of the simple money supply and real output model of the previous sections. As yet, this particular system of Bayesian priors has not acquired a particular name to simplify discussions of it. Since it has been developed and used mainly by economists associated with the University of Minnesota and the Federal Reserve Bank of

<sup>2</sup>The distinctions between structural and BVAR models are sometimes blurred, but usually only slightly. Structural modelers, for example, sometimes also use nonexclusion restrictions and flexible degrees of confidence, though rarely as extensively as BVAR modelers. Exclusion restrictions are also sometimes used in BVAR models, though less extensively and for different reasons than in structural econometric models. Even with their exclusion restrictions, some BVAR models currently used for forecasting still have dozens or even hundreds of variables (including lagged variables) in each equation rather than just the few variables that structural model equations include. (See the outlook papers by Litterman and by Amirzadeh and Todd in this issue, for example.) In BVAR models exclusion restrictions are used mainly to reduce the costs of building and using the models, not to avoid overfitting.

Minneapolis, I will call it the *Minnesota system of prior beliefs* or, even more briefly, the *Minnesota prior*.

#### *Taking a Random Walk, With Confidence*

In general terms we have seen that a Bayesian modeler's prior beliefs take the form of probabilities about which of the possible forecasting models will forecast best. The first step in using the Minnesota prior is the common one of limiting the set of possible models by choosing a group of variables to include in the model and specifying that they will be linked by linear equations. As usual, the choice of variables is dictated in part by which variables the modeler wants to forecast and is also guided by economic reasoning and experience concerning which other variables are available that might be closely related to the variables to be forecasted. In the simple model discussed above [equations (1) and (2)], for example, *MONY* and *RGNP* are to be forecasted by using the current value and the two past values of each (and a constant term). Once the variables have been chosen, the prior beliefs concern the values of the coefficients of each of the variables in each of the linear equations of the model, and they can be expressed in the form of probabilities about which set of values will give the best forecasts. In the Minnesota prior, these probabilities can be described by assigning certain numbers—mainly a best guess and a measure of confidence—to each coefficient in the model. A Minnesota prior for the *MONY-RGNP* model would include a best guess of the coefficients  $k_m, k_r, a_0, a_1, a_2, b_0, b_1, b_2, c_0, c_1, c_2, d_0, d_1,$  and  $d_2$  in equations (1) and (2) as well as a quantitative expression of the modeler's confidence in each of these guesses.<sup>3</sup>

In the Minnesota prior, the best guesses of the coefficients are usually set either exactly or approximately according to the *random walk hypothesis*. This hypothesis capitalizes on a simple statistical observation that is often a forecaster's chief source of embarrassment: many economic (and other) variables seem to behave as though changes in their values are completely unpredictable. For such a variable, the best forecast of its future values is just that they will equal its current value. Even for variables whose changes are thought to be partially predictable, these no-change forecasts can be surprisingly difficult to improve upon.

To implement the random walk hypothesis, the best guesses of the Minnesota prior are that, with one exception, all the coefficients in the equation for any given

variable are zero. The exception is the coefficient on the most recent value of the given variable, and that is guessed to be one. In the *MONY-RGNP* model, this would mean setting the best guesses of  $a_0$  and  $d_0$  at one and those of all the other coefficients at zero. If all these guesses were right, future values of the variables would differ from their current values only because of completely unpredictable random events [represented by  $m_{t+1}$  and  $r_{t+1}$  in equations (1) and (2)].

Since it is an explicitly Bayesian procedure, the Minnesota prior does not place unlimited confidence in the best guesses derived from the random walk hypothesis. Instead the modeler must supply a quantitative measure of confidence in each best guess. Although these measures can be expressed in many equivalent ways, in the Minnesota prior they are usually documented and discussed in terms of what a Bayesian statistician would call the *prior variance of the coefficient*. In less technical (and somewhat loose) language, this measures how likely it is that the coefficient is actually close to the best guess and, in particular, how far above and below the best guess the modeler would have to go before being willing to place 2-to-1 odds on the coefficient actually lying between those values. (The distance away from the best guess in either direction equals the square root of the prior variance, which is known as the *prior standard deviation*.) A small prior variance, or equivalently a narrow 2-to-1 *confidence band*, indicates that the modeler is very sure that the coefficient that gives the best forecasts is close to the best guess. A wide band indicates that the best coefficient could easily be very far away from the modeler's best guess. Narrowing the band, or decreasing the prior variance, is known as *tightening* the prior (around the best guess). Similarly, *loosening* the prior involves widening the band, or increasing the prior variance. A complete description of the modeler's degree of confidence in the best guesses would determine the prior variance of the coefficient of each variable in each equation.

Some BVAR forecasting models have hundreds of coefficients, which means that the number of prior variances to choose is so large that deliberating about each individually is impractical for the modeler. The

<sup>3</sup>In technical terms, the modeler must specify a *joint probability distribution* for the coefficients of the model, and the Minnesota prior assumes that it is a *multivariate normal* distribution. Many of the means and covariances of this prior distribution are often set at zero, but the variances are generally positive, as indicated later.

Minnesota prior includes a system which makes such deliberation unnecessary. It approximates a full set of prior variances almost automatically once the modeler has chosen some of their key features. With one exception, this system proceeds in two stages. First the modeler selects a few restrictions that group the prior variances and mainly determine the relative sizes of the prior variances within each group. Then, for each of those groups, the modeler selects a range of possible values for a scale factor that completes the determination of the prior variances.

The one exception to the two-stage procedure for determining prior variances concerns the prior variances of the constant terms in each equation. [These would be the prior variances of  $k_m$  and  $k_r$  in the *MONY-RGNP* model of equations (1) and (2).] These variances are often simply set to very large numbers, which amounts to saying that at least over a very large range the modeler regards all possible values of the constant term as almost equally likely and is willing to let the constant term be determined by the data alone. The constant term represents the average increase per period, or the *drift*, in the variable.<sup>4</sup> Because drift is determined by the data with almost no influence from the modeler's guess of zero for the constant term, the best guesses of the Minnesota prior are sometimes described not as just the random walk hypothesis but as the *random walk plus drift hypothesis*.

#### *Relative Degrees of Confidence*

Specification of the rest of the prior variances begins with the selection of restrictions on how the variances are related to each other. BVAR modelers use many different types of restrictions. (See Doan, Litterman, and Sims 1984 for many examples.) Most modelers, however, use two basic and closely related types. Both of them are motivated by the notion that the highly lagged (less recent) values of a variable are less likely to be useful in forecasting than the less highly lagged (more recent) values.

In the equation that forecasts any given variable, the first restriction takes the form of weights that shape the prior variances of the coefficients of current and past values of the given variable. These values are known as the *own lags* (of the variable that the equation forecasts). In the simple *MONY-RGNP* model they are  $MONY_t$ ,  $MONY_{t-1}$ , and  $MONY_{t-2}$  in the money supply equation and  $RGNP_t$ ,  $RGNP_{t-1}$ , and  $RGNP_{t-2}$  in the real output equation.

The Minnesota prior asserts that the less important a variable is believed to be for forecasting, the greater is the modeler's confidence in the best guess of its coefficient. Since the more recent values of a variable are considered more important for forecasting than the less recent ones, the prior variances on the own lags should get smaller, or tighter around the best guess, as the lags become longer (that is, as the variable becomes less recent). As lag length increases, this feature of the Minnesota prior combines with the random walk best guess to express increasing confidence that zero coefficients for own lag variables will lead to good forecasts. The restriction is imposed by weighting each own lag prior variance by  $1/(k+1)$ , where  $k$  is the length of the lag (the number of periods before the current period). Making the own lag prior variances proportional to  $1/(k+1)$  means that, in the *MONY* equation, for example, the prior variances of the coefficients of  $MONY_{t-1}$  and  $MONY_{t-2}$  are one-half and one-third, respectively, as large as the prior variance of the coefficient of  $MONY_t$ .

In the equation that forecasts any given variable, the second restriction takes the form of weights that shape the prior variances of the coefficients of current and past values of all the variables besides the given variable. These values are known as the *cross lags* (of the variable that the equation forecasts). In the simple *MONY-RGNP* model they are  $RGNP_t$ ,  $RGNP_{t-1}$ , and  $RGNP_{t-2}$  in the money supply equation and  $MONY_t$ ,  $MONY_{t-1}$ , and  $MONY_{t-2}$  in the real output equation. The prior variances of the coefficients of the cross lag variables have the same relative sizes as the coefficients of the own lag variables. In particular, they are weighted by  $1/(k+1)$ , so that their prior variances also get tighter (smaller) as the length of their lag increases (and hence as their probable importance for forecasting declines). They are also each weighted by an own-versus-cross variance factor, which gives the cross prior variances units comparable to those of the own prior variances.<sup>5</sup>

<sup>4</sup>For the many variables in logarithmic form in most BVAR models, the constant term represents average percentage increase per period.

<sup>5</sup>The own-versus-cross variance weight is  $s_o/s_c$ , where  $s_o$  and  $s_c$  are the standard errors of estimate from regressions of the own and cross variable, respectively, on several of their past values. It is independent of lag length but specific to a given cross variable. It becomes especially important in larger models, where each equation would have not just one cross variable (and its lags) but rather many. Then the prior variance of each lag of each cross variable is converted to units comparable to the own lag prior variances by the own-versus-cross variance factor ( $s_o/s_c$ ) specific to that cross variable.

The combined effect of the random walk best guesses and the first stage of the determination of the confidence levels is illustrated for the money supply equation of the *MONY-RGNP* model in the accompanying figure. The curves show the relative probabilities attached to the coefficients of the *MONY* and *RGNP* variables in that equation. The curve for the current value of *MONY* has its highest value—the highest probability among the possible values of its coefficient—at one, which is the best guess for that variable. Similarly, the curves for the other variables have their highest values at zero, the best guess for those variables. The curves for current values of variables are broad and low, which means that a wide range of possible values for the coefficient has prior probability not much lower than the best guess and that even values fairly far from the best guess are not considered to be extremely unlikely. The curves for the lagged values of variables become progressively more peaked and concentrated around the best guess, which puts low probability on values outside of that narrow range. This reflects the fact that as lag length increases the modeler becomes increasingly confident that a zero coefficient will be consistent with a model that forecasts well.

#### Absolute Degrees of Confidence

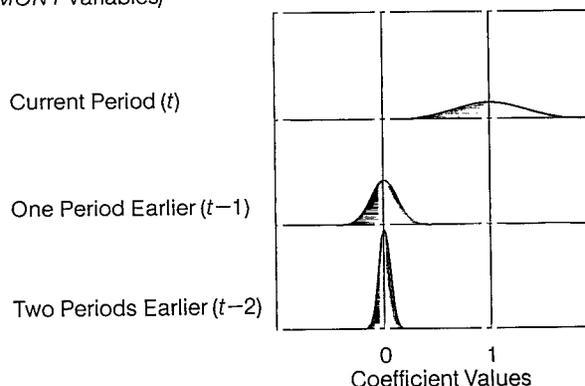
Restrictions such as the two described above determine many, but not all, features of the prior variances of the coefficients in the model. In particular, they define a few large groups of coefficients and, within each group, determine how the coefficients' prior variances are related to each other. In the *MONY-RGNP* model above, I have defined two groups—the own lag coefficients and the cross lag coefficients—and have picked weights that determine the relative sizes of the prior variances within each group.<sup>6</sup>

Once this has been done, all the modeler would need to do to complete the specification of the prior variances is to pick one number, a scale factor called a *hyperparameter*, for each group of coefficients. That hyperparameter would simultaneously multiply all the weights assigned to coefficients in the group and convert these weights from relative to absolute prior variances. For example, if in the *MONY-RGNP* model the hyperparameters  $H_o$  and  $H_c$  were assigned to the own and cross variable groups of prior variances, respectively, then together with the own and cross lag weights already chosen they would give the prior variances shown in the accompanying table. (In the

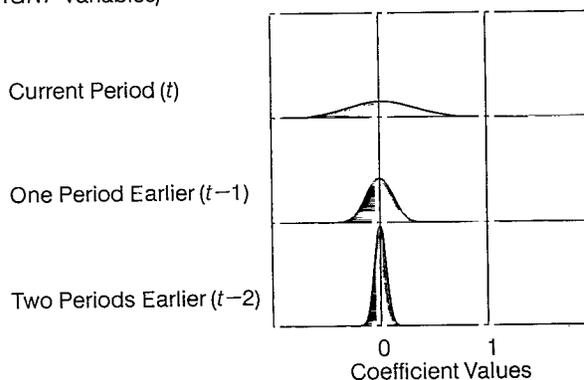
### The Minnesota Prior's Relative Degrees of Confidence in Its Random Walk, Best-Guess Coefficients

Hypothetical Probabilities on the Coefficients in the Money Supply Equation (1)

Coefficients on Own Lags  
(*MONY* Variables)



Coefficients on Cross Lags  
(*RGNP* Variables)



table,  $f_m = s_m/s_r$  and  $f_r = s_r/s_m$ , where  $s_m$  is the standard error of estimate in a regression of *MONY* on several of its past values and  $s_r$  is the corresponding figure for

<sup>6</sup>In this example, the groups do not overlap, but in real BVAR models this is not necessary and often not true. See Doan, Litterman, and Sims 1984, for example.

RGNP.) To be more specific, if  $H_o$  equaled  $\frac{1}{8}$  and  $H_c$  equaled 2, the own lag weights in the table would become  $\frac{1}{8}$ ,  $\frac{1}{16}$ , and  $\frac{1}{24}$  while the cross lag weights would become 2, 1, and  $\frac{2}{3}$ . Assigning hyperparameters to each group of coefficients is the second stage in specifying prior variances according to the Minnesota prior.

The second stage would be very simple—and the entire procedure a standard Bayesian one—if the modeler had firm beliefs about the hyperparameters, or in other words was certain of the absolute size of at least one of the variances within each group of relative variances. In that case, as in the *MONY-RGNP* example, the appropriate hyperparameter would be assigned to each group, completing the specification of the prior probabilities (best guesses and variances) of the model's coefficients. Then the data would be examined with standard Bayesian statistical procedures in order to revise these prior coefficient probabilities and determine the final forecast.

The Minnesota prior, however, generalizes the standard Bayesian approach by not requiring that the modeler have firm beliefs about the hyperparameters. The modeler determines many key features of the prior coefficient probabilities with the best guesses and the restrictions among groups of prior variances. Once that is done, the modeler using the Minnesota prior can claim to be almost ignorant about the absolute levels of the various groups of prior variances. This adds another layer of prior probabilities—probabilities about the few hyperparameters that determine those absolute levels—to the normal Bayesian procedure. In other words, instead of picking a single probability distribution for the model's coefficients, the modeler specifies a group of similar probability distributions, one for each setting of the hyperparameters, and treats all the distributions within the group as equally likely. Ideally, standard Bayesian statistical procedures would then be applied to the data to compute revised (*posterior*) coefficient probabilities for each possible setting of the hyperparameters. The final coefficient probabilities (and hence the final forecast) would be formed as a weighted average of these, with the weight attached to each proportional to the probability that the setting of the hyperparameters that generated it is consistent with the historical data. (See Doan, Litterman, and Sims 1984.)

In principle, adding this ideal weighting process to standard Bayesian statistical procedures should not be difficult, but the great expense of actually computing the average has led most BVAR modelers to use a cheaper

### Moving From Relative to Absolute Degrees of Confidence in a Model's Coefficients

Hypothetical Prior Variances for the Coefficients in the *MONY-RGNP* Model\*

Variables	Equations	
	(1) Money Supply	(2) Real Output
Money Supply Variables ( <i>MONY</i> )		
Current Period ( $t$ )	$H_o$	$H_c f_r$
One Period Earlier ( $t-1$ )	$H_o(\frac{1}{2})$	$H_c(\frac{1}{2}) f_r$
Two Periods Earlier ( $t-2$ )	$H_o(\frac{1}{3})$	$H_c(\frac{1}{3}) f_r$
Real Output Variables ( <i>RGNP</i> )		
Current Period ( $t$ )	$H_c f_m$	$H_o$
One Period Earlier ( $t-1$ )	$H_c(\frac{1}{2}) f_m$	$H_o(\frac{1}{2})$
Two Periods Earlier ( $t-2$ )	$H_c(\frac{1}{3}) f_m$	$H_o(\frac{1}{3})$

\*This is the general formula for the prior variances:

$$\left[ \begin{array}{l} \text{Absolute} \\ \text{Confidence} \\ \text{Scale Factor:} \\ \text{Hyperparameter} \\ (H_o \text{ or } H_c) \end{array} \right] \times \left[ \begin{array}{l} \text{Relative} \\ \text{Confidence} \\ \text{Weight:} \\ 1/(k+1) \\ (k = \text{lag length}) \end{array} \right] \times \left[ \begin{array}{l} \text{Own vs. Cross} \\ \text{Variance} \\ \text{Factor:} \\ f_m \text{ or } f_r \\ \text{(on cross lags only,} \\ \text{to make units} \\ \text{comparable)} \end{array} \right]$$

approximation. Instead of averaging all the standard models that come from all the possible settings of the hyperparameters, they simply try to find the one set of hyperparameters that leads to the best mock forecasts of the historical data. Under certain assumptions the forecasting model and revised probabilities associated with the best mock forecasts will be close to those formed by the ideal weighted average (Doan, Litterman, and Sims 1984). One of those assumptions is that the modeler's beliefs can be thought of (at least approximately) as defining a particular range of hyperparameter values: all values within the range are believed to be equally likely and all values outside it highly unlikely. Another assumption is that within that range the quality of the mock forecasts does not vary too much. This second assump-

tion frequently seems to be satisfied over ranges broad enough to include most plausible hyperparameter values. (See Doan, Litterman, and Sims 1984 for one example.) That has made it relatively easy for several BVAR modelers to accept the first assumption as well.<sup>7</sup>

#### *Finding the Final Model*

In practice, then, the second stage of specifying prior variances according to the Minnesota prior merges into the selection of a final forecasting model; these are simultaneously determined when the modeler finds the hyperparameters that lead to the model that seems to give the best forecasts. To judge the forecasts associated with a given setting of the hyperparameters, the modeler repeatedly replicates how a model based on those hyperparameters would have forecasted in the past.<sup>8</sup> To find the setting that seems to lead to the best forecasts, the modeler simply uses trial and error, testing many settings and picking the one that leads to a model whose replicated forecasting errors are smallest (according to whichever of several standards of accuracy the modeler prefers). The final forecasting model, computed as part of the trial and error process, uses all available historical data to revise the prior probabilities associated with the best setting of the hyperparameters.

The key step in finding the best setting of the hyperparameters is replicating how a model based on any given set of hyperparameters would have forecasted. Because BVAR models are linear and can easily be reestimated when new data become available, BVAR modelers can address this question by simply recreating hundreds of past forecasts.

Suppose that a BVAR modeler had quarterly data from the first quarter of 1955 to the fourth quarter of 1984, or a total of 120 observations, on the variables in the *MONY-RGNP* model. For any given set of hyperparameters, the modeler could use the first 20 observations (the first quarter of 1955 through the fourth quarter of 1959) to estimate a BVAR model, use that model to forecast the 21st observation (the first quarter of 1960), and then compute the difference between the forecasted and actual values of the 21st observation. This difference is known as the *one-step-ahead forecast error* for the 21st observation, because it is the error in a forecast that extends one quarter beyond the period of historical data used to estimate the forecasting model. It replicates the forecast that a BVAR modeler could actually have made in early 1960, as soon as data for the fourth quarter of

1959 became available. Next the BVAR modeler would use the first 21 observations to compute a one-step-ahead error for the 22nd observation, the first 22 observations to compute a one-step-ahead error for the 23rd observation, and so on until all the observations had been used to compute, in this case, 100 one-step-ahead forecast errors for the given setting of the hyperparameters. The same number of one-step-ahead forecast errors would then be computed for each of a number of other possible settings of the hyperparameters. The setting with the smallest average one-step-ahead forecast errors (by any of several measures) could then be selected as the best, or the modeler might also examine similarly computed two-, four-, or eight-step-ahead forecast errors before selecting the best setting of the hyperparameters. This would complete the selection of the hyperparameters as well as of the final forecasting model.

Once the hyperparameters that lead to the best forecasting model have been chosen, they are usually reevaluated only every few years or, more likely, when a change is made in the model, such as the addition or deletion of a variable. More frequent updating of the hyperparameters is expensive, and experience in searching for the best settings of the hyperparameters suggests that it would probably make little difference in the average accuracy of the forecasts.

Although the hyperparameters are infrequently revised, the coefficients associated with the chosen hyperparameters are routinely updated every period, as new data become available. The linear relationships between variables in a BVAR model make this simple and inexpensive.

The Minnesota system of prior beliefs is not simple to use, but it does give forecasters a flexible way to express prior beliefs and an objective procedure for combining those beliefs with historical data to produce forecasts. In

---

<sup>7</sup>An alternative, non-Bayesian interpretation of picking the hyperparameters that lead to the best mock forecasts is that BVAR modelers' restrictions on the relative sizes of prior variances already express a set of prior beliefs strong and clear enough to avoid overfitting. Having done this, the BVAR modeler can just pick the model that seems to forecast most accurately (according to tests on the historical data) without being too concerned that the model's coefficients have been contaminated by accidental patterns in the historical data.

<sup>8</sup>Note that, in the Minnesota prior, hyperparameters are picked according to how well the coefficients they lead to forecast data that were not used to estimate the coefficients. This is known as *out-of-sample* forecasting. A more common procedure for choosing among models is *in-sample* forecasting, where models are judged by how well they explain the same data used to estimate their coefficients. That is not as effective as out-of-sample forecasting at weeding out models whose coefficients have been overfit to the data.

that sense it represents a move away from traditional forecasting procedures toward Bayesian procedures that seemed, by the late 1970s, more promising to at least some economists.

### **Improved Method, Improved Results**

The accuracy of a forecasting model is generally assessed by some measure of the average distance between its guesses of what the future is most likely to be and what the future actually turns out to be. Although the evidence is limited, BVAR models' guesses about the most likely future appear to be at least as accurate as those of competing forecasting procedures.

As noted earlier, structural econometric models are frequently not accurate enough to satisfy the modelers who manage them, and many of these managers routinely use their own subjective views about where the economy is heading to adjust their models' forecasts. Because of this subjectivity, good evidence on the relative accuracy of forecasts by BVAR models and forecasts by structural modelers is hard to come by. Good evidence would consist of, for example, a history of thousands of comparable forecasts from both of these sources. At first thought it might seem that such evidence could be synthesized by doing with structural models what BVAR modelers do—repeatedly using data up to some previous date to estimate the coefficients of the model, pretending to forecast the ensuing periods, and comparing these pretended forecasts to the actual historical data. However, this would not accurately reproduce structural modelers' forecasts. It omits the subjective adjustment they would have made to the model's forecasts. That adjustment cannot be realistically recreated with hindsight.

Instead of using synthetic forecasting records to compare the forecasts made by BVAR models and structural econometric modelers, generally we must look to the history of their actual forecasts. This history is far too short to give decisive evidence, and the fact that the subjective component in the structural forecasts takes on new properties at least every time a new person begins to manage the model means that in some sense it will always be too short. Nonetheless, the historical record gives no reason to regard BVAR forecasts as less accurate. The longest record of reasonably comparable forecasts includes national economic forecasts from several well-known structural econometric forecasting firms as well as from the first BVAR prototype, a six-variable model constructed at the Federal Reserve Bank of Minneapolis

in the late 1970s. This model pioneered the Minnesota prior in an attempt to investigate the usefulness of the BVAR technique. Forecasts from that BVAR prototype, regularly computed over a four-year period with no subjective adjustment or respecification of the model, compare favorably with those from the structural forecasters. Out of a total of over 1,100 forecasts, the BVAR model was most accurate for 39 percent; the next-best model was most accurate for only 23 percent. (See Litterman 1984.)

The explicit documentation of BVAR statistical procedures means that the forecast performance of BVAR models is more likely to improve than that of structural econometric forecasters. Unlike the highly personal procedures for adjusting structural forecasts, the objective BVAR procedures can be improved by economic research and accumulated forecasting experience, and this is already happening. Recent BVAR models of the national economy are much more sophisticated and likely to be accurate than the BVAR prototype that outperformed the commercial forecasting models. (See Doan, Litterman, and Sims 1984.)

If accuracy of guesses about the most likely future were the only criterion, some might consider the advantages of BVAR models modest. However, because BVAR models can forecast a large group of variables relatively accurately without subjective adjustment, they are also likely to be more informative than other forecasting procedures. They probably can produce more accurate answers to a wider variety of questions about complicated or unobservable features of the future economy.

Examples of such questions are, How would an unexpected change in the money supply affect real output? And, What are the odds of a recession occurring next year?<sup>9</sup> The accuracy of any answers to such questions is obviously very hard to measure. But a BVAR model's answers are probably more accurate than a structural model's. A BVAR model produces these types of forecasts using only current data and the same explicitly documented statistical procedure whose measured accuracy compares well with that of other procedures. By contrast, as we have seen, a structural model's forecasts—or at least the forecasts that the model's managers actually give to its users—often include adjustments that

<sup>9</sup>For examples of the use of BVAR models to answer such questions, see the other papers in this issue.

the managers make according to an undocumented and highly personal procedure. Even if the managers appear to be accurate forecasters of simple observable future values, therefore, there is no guarantee that the subjective adjustments they would make to forecast complicated economic interactions or unobservable probabilities would be determined by procedures consistent with those they use to generate their standard forecasts, whose accuracy can be measured. The procedures they use to forecast anything more complicated than a most likely value are not only generally undocumented but also probably untested and almost untestable.

These abstract arguments for the greater informativeness of BVAR models are mild compared to arguments based on the practical ways BVAR and structural models are actually used. With BVAR models, the same model and explicit statistical procedures that produce its accurate forecasts of the most likely values of variables are routinely used to generate answers to a wide variety of more complicated questions, about future probabilities and hypothetical relationships among variables. Structural econometric modelers often avoid forecasting these important but complicated events or forecast them in ways that defy a rigorous statistical interpretation. (See Litterman and Supel 1983.)

### Summary

Research on BVAR models and the Minnesota prior was stimulated in the 1970s by dissatisfaction with the methods and results of structural econometric forecasting and by hopes that explicitly Bayesian forecasting procedures could improve both. Experience with BVAR models is still limited, but they seem to be realizing the initial hopes to some degree. Forecasts straight from BVAR models seem to be at least as accurate as subjectively adjusted forecasts based on structural econometric models. The ability of BVAR models to forecast accurately without adjustment means they can generate objective answers to complicated questions that structural econometric forecasts usually avoid or evade. The methods used to construct BVAR models combine flexible—and therefore probably more accurate—forms for expressing and documenting personal beliefs about the economy with objective, reproducible statistical procedures for combining those beliefs with historical data. This has made these forecasting procedures relatively open to scientific examination and discussion, a powerful process which has led to some useful refinements already and may well lead to many more.

### References

- Doan, Thomas; Litterman, Robert; and Sims, Christopher. 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews* 3 (1): 1-100.
- Hoehn, James G.; Gruben, William C.; and Fomby, Thomas B. 1984. Some time series methods of forecasting the Texas economy. Preliminary Working Paper. Federal Reserve Bank of Dallas.
- Kinal, Terrence, and Ratner, Jonathan. 1983. A VAR forecasting model of a state economy: Its construction and comparative accuracy. Department of Economics, State University of New York at Albany. Manuscript.
- Litterman, Robert B. 1980. A Bayesian procedure for forecasting with vector autoregressions. Working Paper. Massachusetts Institute of Technology. Available from the Research Department, Federal Reserve Bank of Minneapolis.
- \_\_\_\_\_. 1984. Forecasting with Bayesian vector autoregressions—Four years of experience. Paper presented at the American Statistical Association's Annual Meeting, sponsored by the Business and Economic Statistics Section and held at Philadelphia, Penn., August 12-14. Also, forthcoming in proceedings volume and as Research Department Staff Report, Federal Reserve Bank of Minneapolis.
- Litterman, Robert B., and Supel, Thomas M. 1983. Using vector autoregressions to measure the uncertainty in Minnesota's revenue forecasts. *Federal Reserve Bank of Minneapolis Quarterly Review* 7 (Spring): 10-22.
- Lucas, Robert E., Jr.; and Sargent, Thomas J. 1981. Introduction. In *Rational expectations and econometric practice*, vol. 1, pp. xi-x1. Minneapolis: University of Minnesota Press.
- Sims, Christopher A. 1980. Macroeconomics and reality. *Econometrica* 48 (January): 1-48.

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