Federal Reserve Bank of Minneapolis
Research Department

Implementing Bayesian Vector Autoregressions

Richard M. Todd*

Working Paper 384

Revised May 1988

Presented to International Association of Business Forecasters, Annual Conference on Business Forecasting, Pittsburgh, Pennsylvania, December 4, 1987.

---

*Federal Reserve Bank of Minneapolis.

Economic forecasting models are based on a combination of human insight and statistical analysis, but the question of how best to combine theses elements remains, in practice, unresolved. Their combination cannot be avoided. Given the limitations of human brains and economic databases, the potential linkages among economic variables are too numerous to be accurately gauged by either human reasoning or statistical analysis alone. The traditional solution to this degrees of freedom problem has been, more or less, to use reasoning to select a small set of coefficients thought to represent the most important linkages among variables, to assume all other linkages can be ignored (by fixing their coefficients at zero), and then to estimate the chosen coefficients from the data. Bayesian statistical theory suggests a different approach. Include as many linkages--and thus coefficients--as your computer can handle. Then use human insight to overcome the degrees of freedom problem, by allowing the data and a set of prior beliefs to jointly determine the estimated values of all the coefficients.

This paper discusses how the Bayesian approach can be used to construct a type of multivariate forecasting model known as a Bayesian vector autoregression (BVAR).[1] Since the idea of specifying prior beliefs about the numerous coefficients of a multivariate forecasting model is rather daunting, the key to the BVAR approach is to simplify this task. To that end, Doan, Litterman, and Sims (1984) have proposed that a certain family of prior probability distributions, indexed by a fairly small set of so-called hyperparameters, can adequately represent modelers'

prior beliefs in most cases. They have shown how to estimate a BVAR, either when the modeler picks a particular member of their family of prior distributions or in the not uncommon case where the modeler has no strong beliefs (i.e., flat priors) about which members of this family of possible prior distributions are most appropriate. The explanation of their suggestions, in Section 2, forms the core of this paper. However, the BVAR approach to economic forecasting has implications, as well as a set of conventional procedures, for how the data in a model are selected and transformed. For that reason, and for completeness, Section 1 discusses how to specify a BVAR and set up a BVAR database. A 4-variable model is used to illustrate the BVAR approach.

Section 2: Specifying a BVAR

In the simplest form of a BVAR, n lags of each variable in the model appear in each equation. In that case, model specification consists of picking n and a data series for each variable. Even more complicated BVARs generally consist of blocks of equations that have the simple form or slight variations of it, so the links between model specification and database preparation remain quite direct. For that reason, I will discuss the basics of specifying a BVAR through a step-by-step discussion of the setting up a BVAR database. In many ways the considerations and conventions used by BVAR modelers are the same as those used by other forecasters, except that somewhat different criteria are used in deciding which variables to include in the model.

What is the purpose of the model? The first step in specifying any forecasting model, including a BVAR, is to state as

precisely as possible the goal of the forecasting exercise. At a minimum, the forecaster presumably wishes to produce accurate forecasts of at least one variable, and often more. I will refer to these variables as the variables of interest. They should be identified at the outset, and a loss function over their forecast errors could also be specified.

Frequently the forecaster wants not only to forecast accurately but also to claim that his or her forecasts incorporate the effects of one variable on another. (Note that this desire may be independent of, or even conflict with, the desire to forecast a set of variables accurately.) After October 19, 1987, for example, forecasters had to indicate whether their forecasts reflected that day's crash in the stock market. Incorporating such cross variable linkages rules out one fairly effective class of pure forecasting models--those which forecast each variable of interest by means of a univariate ARIMA model. Instead the modeler will need a multivariate model to estimate the linkages among the variables of interest and/or between those variables and a set of other variables which affect the variables of interest but are not of direct interest themselves. I will refer to these other variables as related variables.

The forecaster may be even more ambitious and wish not only to incorporate but also to separately identify the causal linkages among variables. For example, forecasters may wish to say not only that their forecasts reflect the stock market crash but also that they know by just how much the crash affected each variable's forecast. In this case the estimated multivariate

model will have to be augmented by a set of identification restrictions on the model's coefficients and/or covariance matrices. These additional restrictions provide a unique relationship between the forecasting model and a model of the underlying structure of the economy. Meaningful what-if questions can then be posed in the structural form of the model and traced through to the corresponding forecasting model, producing what are known as conditional forecasts. To estimate the effects of the stock market crash, for instance, the forecaster could use the structural model to pose the question of what the current outlook would be if the stock market had held steady at its October 18th level. Without identifying restrictions, what-if questions can't be posed, and the forecasts are said to be unconditional. Modelers can say that their unconditional forecasts incorporate the effects of the stock market crash,[2] but they can't say just what those effects are.

For the purposes of this paper, I will assume that my goal is to accurately forecast, in as much detail as possible, economic activity in the state of Minnesota. I will also assume that my clients, whom I take to be the general public, demand multivariate forecasts (whether or not these are more accurate than multiple univariate forecasts) that incorporate the linkages among Minnesota variables and between Minnesota and national variables. However, my clients are content with unconditional forecasts, allowing me to defer the complexities and controversies of how to impose identifying restrictions on BVARs.

How many variables and linkages can I include in the model? Ideally, my clients would like forecasts of almost every Minnesota economic variable. Furthermore, they want each variable's forecast to incorporate the effects of all the other variables. In non-Bayesian modeling, this ideal cannot be met because of insufficient degrees of freedom. Because there are many data series, the number of potential linkages among them, and hence the potential number of model coefficients, is huge. But each data series is fairly short, so that the number of observations per coefficient in an equation is small, possibly less than one. To conserve degrees of freedom, some variables are left out of the model, and many potential linkages among the remaining variables are turned off by assuming that the coefficients that represent them are zero.

In BVAR modeling, concerns over degrees of freedom play a much less important role in deciding which variables and linkages among variables to include in the model. Instead modelers must worry more about the availability of computing power, support staff, and their own time and energy in deciding whether to enlarge their models.

BVAR models can easily demand more memory than many computers and software packages allow. The prior information a BVAR modeler supplies to the estimation program includes a variance-covariance (VCV) matrix for each equation. Since BVARs typically include several lags of many variables in each equation, the number of coefficients, k, in a given equation can be large by non-Bayesian standards. Since the VCV matrices, in turn, are

k x k, the memory required to store them may exceed the limits of many computer systems (microcomputers, for example) or of software oriented to non-Bayesian methods. For example, although the RATS software available on my bank's mainframe is designed for BVARs and accommodates a fairly generous number of coefficients per equation--about 200--this limit is binding at times and has dictated a block recursive structure for some of the models in use there. On a personal computer, RATS can handle about 50 coefficients per equation. For a quarterly model with 6 lags of each variable in each equation, this imposes a limit of about 9 variables.

BVARs can also put large demands on support staff and on the modeler's time and energy. Whenever a large model is specified, someone has to assemble and maintain the database. This takes a lot of time, or money, or both. In addition, estimating a BVAR requires some programming and generates a lot of output that needs to be analyzed. My impression is that, variable for variable, BVARs are probably much easier to estimate and maintain than many econometric alternatives. Nonetheless, they do require support, and in practice limits on these resources may also constrain the size of BVARs.

Degrees of freedom do play a role in choosing variables for a BVAR. Although they ease the degrees of freedom problem, the prior beliefs imposed on the coefficients of BVAR models express a degree of uncertainty about the value of the coefficient. The more data are available, the more this uncertainty can be reduced. In cases where the prior uncertainty is large and the

data series are short, some variables may have to be dropped from a BVAR to conserve degrees of freedom. I have estimated 6-variable, 6-lag BVARs with as little as 10 years of monthly economic data, but where longer data series can be substituted I would normally exclude business cycle variables with less than 15 or 20 years of available data.

For my Minnesota model, let me begin with a generous wish list of variables of interest. I'll include gross state product, nonfarm employment, nonfarm earned income, farm income, an index of farm prices, farm land values, nonearned income, the unemployment rate, retail sales, the Minneapolis-St. Paul CPI, and possibly some further disaggregations of the above series. I'll tentatively assume that the model will have the simple BVAR form, with n lags of each variable in each equation (n to be determined later).

<u>How Shall I Transform the Data</u>? The BVAR framework does not have strong implications for whether data series should be modeled as seasonally adjusted or not, logged or not, deflated or real, etc. Nonetheless, some conventions appear to have been followed in the majority of BVAR applications. In addition, the BVAR framework itself allows certain new ways of dealing with these issues, especially seasonality.

<u>Seasonal adjustment</u>. Most BVARs to date have been estimated with seasonally adjusted data. I suspect that this is because most BVAR applications so far have been for general purpose macro or regional forecasting, where success consists of beating your competitors' forecasts of the seasonally adjusted

versions of variables like GNP, CPI, and unemployment. To keep things simpler, I will assume that my clients also want me to forecast the seasonally adjusted versions of my Minnesota variables.

As BVARs spread to other applications--such as predicting government revenues or corporate cash flows--forecasts of the nonseasonally adjusted versions of variables will be more relevant. The question then arises: Is it more accurate to directly model the unadjusted series, or should I first adjust, then model, then unadjust? Some research now underway touches on that question, but I don't think there's an answer yet. BVARs do open up one interesting possibility, which is to model the unadjusted data directly and introduce seasonal patterns into the prior beliefs about the coefficients (Canova 1987, Ballabriga 1987).

Logging. The general practice in BVAR models has been to log most variables, except those in ratio form or those taking on negative values. Many economic data series--GNP and Minnesota nonfarm employment, for example--seem to grow nearly exponentially with proportional disturbances. Logging them is convenient because it allows their behavior to be better approximated by a linear model with constant-variance additive disturbances. Even series that show little evidence of exponential growth or proportional disturbances may be logged, if their affects on other variables are believed to be more nearly multiplicative than additive. In U.S. models the exchange rate is often logged partly for these reasons, and in my Minnesota model I will log the index of ag prices for this reason.

Some series are usually not logged, however. Series in ratio or proportional form--such as interest rates or unemployment rates--are usually not logged. Their ratio form already captures some multiplicative effects and their U.S. time series may show little evidence of exponential growth.[3] For these reasons, I will not log the Minnesota unemployment data. Series that have negative values cannot, of course, be logged. For this reason, I will work with levels rather than logs of Minnesota net farm income.

Deflating. With regard to deflating nominal variables, BVAR modelers seem to follow the convention of convention. That is, if the users of the forecast are primarily interested in the deflated version of a variable, then that form is directly incorporated in the model (rather than being computed from the forecasts of the nominal version and a deflator). Thus GNP is generally modeled in real terms (despite the occasional interest in nominal GNP). If the nominal version is of more direct interest--as is often true of stock price indices or the money supply--it is included in the model.

Sometimes this rule of thumb is not workable. For example, I don't know whether the users of my Minnesota forecasts are more interested in real or nominal retail sales. Also, my previous experience with this series has indicated that modeling it in nominal form may lead to implicit forecasts of real growth that seem unreasonable. Therefore, I will deflate retail sales.

Detrending. Although many data series, even those that have been logged, seem to grow over time, BVAR modelers rarely detrend their data series. Trend effects are allowed to enter

into the estimated coefficients of the many distributed lags in the model. Frequently this results in explosive roots in the estimated model.

Are Appropriate Data Series Available? At this point, I know which variables I want in my model, and in which forms. Now I have to see if I the necessary raw data series are available or can be created.

Some series are just not available at all. They must be dropped or replaced by proxies. In my case, gross state product figures for Minnesota are not available now. To some extent, they can be proxied by employment and income.

Some raw series are available, but not frequently enough. The available series on my wish list are reported at a variety of intervals--monthly, quarterly, and annually. I could put them on a common basis with an annual model, but my clients aren't satisfied with this option. An alternative is use interpolation procedures to estimate quarterly values for series, such as farm land values, available only annually. The model could than be standardized on a quarterly calendar. My clients approve, and I proceed on this basis.[4] To interpolate quarterly land values, I need to find a set of quarterly data series that provide information about the likely intrayear movements in land values. I can use some of the series I have already chosen, such as farm prices and incomes. I may also wish to gather some additional series--such as farm mortgage interest rates--just for this purpose.[5]

Some raw series are too short or inconsistent. My Minnesota employment figures go back, in one form or another, to 1939, but income data begins in 1958, retail sales data in 1964, and unemployment data in 1970. I will pick 1958 as my starting date. I throw away the earlier data on employment (except that I've already used it to seasonally adjust) and estimate regression equations for retail sales (using post-1964 data) and unemployment (using post-1970 data) that allow me to backcast them to 1958. In these regressions, as in my interpolations, I may use a combination of variables in the model and variables gathered solely for this purpose. Similar techniques can be used to splice a more recent to a less recent version of a data series that is not available in a statistically consistent form for the whole data period. Mendesh (1987) describes one not entirely successful application of this approach to M1 data in the U.S.

Finally, some series are released with an unusually long delay. If this delay is not too much longer than for the other variables in the model, the data may still be useful. This is the case for the quarterly data on Minnesota earned income, which are released about four months after the end of the quarter they cover. Most of my other series are available by one month after a quarter, but the one quarter lag of the income data does not prevent them from contributing to my forecasts of the other variables. This will probably not be true of the new data series on state gross product that may be released soon. As I understand it, these series will be annual and available about 13 to 15 months after the fact. Even if I could get a good historical

gross state product series and interpolate it to a quarterly level, the delayed release of this series would probably make it useless in my forecasting model.

How Should Relate Series Be Handled? Recall that related series, though not of direct interest themselves, may belong in the model because they affect the forecasts of the variables of interest. The usual BVAR methodology requires that forecasting equations for these related series appear somewhere in the model, so that the model is self-contained in the sense that it can produce forecasts autonomously. However, the modeler has wide latitude in choosing exactly how to model the related series.

The essential feature of a related series is that it contributes to the forecasts of the variables of interest, either by making them more accurate or by allowing modelers to claim that the forecasts incorporate all information their clients believe relevant. Note that to improve the forecasts of the variables of interest, the related series could be either causally related or merely correlated with the variables of interest. In my Minnesota model, for example, national output, income, and employment variables can be viewed as shifting the demand for Minnesota products, thereby causing changes in Minnesota's employment and income. The role of stock market indices, such as the S&P 500, is less clear. They seem to contribute significantly to making both U.S. and Minnesota growth forecasts more accurate, but economists disagree over whether this is because they cause or merely prefigure changes in the real sector. A sharper example was related to me by Thomas Cargill, who found that state gambling revenues in

Nevada could be used to improve forecasts of GNP. Probably no one thinks that Nevada gambling revenues cause big changes in GNP, but it is easy to believe that they might be correlated with--and thus serve as a proxy for--some unmeasured dimension of consumer confidence.

For my Minnesota model, I will choose my related series from a list of important national economic variables. These include the national counterparts of all my Minnesota variables of interest, plus GNP and selected GNP components, the interest rate on 3 month T-bills, the money supply, the exchange rate, the S&P 500 index, and and index of commodity prices.

The usual BVAR estimation method requires that forecasting equations for the related series be included in the model. The method relies on computing realistic simulations of the out-of-sample forecasting performance of the model. That, for example, means estimating the model through the first quarter of 1964, forecasting the next eight quarters, computing the resulting eight forecast errors, and then repeating the process for each successive quarter. If the modeler tries to avoid modeling the related series by simply plugging in someone else's forecasts of these variables, it will probably be impossible to find out what that other forecaster forecasted, or would have forecasted, as of the first quarter of 1964. As a result, it will be impossible to compute realistic statistics on the model's out-of-sample forecasting performance. Statistics computed by using the actual rather than forecasted values of the related series--in effect, under the assumption of perfect foresight of the related series--

will generally be too optimistic. This can lead to the estimation of a model which is suboptimal in computing current forecasts, when perfect foresight of the related series is impossible.

Some important BVAR applications also require that the related series be included in the model. For example, BVAR forecasters frequently use stochastic simulation to compute confidence bands around forecasts or probabilities of events like a recession. Such probability assessments cannot be accurately computed without a stochastic model for the related series.

Although the related series need to be included in the model, they do not have to be modeled in the BVAR style. The modeler is free to make liberal use of exclusion restrictions in the equations for the related variables, since by definition the clients don't care whether the forecasts of the related series incorporate feedback from other variables. At one extreme, then, the modeler may choose to model each related series as a univariate autoregression. At the other extreme, the modeler may simply enlarge the BVAR and treat the related series as though they were variables of interest. Each equation would contain n lags of the variables of interest and of the related variables. Even committed BVAR modelers may stop short of this extreme, however, especially if it generates more coefficients than can be estimated with the modeler's computer package. An intermediate possibility is to follow a block recursive pattern: one BVAR with equations for the related variables only, with no variables of interest on the right-hand side, and a second BVAR with equations for the variables of interest, with some or all of the related variables

on the right-hand side. A variety of other forms is also possible, and the out-of-sample forecasting statistics used in estimating the model can guide the selection of a final model. I will use these statistics in deciding whether to estimate a single BVAR for the Minnesota and U.S. series or to adopt a block recursive structure by excluding Minnesota variables from the U.S. variable's equations.

Is the Model Manageable? I have now drawn up a fairly long wish list of variables of interest and related series. It's time to check whether my staff, my computer, and I can handle so big a model.

My computing resources are probably adequate. They allow about 200 coefficients per equation. Since I am working with quarterly data, I can probably get by with 4 lags of each variable. (I recommend at least enough lags to cover half a year, and preferably at least a year's worth.) This means I can put about 50 variables in the model and still not have to exclude any variable from any equation. If I want to experiment with 8 lags--which might be a good idea--I will have to cut down to about 25 variables. Even that is not very restrictive.

In this case my human resources are binding, however. As usual, I have not allowed enough time to prepare for this conference. In addition, I am not clever enough to clearly and concisely present a large model to you as an illustration of the BVAR technique. As a result, I will prune my model way back. My only variable of interest will be Minnesota nonfarm employment. My related series will be the S&P 500 index, the interest rate on

3 month T-bills, and GNP. All will be modeled as quarterly averages, with data since 1958. Only GNP will be deflated. GNP and Minnesota employment will be seasonally adjusted, but the other two series are already virtually nonseasonal in their raw form. All except interest rates will be logged.

Section 2: Estimating a BVAR

As its name suggests, the Bayesian vector autoregression procedure is motivated by Bayesian statistical theory. However, BVAR methods do not rigorously follow Bayesian guidelines. They can instead be interpreted as approximations to exact Bayesian methods.

Without Bayesian-like restrictions, my proposed Minnesota model would probably be subject to degrees of freedom problems, leading to what is known as overfitting of the coefficients. With four lags of each of my four variables in each of my four equations, plus a constant term, I would have an unrestricted VAR (UVAR) system with 68 coefficients, or 17 per equation. Because the right-hand side variables are the same for each equation, the theory of seemingly unrelated regression (SURE) implies that I can optimally estimate my entire UVAR by applying OLS to each equation separately. I have 119 quarters of data for each equation, or 6 to 7 for each of the 17 coefficients in the equation. Those 17 coefficients are probably more than enough to fit quite well the in-sample data on Minnesota employment. In fact, the fitted coefficients may fit the data too well, in the sense that meaningless accidental patterns in the historical time series may strongly influence the estimated coefficients (see Todd [1984]

for a nontechnical discussion of this idea). This is called overfitting.

A comparison of simulated out-of-sample forecasts of this UVAR and a set of four 4-lag univariate autoregressions suggests that overfitting is at least mild problem in the UVAR (see Table 1). One step ahead, its forecasts for all variables are less accurate than those of the system of univariate equations. For the national variables, this remains true at longer horizons, and the UVAR GNP forecasts are clearly inferior. For Minnesota employment, however, the two models forecasts are nearly equally accurate, and the UVAR does better at long horizons. This reinforces the notion that national variables are important for Minnesota forecasts.

In larger UVAR models, where the number of coefficients grows rapidly, overfitting is more of a problem. The number of coefficients may even exceed the number of observations.

The BVAR method is designed to offset overfitting by imposing Bayesian prior restrictions on the coefficients. In effect, the information that OLS extracts from the data is supplemented by the modeler's beliefs about the likely values of the coefficients. For practical reasons, these prior beliefs have so far been imposed by methods that only approximate true Bayesian procedures.

One reason that BVAR methods are only approximately Bayesian is that the coefficients of all BVARs to date have, I believe, been estimated with equation-by-equation estimators rather than system estimators. Although optimal for UVARs, single

equation methods are suboptimal for BVARs unless the prior variance-covariance matrices of the coefficients are identical for each equation (up to a scale factor given by the ratio of disturbance term variances). It is unlikely that a modeler's prior beliefs would satisfy this condition exactly, so the typical BVAR single equation estimation procedure is inefficient.

Full system estimation of a BVAR, though efficient, is computationally demanding. In my Minnesota model, it would involve the 68 x 68 variance-covariance matrix of all the coefficients, instead of the four 17 x 17 variance-covariance matrices of the separate equations. Some efforts are being made to find efficient algorithms for system estimation of BVARs, but I am not aware of any completed applications yet.

For each BVAR equation, the prior information takes the form of a "known" variance of the disturbance term and a normal distribution for the coefficients.

Though treated as known, the disturbance term variances in the prior are usually computed from the data. [Alternatively, Doan, Litterman, and Sims (1984) propose treating this variance as an additional hyperparameter.] For each BVAR variable, Doan, Litterman, and Sims (1984) propose setting the standard deviation of its disturbance term to 0.9 times the standard error of the residuals in a regression of the variable on six lags of itself. The factor 0.9 allows for a 10 percent reduction in the standard deviation when other variables are added to the equation. All of these standard deviations are also used in scaling the prior covariance matrix of the coefficients of each equation. Therefore

a standard deviation must be computed even for related series that are not modeled with BVAR equations.

The mean of the normal distribution of each equation's coefficients is typically set according to the so-called random walk, or random walk with drift, prior. The intuition behind this prior is that most economic series are reasonably well approximated by a random walk around a trend.

To help capture the trend, a constant is included in each equation. For logged variables, the constant can pick up the trend growth rate. (As it happens, the variables that typically have not been logged also typically exhibit little trend.) Usually, the prior means of all the constant terms are zero. Their prior variances are usually scaled by a hyperparameter dedicated to this purpose. Alternatively, their variances can be set at a large multiple (100,000 for example) of the disturbance term variance, to express prior ignorance about the constant term and thereby let the data determine its value.

To represent a discrete time random walk, the prior mean of the coefficient on the first lag of the dependent variable is set to 1.0. All other coefficients are given a prior mean of 0.0.

It may be more plausible that a variable behaves like a random walk in continuous, not discrete, time. In this case the implications of the random walk prior for the BVAR coefficients are somewhat different. BVARs are generally fit to time-averaged discrete data series. If the underlying continuous time version of the variable follows a random walk, the time-averaged version will show a more complicated pattern of lag coefficients. [See

Christiano and Eichenbaum (1987), pp. 75-77 and Working (1960).] The coefficient of the j^{th} lag of the dependent variable will have a nonzero mean given by $(1-\alpha)\alpha^{j-1}$ where $\alpha = \sqrt{3}-2$. The prior means of the other coefficients in the equation remain zero, however.

Specifying the prior variance-covariance matrices for the coefficients of each equation is the heart of the BVAR method and the core of a typical BVAR computer program. Usually the prior VCVs are chosen from a standard family of VCVs. This family, which is indexed by about ten so-called hyperparameters, has gradually emerged from the work of Sims, Litterman, and Doan. Each setting of the hyperparameters identifies a particular VCV in this family, and thus completes the specification of the prior normal distribution of the coefficients. Viewed another way, the prior VCV can be written as a function of the hyperparameters (and of the disturbance term variances). Doan, Litterman, and Sims (DLS) have thus considerably simplified the task of specifying prior VCVs, reducing the number of decisions from hundreds or thousands to about ten.

The DLS family of VCVs is based on two simple beliefs about the coefficients of a multivariate time series forecasting model. One is the belief that, in forecasting a given variable, its own past values are likely to be more useful than the past values of other variables. The other is that, for both the forecasted variable and the other variables used to forecast it, more recent values are likely to be more useful than more distantly lagged values. In other words, our confidence that a coefficient is nearly zero increases if the coefficient applies to a more

distant lag or to a variable other than the dependent variable, as indicated in Figure 1.

These two simple beliefs are expressed by means of two or three hyperparameters. One, which I will call DAMP, controls the rate at which the prior variances shrink toward zero as lag length increases. This is generally done by making the prior variances of coefficients on j^{th} lags proportional to either j^{-DAMP} or $DAMP^{(j-1)}$. Although the most prominent BVAR software package (RATS, by VAR Econometrics) has a built in facility for varying DAMP (which it refers to as the lag decay parameter), in many applications the formula j^{-1} is used.

Two other basic hyperparameters govern the relative prior variances of coefficients of own lags (lags of the dependent variable) versus coefficients of cross lags (lags of other variables). Own lag variances are all scaled by the hyperparameter DLS call π_1, and cross lag variances are all scaled by the hyperparameter they call π_2. π_1 and π_2 can strongly affect the forecasting performance of a BVAR.

Over the years, the DLS family of VCVs has been refined by the addition of more hyperparameters. The prior variances of the constant terms, for example, are scaled by a hyperparameter they call π_3. Other hyperparameters influence the time variation of the coefficients (π_7,π_8), the sum of the coefficients (π_6), and relative variances among the cross variables (π_4).

Although DLS have taken us a long way toward specifying a prior VCV for our coefficients, it is usually the case that even experienced BVAR modelers cannot pick a setting of the hyperparam-

eters that adequately represents their beliefs. In fact, most of the easily agreed to characteristics of a prior VCV are common to all members of the DLS family, and DLS focus attention on the hyperparameters partly because they index the remaining uncertainty about the nature of the prior. In Bayesian terms, the prior distribution of the coefficients is a mixture of normal distributions, and the hyperparameters index the family of normal distributions in the mixture. (More correctly, the distribution of the coefficients is a mixture of conditional normal distributions, each conditional on the assumed variance of the equation disturbance term.) A Bayesian could proceed by expressing a fairly flat prior distribution over the hyperparameters, computing the implied mixture of coefficient distributions, and using that mixture to specify the prior VCV of the coefficients.

This fully Bayesian approach has never, as far as I know, been implemented. Instead DLS have suggested a procedure for using the historical data to select a particular set of hyperparameters (and thus also a particular VCV for each equation). This procedure appears to be contradictory. A prior is the distribution the modeler believes before examining the data, so how can the data be used to pick the prior? Strictly speaking, they cannot. However, under certain conditions, this data-based method for selecting a prior is justified as an approximation to the Bayesian mixture-of-distributions procedure outlined above (see the Appendix).

The key ingredient in the DLS procedure for selecting a set of hyperparameters is realistic simulation of the model's out-of-sample forecasting performance. A set of hyperparameters is chosen. Then the model is estimated, using Kalman filter techniques to incorporate the prior associated with those hyperparameters, through a startup date. A forecast is made from that date and compared to the actual data to compute its forecast errors. The Kalman filter is applied again to update the estimates of the coefficients with data through date startup plus one, and the forecasts and forecast errors for that date are computed. This continues until the end of the historical data are reached. The whole procedure is then repeated for many other settings of the hyperparameters, so that each hyperparameter setting has a simulated forecasting track record. The setting with the best track record wins. That is, its prior is singled out and used to estimate the final forecasting model.

A critical aspect of this procedure, obviously, is deciding which forecasting track record is best. Several criteria, or metrics, have been proposed and used, including minimization of the log determinant of the VCV of the simulated one-step-ahead forecast errors and minimization of weighted averages of root mean squared forecast errors one or more steps ahead. The choice of a criterion is to some extent problem specific and could be based on the modeler's forecast error loss function. In my Minnesota model, for example, I could focus on my variable of interest, using as my criterion the root mean squared error of one-step-ahead forecasts of Minnesota nonfarm employment.

However, the criterion that is most consistent with the notion that BVAR methods are approximately Bayesian is maximization of the quasilikelihood statistic proposed by Doan, Litterman, and Sims (1984). For any given equation, maximizing this statistic is equivalent (under certain assumptions) to minimizing a weighted sum of the out-of-sample one-step-ahead forecast errors, with weights that are inversely proportional to the conditional forecast error variances. An equation's forecasting performance thus depends both on its model's accuracy and on its model's ability to detect when forecasts are subject to abnormal uncertainty. An analogous statistic could be computed for the model as a whole if efficient system Kalman filtering algorithms were available. Since these algorithms are not available yet, a model's forecasting performance (that is, its performance under a given set of hyperparameters) is ranked by summing the quasilikelihoods of its individual equations. This is the criterion I will use to select hyperparameters for my Minnesota model.

The remaining task--searching for a set of hyperparameters that optimize the criterion--is computationally intensive. Even if the final model selected will be small enough to estimate and use on a microcomputer, it is helpful to use a mainframe at this stage. This is especially true if the number of hyperparameters is large, say 5 or more.

The computational burden of hyperparameter search depends on two factors, the number of hyperparameters to be picked and the computational complexity of computing out-of-sample forecasts for a given set of hyperparameters. The latter factor

depends primarily on the number of coefficients per equation and the length of the historical period over which out-of-sample forecasts will be computed. The former factor determines the dimension of the space to be searched. This is important because it can rule out a nearly surefire method of optimizing the criterion--checking all hyperparameter settings on a fine grid that encompasses all reasonable values. For example, five hyperparameters and 5 grid points for each implies a grid with 5^5, or 3125, settings of the hyperparameters. Except for small models with short out-of-sample forecasting periods, checking all of those hyperparameter settings would require hours--possibly days--of mainframe CPU time. Similarly, computational burdens also limit the possibilities of computing the numerical derivatives needed for hill-climbing algorithms.

With grid search and hill climbing effectively ruled out, two alternatives have generally been used. The most common procedure is sometimes referred to a axial search. The modeler takes an initial guess at the optimal setting of the hyperparameters. Then a grid of settings of the first hyperparameter is searched, holding the remaining hyperparameters fixed. The best point in this grid is used to pick the optimal value of the first hyperparameter. The procedure is repeated for each hyperparameter in succession, always with the other hyperparameters held constant. A grid of 5 points for each of 5 hyperparameters thus implies 25 points under axial search, rather than 3125 under a multidimensional grid search. And if the criterion is a fairly smooth symmetrical function of the hyperparameters, axial search

will get close to the optimal setting of the hyperparameters. (See Figure 1.)

In general, however, axial search can go astray, as illustrated in Figure 2. As the figure suggests, it may be valuable to do several axial searches, varying the order in which the hyperparameters are searched. Table 2a shows the differences between 2 axial searches over 10 hyperparameters for the 4-lags version of my Minnesota model. The first search, leading to Model 1, ends with fairly typical optimal hyperparameter values (though the hyperparameter on the variance of the constant term is somewhat low). The search leading to Model 3, over a different ordering of the hyperparameters, leads to a lower likelihood and somewhat unusual degrees of time variation, decay of coefficients, and tightness on sums of coefficients.

Sims (1986) has recently proposed an alternative search procedure. First the criterion is evaluated on a small grid (up to 50 points) of the hyperparameter space. Then Bayesian procedures and some assumptions about the rate of change of the criterion function are used to interpolate the value of the function for any other hyperparameter setting. These interpolations are not computationally expensive, so a hill climbing routine can be used to find optimal hyperparameter settings for the interpolated function.

Tables 2b-2d show that all three BVAR models achieve better performance statistics for Minnesota employment than the univariate and unrestricted VAR systems, at least one year ahead. Results for the national variables are less clear. One

step ahead, the BVAR models general forecasted more accurately than the univariate and unrestricted VAR models. At longer horizons, however, BVAR Model 1 is dominated by the univariate model and, for TBILL, by the unrestricted VAR. Model 2, which had the highest likelihood, fared better. It almost uniformly forecasted better than the univariate and unrestricted BVAR models. (I also estimated 8-lag univariate and unrestricted VAR models. The former performed about the same as its 4-lag counterpart. The latter performed much worse than any model shown here.) Based on its higher likelihood and overall superior forecasting performance, Model 2 is my choice as the best forecasting model.

The optimal hyperparameter settings from the axial search reported in Table 2a were used to estimate the Minnesota BVAR models 1 and 2. I am now ready to use them, especially model 2, for forecasting my variable of interest. Table 3 shows the results, under two assumptions. With data only through the third quarter of 1987, both models forecast strong GNP and Minnesota employment growth ahead. When preliminary estimates of fourth quarter average stock prices and interest rates are added to the data (to reflect October's stock market crash), the models forecast an immediate but brief downturn in GNP and a slowdown in Minnesota employment growth.

My model is up and running, and my discussion of implementing Bayesian vector autoregressions would now have to turn to applications. Perhaps another day.

Footnotes

[1]This paper does not focus on other topics related to BVARs, such as their conceptual motivation [see Doan, Litterman, and Sims (1984) or Todd (1984)], their forecasting performance [see Litterman (1986a) and McNees (1986)], or their use in resolving questions concerning economic structure or policy [see Sims (1982,1986,1987), Litterman (1985), Cooley and LeRoy (1985), Runkle (1987), and the references within these articles].

[2]This assumes that the unknown underlying structure of the economy will not change through the forecasting period (or, in the case of a time-varying coefficients model, will continue to evolve as it has in the past). This may not be a good assumption in the case of the 1987 stock market crash, which may very well induce new private sector strategies and public sector regulations. In addition, the mere size of the crash itself may be inconsistent with the stochastic structure of many models.

[3]However, in countries where interest rates have taken on more extreme values than in the U.S., the log transformation may be more useful and has the advantage of capturing the multiplicative relationship between interest rates and asset values.

[4]There is a long literature on optimal interpolation of economic time series. One early BVAR modeler, Robert Litterman, made extensive use of interpolated NIPA data to build a monthly model of the U.S. economy (Litterman 1984). His interpolation method (Litterman 1983) generalizes the earlier methods of Chow and Lin (1971) and Fernandez (1981).

[5]Amirizadeh (1985) describes many examples in which series both inside and outside the model are used to interpolate data from a quarterly to a monthly level.

Appendix:

A Bayesian Interpretation of Searching Over Hyperparameters

The following is patterned after the discussion in Doan, Litterman, and Sims (1984). I am responsible for any errors.

Suppose the data y are conditionally distributed with density function $p(y|\theta)$, where θ is a vector of coefficients. If we can then specify a prior density $q(\theta)$ for these coefficients, we can form the joint density $p(y|\theta)q(\theta)$ of the data and coefficients, calculate the likelihood function by plugging observed values of y into p, and apply Bayes's Rule to get a posterior p.d.f. for θ. This is the standard Bayesian approach.

Suppose now, however, that our prior for θ can be written as a function of a smaller set of parameters π, so that $q_1(\theta|\pi)$ represents our prior p.d.f. for θ, conditional on the vector π. Suppose further that we are unsure of our beliefs about π, holding the prior p.d.f. $q_2(\pi)$. Then our prior p.d.f. for θ is actual a mixture of p.d.f.s of the form $q_1(\theta|\pi)$, where the mixture is formed by weighing $q_1(\theta|\pi)$ by $q_2(\pi)$ and integrating, or

$$q(\theta) = \int q_1(\theta|\pi)q_2(\pi)d\pi.$$

We are free to formulate our prior for θ in this way and then proceed as above.

Alternatively, suppose we leave q_2 unspecified and regard y and θ as conditionally (on π) jointly distributed as $p(y|\theta)q_1(\theta|\pi)$. If we integrate θ out, we obtain $m(y|\pi)$, the marginal distribution for y conditional on π. For a fixed data

vector y, $m(y|\pi)$ plays the role of a likelihood function for π. And, if we know our prior distribution for π is relatively flat in the region where $m(y|\pi)$ is large, we can say, even without specifying anything more about our prior for π, that our posterior for π is roughly proportional to the "likelihood" $m(y|\pi)$. Our posterior for θ, in turn, is a weighted average of those obtained (by the usual Bayesian methods) conditional on values of π, with weights given by the posterior on π and hence also roughly proportional to $m(y|\pi)$.

This process of weighing the conditional posterior for θ by $m(y|\pi)$ has never, to my knowledge, been carried out by a BVAR modeler. Instead, the value π^* which maximizes $m(y|\pi)$ has been found (at least approximately), and the posterior value θ^* for θ conditional on π^* has been computed and used as the fitted coefficient vector. On the surface, this is a procedure for using the data to select a prior and is thus not Bayesian procedure. However, if the likelihood for π, $m(y|\pi)$, is large only within some region P (and falls off rapidly elsewhere), and if the posterior for θ conditional on π is not very sensitive to variations of π within P, then the difference between θ^* and true posterior value of θ formed by weighing conditional posteriors for θ by $m(y|\pi)$) is small. In this sense, picking π^* and computing θ^* can approximate the true Bayesian flat-π-prior estimate of θ. At least in some well-known examples of BVAR models [Doan, Litterman, and Sims (1984)] the conditions for approximate optimality of θ^* appear to be met. Estimating BVARs by picking optimal hyperparameters in this way to some extent implies a belief that these conditions are not too seriously violated.

References

Amirizadeh, Hossain. 1985. "Technical Appendix for the Bayesian Vector Autoregressive Model of the U.S. Economy." Research Department, Federal Reserve Bank of Minneapolis.

Ballabriga, Fernando. 1987. "The Relative Importance of International Shocks for the Spanish Economy." University of Minnesota, manuscript.

Canova, Fabio. 1987. "Forecasting Seasonal Series with Bayesian Methods." University of Minnesota, manuscript (May).

Chow, Gregory C., and An-Loh, Lin. 1971. "Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series." Review of Economics and Statistics 53, 372-75.

Christiano, Lawrence J., and Eichenbaum, Martin. 1987. "Temporal Aggregation and Structural Inference in Macroeconomics." Carnegie-Rochester Conference Series on Public Policy 26, 63-130.

Cooley, Thomas F., and LeRoy, Stephen F. 1985. "Atheoretical Macroeconomics: A Critique." Journal of Monetary Economics 16, 283-308.

Doan, Thomas A., and Litterman, Robert B. 1986. User's Manual, RATS Version 2.00. VAR Econometrics.

______________, and Sims, Christopher. 1984. "Forecasting and Conditional Projection Using Realistic Prior Distributions." Econometric Reviews 3, 1-100.

Fernandez, Roque B. 1981. "A Methodological Note on the Estimation of Time Series." Review of Economics and Statistics 63, 471-76.

Litterman, Robert B. 1983. "A Random Walk, Markov Model for the Distribution of Time Series," Journal of Business and Economic Statistics 1, 169-73.

__________. 1984. "Forecasting and Policy Analysis with Bayesian Vector Autoregression Models." Federal Reserve Bank of Minneapolis Quarterly Review, Fall, 30-41.

__________. 1985. "How Monetary Policy in 1985 Affects the Outlook." Federal Reserve Bank of Minneapolis Quarterly Review, Fall, 2-13.

__________. 1986a. "Forecasting with Bayesian Vector Autoregressions--Five Years of Experience." Journal of Business and Economic Statistics 4, 25-38.

__________. 1986b. "A Statistical Approach to Economic Forecasting." Journal of Business and Economic Statistics 4, 1-4.

__________. 1986c. "Specifying Vector Autoregressions for Macroeconomic Forecasting." Chapter 6 in P. Goel and A. Zellner, eds. (1986), Bayesian Inference and Decision Techniques. Elsevier Science Publishers.

McNees, Stephen K. 1986. "Forecasting Accuracy of Alternative Techniques: A Comparisons of U.S. Macroeconomic Forecasts." Journal of Business and Economic Statistics 4, 5-15.

Mendesh, Susan. 1987. "A Warning About M1." Federal Reserve Bank of Minneapolis, memo (Aug. 25)

Runkle, David E. 1987. "Vector Autoregressions and Reality." Journal of Business and Economic Statistics 5, 437-42.

Sims, Christopher A. 1982. "Policy Analysis with Economic Models." Brookings Papers on Economic Activity, 107-64.

______. 1986a. "Are Forecasting Models Usable for Policy Analysis?" Federal Reserve Bank of Minneapolis Quarterly Review, Winter.

______. 1986b. "BAYESMITH: A Program for Multivariate Bayesian Interpolation." University of Minnesota, Center for Economic Research, Discussion Paper 234 (September).

______. 1987. "A Rational Expectations Framework for Short-Run Policy Analysis." William A. Barnett and Kenneth J. Singleton, eds. (1987), New Approaches to Monetary Economics. Cambridge University Press.

Sims, Christopher A., Stock, James, and Watson Mark. 1986. "Inference in Linear Autoregressive Models with Unit Roots." University of Minnesota, manuscript.

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

Working, Holbrook. 1960. "Note on the Correlation of First Difference of Averages in a Random Chain." Econometrica 28, 916-18.

TABLE 1a: FORECASTING PERFORMANCE OF SYSTEM OF 4-LAG UNIVARIATE EQUATIONS

FORECAST STATISTICS FOR SERIES STOCKS 1

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.44129199E-02	0.42502138E-01	0.56529307E-01	0.91949525	94
2	0.90060410E-02	0.76055179E-01	0.97561945E-01	0.95152243	93
3	0.13454169E-01	0.10344890	0.12923917	0.96814327	92
4	0.16141438E-01	0.12081456	0.15058548	0.96612738	91
5	0.18712443E-01	0.13713138	0.17071071	0.96600106	90
6	0.21700713E-01	0.14884193	0.18911620	0.96774744	89
7	0.24127009E-01	0.15938954	0.20380418	0.97004361	88
8	0.24719777E-01	0.16682311	0.21228678	0.97025187	87
9	0.24937701E-01	0.17093978	0.21570405	0.95989219	86
10	0.25664570E-01	0.17983987	0.22225952	0.95251589	85
11	0.27211717E-01	0.18355103	0.22822183	0.94390414	84
12	0.28204854E-01	0.18342790	0.23081953	0.92956019	83

FORECAST STATISTICS FOR SERIES TBILL 2

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.10744089E-01	0.63635907	0.99507047	0.97783979	94
2	-0.28423123E-01	1.1173874	1.5945973	1.0338853	93
3	-0.46950444E-01	1.3271041	1.8077188	1.0209214	92
4	-0.70139866E-01	1.5833610	2.0769228	1.0178494	91
5	-0.94756936E-01	1.8466537	2.3556417	1.0166868	90
6	-0.11252579	2.1243604	2.6516255	1.0175021	89
7	-0.11957813	2.4048976	2.9585040	1.0331561	88
8	-0.12470743	2.5288659	3.0916833	1.0372513	87
9	-0.13603401	2.5632563	3.1929793	1.0334242	86
10	-0.14381942	2.7354623	3.3796909	1.0367232	85
11	-0.14543923	2.8960808	3.5406308	1.0424161	84
12	-0.14503393	2.9584100	3.6232216	1.0444460	83

FORECAST STATISTICS FOR SERIES GNP 3

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.17570410E-03	0.77680328E-02	0.98777363E-02	0.79684875	94
2	0.39077426E-03	0.12310705E-01	0.15663106E-01	0.72366887	93
3	0.61226834E-03	0.17298338E-01	0.21433504E-01	0.69999839	92
4	0.80596419E-03	0.20828869E-01	0.26260504E-01	0.67236302	91
5	0.88962953E-03	0.25027546E-01	0.30763167E-01	0.65346167	90
6	0.91657657E-03	0.28125063E-01	0.34389082E-01	0.62959772	89
7	0.95143430E-03	0.31414946E-01	0.37761297E-01	0.60953130	88
8	0.71383881E-03	0.34583572E-01	0.40548729E-01	0.58997202	87
9	0.32904849E-03	0.36886651E-01	0.42487490E-01	0.56694144	86
10	-0.71850752E-04	0.38161105E-01	0.43948990E-01	0.54224725	85
11	-0.47867238E-03	0.39811195E-01	0.45218275E-01	0.51909701	84
12	-0.92070822E-03	0.40445501E-01	0.46587041E-01	0.49970574	83

FORECAST STATISTICS FOR SERIES EMPYMNS 4

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.10791717E-02	0.53675996E-02	0.64582705E-02	0.62954228	94
2	0.24459644E-02	0.10848400E-01	0.13116183E-01	0.66134598	93
3	0.38908519E-02	0.16458232E-01	0.19579521E-01	0.67165396	92
4	0.52632735E-02	0.21999314E-01	0.25701500E-01	0.67459082	91
5	0.65432795E-02	0.27043109E-01	0.31326160E-01	0.67199097	90
6	0.77887125E-02	0.31471280E-01	0.36392367E-01	0.66511334	89
7	0.89933605E-02	0.35636953E-01	0.40921063E-01	0.65536163	88
8	0.10109570E-01	0.38966145E-01	0.44899233E-01	0.64370795	87
9	0.11226806E-01	0.41347767E-01	0.48208264E-01	0.62860079	86
10	0.12254826E-01	0.43229375E-01	0.50970997E-01	0.61195080	85
11	0.13223554E-01	0.44512279E-01	0.53262566E-01	0.59425100	84
12	0.14222564E-01	0.45364438E-01	0.55285882E-01	0.57681050	83

CURRENT VALUE OF THE METRIC IS
-592.0565842
NOTE THE LOOP HAS ENDED AND THE PROGRAM SHOULD NOW END NORMALLY
HE LOOP HAS ENDED AND THE PROGRAM SHOULD NOW END NORMALLY
END

TABLE 1b: FORECASTING PERFORMANCE OF [redacted] 4-LAG UNRESTRICTED VAR

FORECAST STATISTICS FOR SERIES STOCKS 1

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.64694356E-02	0.56751922E-01	0.69585327E-01	1.1318420	94
2	0.17736743E-01	0.10303019	0.12422263	1.2115443	93
3	0.26725381E-01	0.13366990	0.16293498	1.2205619	92
4	0.31752133E-01	0.15047637	0.18640531	1.1959405	91
5	0.38491230E-01	0.16913933	0.21208120	1.2001043	90
6	0.48161679E-01	0.18509641	0.23810691	1.2184432	89
7	0.57521604E-01	0.19628924	0.25839812	1.2298935	88
8	0.64125931E-01	0.20885106	0.27720111	1.2669414	87
9	0.68350946E-01	0.22064346	0.29740094	1.3234468	86
10	0.69878834E-01	0.23570159	0.31640937	1.3560047	85
11	0.69121985E-01	0.25191640	0.34551420	1.4290144	84
12	0.65056825E-01	0.27150705	0.38216381	1.5390564	83

FORECAST STATISTICS FOR SERIES TBILL 2

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.70387902E-01	0.72140918	1.0716148	1.0530586	94
2	-0.13341794	1.2586961	1.7903711	1.1608187	93
3	-0.14002256	1.4172188	1.9552289	1.1042287	92
4	-0.14276755	1.5966581	2.2257944	1.0908077	91
5	-0.16651768	1.8768557	2.5755833	1.1116128	90
6	-0.17980173	2.1436248	2.8387486	1.0893064	89
7	-0.15489196	2.2513894	3.0805749	1.0757852	88
8	-0.85740564E-01	2.3709643	3.1562491	1.0589130	87
9	-0.17771405E-01	2.4230443	3.1791546	1.0289498	86
10	0.74239393E-02	2.4642894	3.2862272	1.0080531	85
11	0.18654165E-01	2.5397968	3.3671545	0.99134204	84
12	0.53176012E-01	2.5657213	3.3852398	0.97586300	83

FORECAST STATISTICS FOR SERIES GNP 3

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.81345505E-03	0.92385880E-02	0.12096472E-01	0.97586126	94
2	0.15835237E-02	0.14849571E-01	0.19575944E-01	0.90445033	93
3	0.26956973E-02	0.20600086E-01	0.27233923E-01	0.88943473	92
4	0.41698089E-02	0.27198716E-01	0.35305520E-01	0.90394785	91
5	0.57661137E-02	0.33522560E-01	0.43583759E-01	0.92579270	90
6	0.72477327E-02	0.39464848E-01	0.52031886E-01	0.95260343	89
7	0.85765682E-02	0.44988504E-01	0.61003090E-01	0.98469321	88
8	0.10053677E-01	0.51197739E-01	0.71922229E-01	1.0464472	87
9	0.11360671E-01	0.59164359E-01	0.84437222E-01	1.1267072	86
10	0.12129087E-01	0.66671978E-01	0.99412117E-01	1.2265571	85
11	0.11529283E-01	0.74478869E-01	0.11703721	1.3435645	84
12	0.99343588E-02	0.82881626E-01	0.13764736	1.4764448	83

FORECAST STATISTICS FOR SERIES EMPYHNS 4

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.14234481E-03	0.52966308E-02	0.66608800E-02	0.64929233	94
2	0.41964464E-03	0.94588151E-02	0.11549874E-01	0.58236933	93
3	0.61662928E-03	0.13755369E-01	0.16687664E-01	0.57245198	92
4	0.13490653E-02	0.18894005E-01	0.23195708E-01	0.60882095	91
5	0.22083735E-02	0.24733789E-01	0.30678605E-01	0.65809998	90
6	0.28418103E-02	0.30665487E-01	0.38385592E-01	0.70154186	89
7	0.32666174E-02	0.36326160E-01	0.46396025E-01	0.74304459	88
8	0.34885120E-02	0.41518127E-01	0.55188549E-01	0.79122304	87
9	0.37528284E-02	0.47678737E-01	0.65189030E-01	0.85001766	86
10	0.38819301E-02	0.54361450E-01	0.77858764E-01	0.93477387	85
11	0.30685420E-02	0.61980576E-01	0.93047568E-01	1.0381327	84
12	0.12844247E-02	0.70609216E-01	0.11052038	1.1530849	83

CURRENT VALUE OF THE METRIC IS
655.9263919
NOTE THE LOOP HAS ENDED AND THE PROGRAM SHOULD NOW END NORMALLY
HE LOOP HAS ENDED AND THE PROGRAM SHOULD NOW END NORMALLY
END

Table 2a: Optimal Hyperparameters and Likelihood for 3 specifications and searches

(D&S notation)	OWN (π_1)	CROSS (π_2)	TVAR (π_7)	SUM (π_6)	TITE (π_5)	WT (π_4)	CONST (π_3)	DECAY (π_8)	BEGWT*	MNWT**	Likelihood
ORDERINGS											
First order of search	1	2	9	8	6	10	5	7	3	4	
Second order of search	9	8	3	1	4	2	7	6	10	5	
RESULTS											
First order, 4 lags (MODEL 1)	.01	.01	.00001	10.0	5.0	1.0	.001	1.0	0.2	0.25	762.38
First order, 8 lags (MODEL 2)	.01	.01	.0001	10,000.0	5.0	.01	.001	.999	1.5	10.0	789.55
Second order 4 lags (MODEL 3)	.05	.01	.0001	10,000.0	5.0	.01	100.0	.999	0.2	10.0	760.73

* Extends no-change forecast in prior to coefficient covariances

** Larger values increasingly constrain Minnesota coefficients in U.S. equations to be near zero. A value of 10 implies very little feedback from Minnesota to U.S.

Table 2b

Forecasting Performance of BVAR Model 1

FORECAST STATISTICS FOR SERIES STOCKS 1

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.96457942E-02	0.41427079E-01	0.55084460E-01	0.89402629	94
2	0.22915027E-01	0.75341238E-01	0.94585965E-01	0.92249767	93
3	0.37016104E-01	0.10077463	0.12506047	0.94281814	92
4	0.49051095E-01	0.11948865	0.14578003	0.93529655	91
5	0.60568702E-01	0.13870431	0.16640435	0.94163263	90
6	0.71452074E-01	0.15374952	0.18655456	0.95463890	89
7	0.80870501E-01	0.16654578	0.20355499	0.96085754	88
8	0.88026369E-01	0.17331010	0.21617859	0.98803932	87
9	0.94479909E-01	0.17740964	0.22330196	0.99370323	86
10	0.10122276	0.18107277	0.23258063	0.99674804	85
11	0.10854837	0.18970026	0.24344236	1.0060548	84
12	0.11542238	0.19562545	0.25367343	1.0211961	83

FORECAST STATISTICS FOR SERIES TBILL 2

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.66237272E-01	0.64805951	1.0024518	0.98509333	94
2	-0.13208265	1.1374443	1.6338264	1.0593197	93
3	-0.15719414	1.3765318	1.8100463	1.0226909	92
4	-0.18117597	1.5773181	2.0916440	1.0250439	91
5	-0.20117751	1.8054276	2.4156204	1.0425760	90
6	-0.19145483	2.0679167	2.7231837	1.0449609	89
7	-0.16944901	2.3409463	3.1175659	1.0887030	88
8	-0.15550373	2.5476163	3.3218537	1.1144729	87
9	-0.14695365	2.6407432	3.4365845	1.1122689	86
10	-0.13016236	2.8240013	3.6683321	1.1252642	85
11	-0.10348479	2.9999127	3.8970086	1.1473392	84
12	-0.76784764E-01	3.1334149	4.0116873	1.1544490	83

FORECAST STATISTICS FOR SERIES GNP 3

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.64374095E-03	0.79779010E-02	0.98313782E-02	0.79312890	94
2	-0.12042653E-02	0.12271489E-01	0.15022276E-01	0.69406116	93
3	-0.12648121E-02	0.17098211E-01	0.20820094E-01	0.67996502	92
4	-0.97842509E-03	0.22186438E-01	0.26025275E-01	0.66634032	91
5	-0.42922274E-03	0.26643925E-01	0.31549552E-01	0.67016582	90
6	0.36377118E-03	0.30551046E-01	0.36660280E-01	0.67117898	89
7	0.15903903E-02	0.35103609E-01	0.41922874E-01	0.67670621	88
8	0.22722351E-02	0.39164592E-01	0.47713133E-01	0.69421198	87
9	0.33118754E-02	0.44512570E-01	0.53332177E-01	0.71164998	86
10	0.45271268E-02	0.49767265E-01	0.59664058E-01	0.73614140	85
11	0.56902127E-02	0.54906141E-01	0.66182659E-01	0.75976406	84
12	0.67937170E-02	0.60463561E-01	0.73629265E-01	0.78976829	83

FORECAST STATISTICS FOR SERIES EMPYHRS 4

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.49030206E-03	0.38388829E-02	0.47857412E-02	0.46650669	94
2	-0.11091418E-02	0.74209033E-02	0.91212427E-02	0.45991254	93
3	-0.14448603E-02	0.10900423E-01	0.13522790E-01	0.46388446	92
4	-0.17157648E-02	0.15542160E-01	0.18689575E-01	0.49054787	91
5	-0.18296199E-02	0.20765361E-01	0.24236411E-01	0.51990571	90
6	-0.15728967E-02	0.25839602E-01	0.29645826E-01	0.54181237	89
7	-0.10286120E-02	0.30648695E-01	0.35020676E-01	0.56086537	88
8	-0.42042340E-03	0.35660001E-01	0.40454940E-01	0.57999134	87
9	0.53645708E-03	0.39522225E-01	0.45409275E-01	0.59210400	86
10	0.16083528E-02	0.43050071E-01	0.50102495E-01	0.60153155	85
11	0.26489130E-02	0.47376026E-01	0.55095144E-01	0.61469710	84
12	0.37500124E-02	0.51762176E-01	0.60092712E-01	0.62696127	83

Table 2c

Forecasting Performance of BVAR Model 2

FORECAST STATISTICS FOR SERIES STOCKS 1

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.22856762E-02	0.42061257E-01	0.55767352E-01	0.90710142	94
2	0.31050540E-02	0.75705907E-01	0.95092600E-01	0.92743967	93
3	0.11514946E-02	0.99360867E-01	0.12482402	0.93506897	92
4	-0.45228540E-02	0.11010722	0.14087597	0.90383306	91
5	-0.11277940E-01	0.12299407	0.15686149	0.88763245	90
6	-0.18665904E-01	0.13415047	0.17220939	0.88123173	89
7	-0.26924196E-01	0.14601013	0.18572286	0.88390222	88
8	-0.36636097E-01	0.15387294	0.19664602	0.89876613	87
9	-0.46030556E-01	0.16023759	0.20279405	0.90244214	86
10	-0.56731366E-01	0.16780003	0.21035922	0.90151592	85
11	-0.66116721E-01	0.17109914	0.21814299	0.90221901	84
12	-0.75970975E-01	0.17690725	0.22369054	0.90085028	83

FORECAST STATISTICS FOR SERIES TBILL 2

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.15492653E-02	0.63847701	0.96049839	0.94386636	94
2	0.17478035E-01	1.1022750	1.5295039	0.99168087	93
3	0.67086400E-01	1.2430841	1.6536204	0.93389326	92
4	0.12194619	1.4434553	1.9036699	0.93294236	91
5	0.18567637	1.6671164	2.1836558	0.94245829	90
6	0.26447798	1.9470373	2.4374892	0.95533204	89
7	0.34042350	2.1750300	2.7405140	0.95703056	88
8	0.40882801	2.2568759	2.9055395	0.97480058	87
9	0.47863532	2.3317609	2.9788709	0.96412692	86
10	0.55300760	2.4726919	3.1269580	0.95919717	85
11	0.63677045	2.6160178	3.2846774	0.96705952	84
12	0.72548312	2.7045816	3.3794420	0.97419167	83

FORECAST STATISTICS FOR SERIES GNP 3

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.11946314E-03	0.75846571E-02	0.96080025E-02	0.77510846	94
2	0.22709227E-03	0.11573013E-01	0.14943110E-01	0.69040351	93
3	0.44770294E-03	0.16283800E-01	0.20675239E-01	0.67523418	92
4	0.50809497E-03	0.20209497E-01	0.25525459E-01	0.65354325	91
5	0.31461501E-03	0.24194079E-01	0.29988990E-01	0.63701684	90
6	0.20589413E-04	0.27835083E-01	0.33678184E-01	0.61475176	89
7	-0.40121485E-03	0.30892641E-01	0.36770968E-01	0.59354570	88
8	-0.12175093E-02	0.33500818E-01	0.39440695E-01	0.57385045	87
9	-0.22155679E-02	0.35352302E-01	0.41112388E-01	0.54859245	86
10	-0.31990707E-02	0.37023205E-01	0.42722974E-01	0.52712053	85
11	-0.42341280E-02	0.38577541E-01	0.44125157E-01	0.50654823	84
12	-0.53656837E-02	0.39503988E-01	0.45586761E-01	0.48897648	83

FORECAST STATISTICS FOR SERIES EMPYMNS 4

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.17109189E-02	0.43507330E-02	0.53161874E-02	0.51821378	94
2	0.39960907E-02	0.87830836E-02	0.10600580E-01	0.53450025	93
3	0.60735472E-02	0.13618909E-01	0.16379234E-01	0.56187163	92
4	0.10024067E-01	0.19389729E-01	0.22983272E-01	0.60324510	91
5	0.13231756E-01	0.25262979E-01	0.29746641E-01	0.63810805	90
6	0.16595874E-01	0.31109084E-01	0.36304747E-01	0.66351198	89
7	0.19977011E-01	0.34944198E-01	0.42601706E-01	0.68227757	88
8	0.23136702E-01	0.41041318E-01	0.48085716E-01	0.69223035	87
9	0.26300392E-01	0.46268755E-01	0.53197008E-01	0.69565035	86
10	0.29426213E-01	0.49954740E-01	0.57470080E-01	0.68990692	85
11	0.32534462E-01	0.52764902E-01	0.61303317E-01	0.68396174	84
12	0.35664753E-01	0.55455294E-01	0.64764457E-01	0.67570260	83

Table 2d

Forecasting Performance of BVAR Model 3

FORECAST STATISTICS FOR SERIES STOCKS 1

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.35541065E-02	0.43619641E-01	0.55569037E-01	0.90387560	94
2	0.70145339E-02	0.78614110E-01	0.95721289E-01	0.93357049	93
3	0.92834583E-02	0.10287171	0.12703585	0.95163204	92
4	0.95860105E-02	0.11990585	0.14613505	0.93757428	91
5	0.10630259E-01	0.13571973	0.16582809	0.93837175	90
6	0.11948344E-01	0.14836693	0.18378172	0.94044977	89
7	0.12352147E-01	0.15846075	0.19714965	0.93837011	88
8	0.11074612E-01	0.16263327	0.20623418	0.94258861	87
9	0.92638316E-02	0.16770949	0.21042021	0.93637889	86
10	0.75334383E-02	0.16800670	0.21498613	0.92134504	85
11	0.59935856E-02	0.17423073	0.22132618	0.91538437	84
12	0.34810444E-02	0.17841622	0.22663072	0.91269092	83

FORECAST STATISTICS FOR SERIES TBILL 2

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.41453112E-01	0.64017480	0.99358674	0.97638176	94
2	-0.81996515E-01	1.1394534	1.5930433	1.0320778	93
3	-0.96030376E-01	1.3679289	1.7882523	1.0099275	92
4	-0.11252442	1.5891316	2.0644306	1.0117273	91
5	-0.11541352	1.8382827	2.3381579	1.0091408	90
6	-0.96483139E-01	2.0928594	2.5789440	0.98961215	89
7	-0.74925399E-01	2.2902063	2.8535669	0.99651044	88
8	-0.55075583E-01	2.3617689	2.9809486	1.0001001	87
9	-0.29396955E-01	2.3676690	3.0320875	0.98135074	86
10	0.70387229E-02	2.4539561	3.1106811	0.95420421	85
11	0.50948545E-01	2.5182793	3.1910632	0.93949804	84
12	0.93502966E-01	2.5722453	3.2281564	0.93058057	83

FORECAST STATISTICS FOR SERIES GNP 3

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	-0.60704319E-03	0.77246191E-02	0.95961864E-02	0.77415521	94
2	-0.13451009E-02	0.11333701E-01	0.14273644E-01	0.65947277	93
3	-0.19275345E-02	0.15273835E-01	0.19227774E-01	0.62796131	92
4	-0.25165850E-02	0.18371969E-01	0.23179269E-01	0.59347237	91
5	-0.31186671E-02	0.21601460E-01	0.27108948E-01	0.57583987	90
6	-0.36654863E-02	0.24334274E-01	0.30372252E-01	0.55605732	89
7	-0.42829020E-02	0.27108428E-01	0.33580945E-01	0.54205333	88
8	-0.52181970E-02	0.29454842E-01	0.37036705E-01	0.53887311	87
9	-0.62393715E-02	0.31473511E-01	0.39974194E-01	0.53348471	86
10	-0.73465038E-02	0.34458918E-01	0.43510503E-01	0.53683714	85
11	-0.87215344E-02	0.37332458E-01	0.47108468E-01	0.54079605	84
12	-0.10336060E-01	0.39937586E-01	0.50844328E-01	0.54537063	83

FORECAST STATISTICS FOR SERIES EMPYMS 4

STEP	MEAN ERROR	MEAN ABS. ERROR	RMS ERROR	THEIL U	N. OBS
1	0.54594528E-03	0.38844228E-02	0.48762372E-02	0.47532811	94
2	0.12777084E-02	0.76160731E-02	0.92663490E-02	0.46722911	93
3	0.22617341E-02	0.11106328E-01	0.13648983E-01	0.46821335	92
4	0.32156990E-02	0.15303952E-01	0.18582637E-01	0.48774104	91
5	0.41151689E-02	0.19488001E-01	0.23597387E-01	0.50619773	90
6	0.51190481E-02	0.23722275E-01	0.28420107E-01	0.51941090	89
7	0.61616246E-02	0.27888767E-01	0.33200984E-01	0.53172252	88
8	0.70654911E-02	0.32061078E-01	0.37841654E-01	0.54252509	87
9	0.80697252E-02	0.35857706E-01	0.42040031E-01	0.54017151	86
10	0.89407464E-02	0.38512914E-01	0.45975726E-01	0.55198548	85
11	0.96242604E-02	0.41936904E-01	0.49987859E-01	0.55771506	84
12	0.10233763E-01	0.44356588E-01	0.53757773E-01	0.56086738	83

Table 3

Forecasts of Growth Rates for GNP and Minnesota Employement

	GNP				MINNESOTA NONFARM EMPLOYMENT			
	With Data Through 87:3 only		With Additional Data on Stock Prices and T-Bill Rates of 87:4		With Data Through 87:3 only		With Additional Data on Stock Prices and T-bill Rates for 87:4	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
87:4	4.7	4.5	-1.1	12.0	3.6	3.5	4.5	4.0
88:1	4.4	4.2	2.4	0.9	3.6	3.5	2.1	1.7
88:2	3.6	2.9	2.5	0.4	3.7	3.1	1.4	0.7
88:3	3.0	1.7	3.7	1.1	3.5	2.2	1.7	-0.1
88:4	2.7	1.1	5.2	1.9	3.1	1.5	2.8	0.2
89:1	2.5	1.3	5.5	3.6	2.8	1.2	3.7	1.5
89:2	2.4	1.5	5.0	4.5	2.6	0.9	4.0	2.5
89:3	2.3	1.8	4.3	4.4	2.4	0.9	4.1	3.1
89:4	2.3	2.2	3.8	3.9	2.3	1.0	4.0	3.1

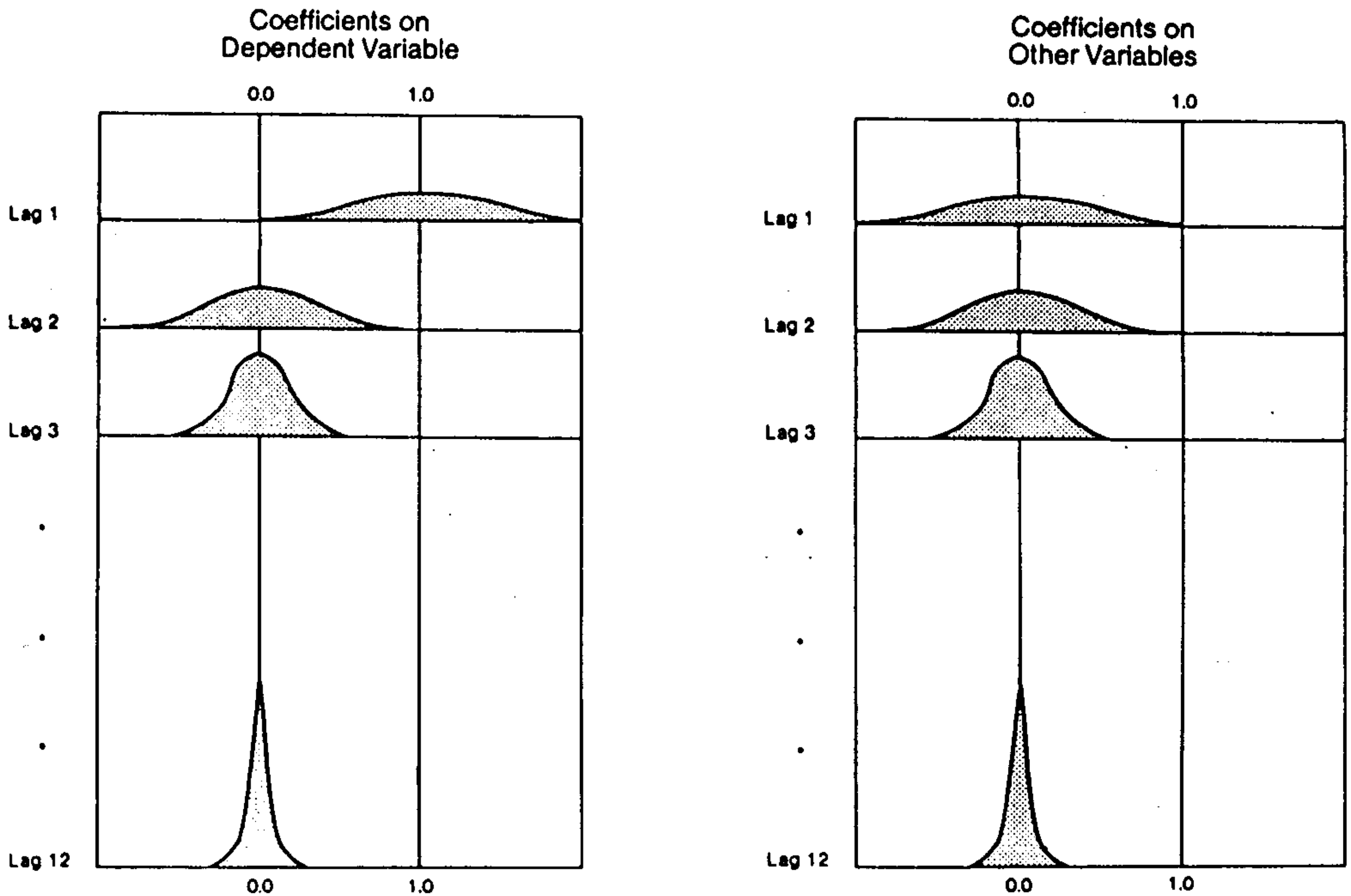

Fig. 1. A schematic representation of the prior.

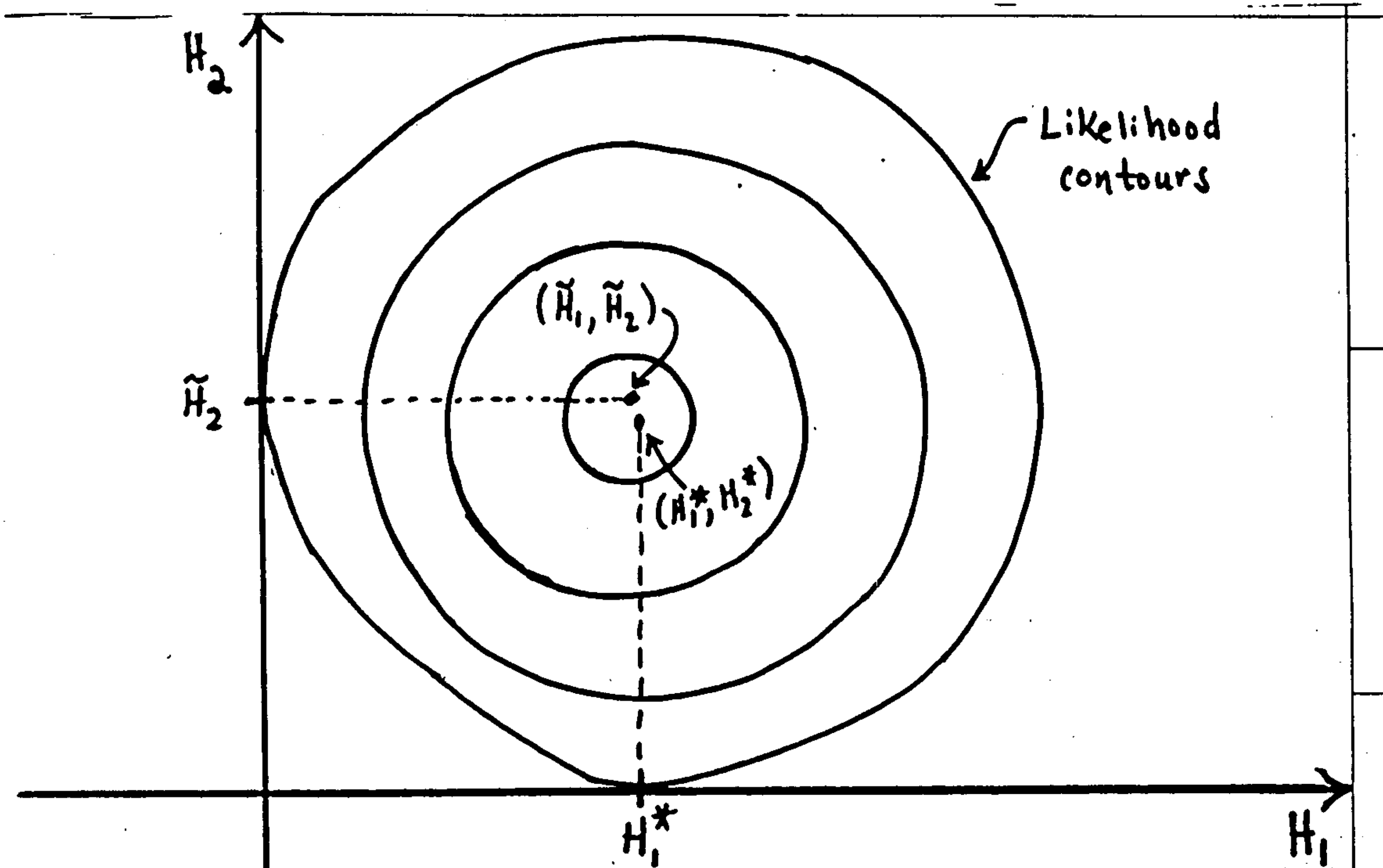

Fig. 2a: Successful Axial Search

* indicates results of searching H_1, then H_2

~ indicates results of searching H_2, then H_1

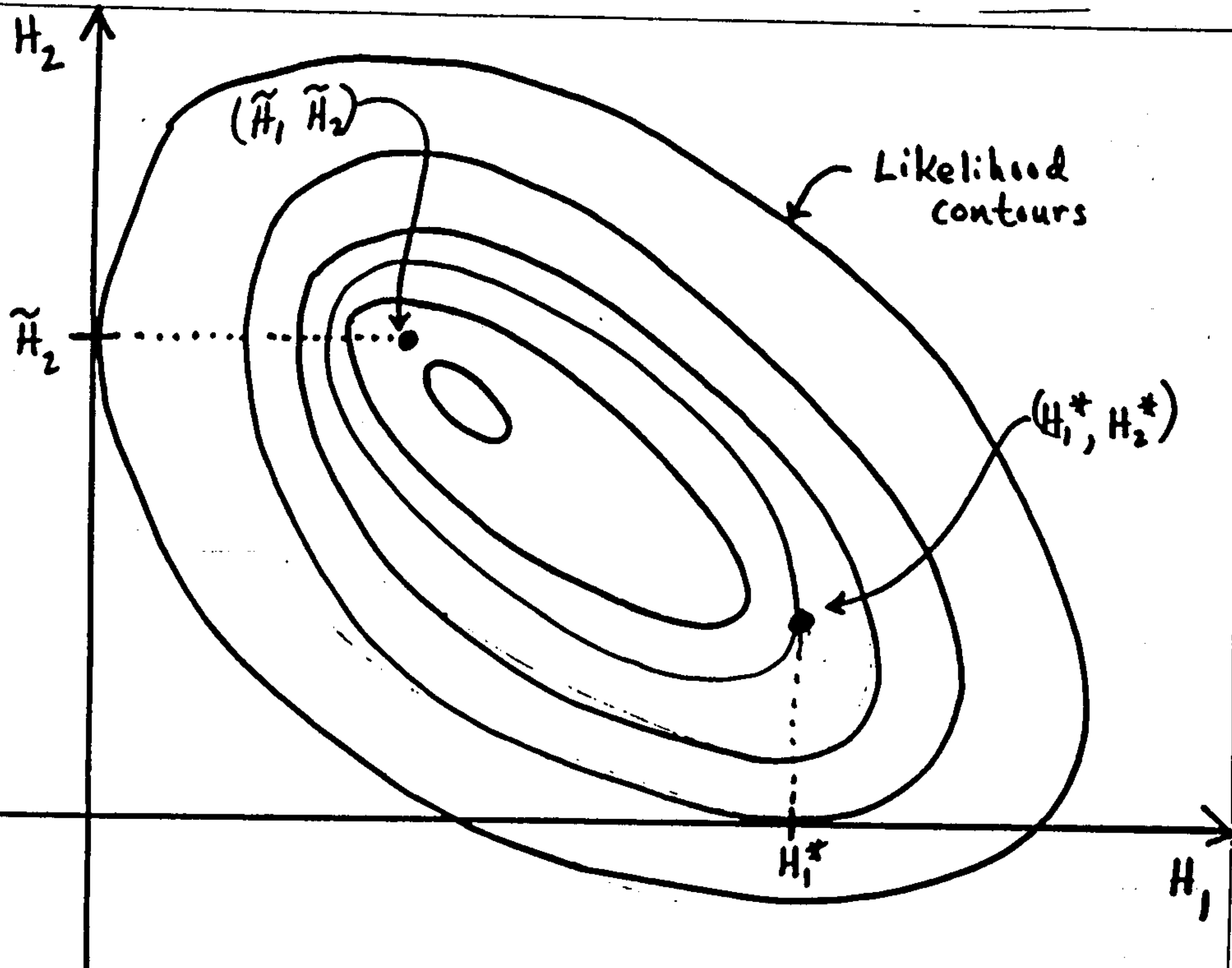

Fig 2b: Potentially Problematic Axial Search
* indicates results of searching H_1, then H_2
~ indicates results of searching H_2, then H_1

Software Appendix

RATS Programming Language (VAR Econometrics)

```
*** PROGRAM RGPROG(PITTSBRG)                                    ***
*** ESTIMATES HYPERPARAMETERS FOR MODEL WITH MN NONFARM EMPLOYMENT,
*** REAL GNP, S&P 500 INDEX, AND INTEREST RATE ON 3 MONTH T-BILLS.
***
*** FILE RSTS.IMRMT.BNDECST SHOULD BE ON TAPE 1 !!!!!
***
BMA(NOPRINT) COMPILE 4000
BMA(NOPRINT) GLOBAL 500
CHECK(NOTRACE) UNDEFINED
CAL 55 1 4
IEVAL FOREC=0
IEVAL ULAGS = 6
IEVAL NVARS = 4
IEVAL NENDOG = 4
IEVAL NEXOG = NVARS-NENDOG
IEVAL CALBEG = (55,1)
IEVAL DBEG = (58,1)
IEVAL ENDDATA = (87,3)
IEVAL CALEND = ENDDATA+18
IEVAL BEGTHEIL = DBEG+24
IEVAL FH = 12;* FH IS FORECAST HORIZON IN THEIL STATS
*
* THIS BLOCK OF STUFF USED IN COMPUTING A FORECASTING METRIC
EVAL METRIC = 0.0
EVAL OMETRIC = 100000.0
DEC SYMM VCV4(4,4)
*
DEC RECT FF
DEC INDEX HYPORD(10)
INPUT HYPORD
2 1 9 10 7 5 8 4 3 6
DEC INDEX HYPDIM(10)
INPUT HYPDIM
5 6 5 6 5 5 8 6 5 10
DEC VEC VCROSS(10)
INPUT VCROSS
0.01 .00001 .0001 .001 0.1 0.5 1.0 0.0 0.0 0.0
DEC VEC VOWN(10)
INPUT VOWN
0.1 .0001 .001 .01 0.05 .1 1.0 10.0 0.0 0.0
DEC VEC VTVAR(10)
INPUT VTVAR
.000001 .00000001 .0000001 .00001 .0001 .00001
.0001 .001 .01 .1
DEC VEC VSUM(10)
INPUT VSUM
10. 1.0 100.0 500.0 1000.0 10000.0 10000.0 100000.0 0.0 0.0
DEC VEC VTITE(10)
INPUT VTITE
5.0 0.01 0.1 10.0 100.0 5.0 10.0 100.0 0.0 0.0
DEC VEC VWT(10)
INPUT VWT
1.0 0.001 0.01 1.0 3.0 6.0 10.0 0.0 0.0 0.0
DEC VEC VCON(10)
INPUT VCON
10000 0.001 0.01 1.0 10.0 100.0 1000.0 100000.0 0.0 0.0
DEC VEC VDECAY(10)
INPUT VDECAY
1.0 .9 .95 .99 .999 .9999 0.0 0.0 0.0 0.0
DEC VEC VBEGWT(10)
INPUT VBEGWT
```

```
0.2 .3 .5 .75 1.5 2.0 0.0 0.0 0.0 0.0
DEC VEC VMMWT(10)
INPUT VMMWT
1.0 2.0 5.0 10.0 100.0 1000.0 100000.0 0.0 .25 .5
***                                                                    ***
IEVAL NLAGS = 6
IEVAL NCNTEMP = 0;*NUMBER OF CONTEMPORANEOUS EXOG TERMS
IEVAL NCOEFF = NVARS*NLAGS + NCNTEMP + 1
DEC RECT LIKEHD(2,NENDOG)
MAT LIKEHD = CONST(0.0)
DECLARE VECTOR SUMCON(NCOEFF) TMPCON(NCOEFF) FAC(1)
DECLARE RECT XXOVER(NCOEFF*(NCOEFF+1)/2,NENDOG)
DECLARE RECT TVOVER(NCOEFF*(NCOEFF+1)/2,NENDOG)
DEC SYM XX1 XX2 XX3 XX4
DEC SYM TV1 TV2 TV3 TV4
OVERLAY XXOVER(1,1) WITH XX1(NCOEFF,NCOEFF)
OVERLAY XXOVER(1,2) WITH XX2(NCOEFF,NCOEFF)
OVERLAY XXOVER(1,3) WITH XX3(NCOEFF,NCOEFF)
OVERLAY XXOVER(1,4) WITH XX4(NCOEFF,NCOEFF)
OVERLAY TVOVER(1,1) WITH TV1(NCOEFF,NCOEFF)
OVERLAY TVOVER(1,2) WITH TV2(NCOEFF,NCOEFF)
OVERLAY TVOVER(1,3) WITH TV3(NCOEFF,NCOEFF)
OVERLAY TVOVER(1,4) WITH TV4(NCOEFF,NCOEFF)
ALLOCATE 2*NVARS+2 CALEND 2*NENDOG NCOEFF+2
CLEAR
EQV 1 TO NVARS+1+NEXOG+1
STOCKS TBILL GNP EMPYMNS KONSTANT PGNP
DATA(FORMAT=RATS) DBEG ENDDATA STOCKS TO EMPYMNS
DATA(FORMAT=RATS) DBEG ENDDATA PGNP
SET KONSTANT CALBEG CALEND = 1.0
*
DECLARE INDEX LOGDATA(NVARS)
INPUT LOGDATA
1 0 1 1
DO I = 1, NVARS
 IF LOGDATA(I).EQ.1
 BEGIN
   LOG I DBEG ENDDATA I DBEG
 END
END DO I
*PRINT(DATES) CALBEG CALEND 1 TO NVARS+1+NEXOG+1
*
DEC VECT SAMVAR(NVARS) DIAGVAR(NCOEFF) OVERDIAG
DEC VECT CVECT1 CVECT2 OVERTMP
DEC SYMM XXSYMM TVSYMM
DECLARE VECT MEANS(NENDOG)
DECLARE RECT WEIGHT(NENDOG,NVARS)
*** NOTE: DIMENSIONS OF COEFFSAV CONFORM TO THOSE OF COEFF, WHICH
***       RATS SPECIFIES AS (4) BY (3), WHERE (4) AND (3) ARE THE
***       FOURTH AND THIRD PARAMETERS ON THE ALLOCATE CARD
DEC RECT COEFSAV(NCOEFF+2,2*NENDOG)
MAT COEFSAV = CONST(0.0)
INPUT MEANS
1.0 1.0 1.0 1.0 1.0 1.0 1.0
INPUT WEIGHT
.0 2.0 2.0 6.0
1.0 0.0 1.0 3.0
1.0 1.0 0.0 3.0
1.0 1.0 1.0 0.0
*
DO I = 1, NVARS
```

```
  OLS(NOPRINT) I DBEG+ULAGS ENDDATA
  #  -I 1 ULAGS CONSTANT
*  MULTIPLICATION OF SEESQ BY .9 ALLOWS FOR IMPROVEMENT WITH MORE VARIABLES.
  EVAL SAMVAR(I) = SEESQ*.9
END DO I
*
DO J = 1,NENDOG
  EQUATION(NOCONST,MORE) J J
#-STOCKS 1 NLAGS -TBILL 1 NLAGS -GNP 1 NLAGS -EMPYHNS 1 NLAGS KONSTANT
END DO J
*
EVAL ALPHA = SQRT(3.0) - 2.0
*
*BEGIN LOOP OVER HYPERPARAMETER GRID
*
CHA COPY 8
DO HYPL = 1,10
IEVAL HYP = HYPORD(HYPL)
  EVAL CROSS = VCROSS(1)
  EVAL OWN = VOWN(1)
  EVAL TVAR = VTVAR(1)
  EVAL SUM = VSUM(1)
  EVAL TITE = VTITE(1)
  EVAL WT = VWT(1)
  EVAL CON = VCON(1)
  EVAL DECAY = VDECAY(1)
  EVAL BEGWT = VBEGWT(1)
  EVAL MNWT = VMNWT(1)
  DO LOOP = 1,HYPDIM(HYP)
  *DO LOOP = 1,1
    IF HYP.EQ.1; EVAL CROSS = VCROSS(LOOP)
    IF HYP.EQ.2; EVAL OWN = VOWN(LOOP)
    IF HYP.EQ.3; EVAL TVAR = VTVAR(LOOP)
    IF HYP.EQ.4; EVAL SUM = VSUM(LOOP)
    IF HYP.EQ.5; EVAL TITE = VTITE(LOOP)
    IF HYP.EQ.6; EVAL WT = VWT(LOOP)
    IF HYP.EQ.7; EVAL CON = VCON(LOOP)
    IF HYP.EQ.8; EVAL DECAY = VDECAY(LOOP)
    IF HYP.EQ.9; EVAL BEGWT = VBEGWT(LOOP)
    IF HYP.EQ.10; EVAL MNWT = VMNWT(LOOP)
***
NOTE ZZZZ CURRENT SETTINGS OF HYPERPARAMETERS AND LOOP INDICES
NOTE CROSS
EVAL CROSS
NOTE OWN
EVAL OWN
NOTE TVAR
EVAL TVAR
NOTE SUM
EVAL SUM
NOTE TITE
EVAL TITE
NOTE WT
EVAL WT
NOTE CON
EVAL CON
NOTE DECAY
EVAL DECAY
NOTE BEGWT
EVAL BEGWT
NOTE MNWT
```

```
EVAL MWHT
NOTE GRID SEARCH INDICES
EVAL HYP
EVAL LOOP
*** COPY 58-60 DATA INTO 55-57 FOR USE IN INITIALIZING KALMAN FILTER
*** THIS MUST BE INSIDE LOOP TO RESET EACH TIME
DO II = 1,NVARS+1
 SET(SCRATCH) II CALBEG+1 CALBEG+11 = X(T+12,II)
END DO II
***
EVAL WEIGHT(1,4) = 6.0
EVAL WEIGHT(2,4) = 3.0
EVAL WEIGHT(3,4) = 3.0
DO I = 1,3
  EVAL WEIGHT(I,4) = WEIGHT(I,4)*MWHT
END DO I
***
*A POINT ON THE HYPERPARAMETER GRID HAS NOW BEEN SELECTED.
*THE FOLLOWING CODE ESTIMATES THE MODEL WITH THOSE HYPERPARAMETERS,
*GENERATES ITS FORECASTING STATISTICS, AND COMPARES THOSE STATISTICS TO
*THE BEST PREVIOUS STATISTICS.  IF THE NEW STATISTICS ARE BETTER, THE
*FIRST MEMBER OF THE HYPERPARAMETER VECTOR CORRESPONDING TO THE CURRENT
*VALUE OF "HYP" IS CHANGED TO EQUAL THE CURRENT SETTING OF THAT HYPER-
*PARAMETER. THEN LOOP AND/OR HYP ARE UPDATED AND THE LOOP CONTINUES.
***
MAT COEFSAV = CONST(0.0)
SYSTEM 1 TO NENDOG
KFSET(NOSCALE,LIKELY=LIKEHD,CONSTANT,NAME='XX')
# SAMVAR(1) SAMVAR(2) SAMVAR(3) SAMVAR(4)
TVARYING TV1 TV2 TV3 TV4
END(SYSTEM)
*
DO IEQN = 1,NENDOG
  OVERLAY COEFF(1,IEQN) WITH CVECT1(NCOEFF)
  OVERLAY COEFSAV(1,IEQN) WITH CVECT2(NCOEFF)
  MATRIX CVECT1 = CONST(0.0)
  EVAL CVECT1(1+(IEQN-1)*NLAGS)= MEANS(IEQN)
  OVERLAY CVECT1(1+(IEQN-1)*NLAGS) WITH OVERTMP(NLAGS)
  EWISE OVERTMP(I) = (1-ALPHA) * (ALPHA)**(I-1)
  MAT CVECT2 = CVECT1
END DO IEQN
EVAL DIAGVAR(NCOEFF) = CON * OWN
DO ISER = 1, NENDOG
  OVERLAY XXOVER(1,ISER) WITH XXSYMM(NCOEFF,NCOEFF)
  OVERLAY TVOVER(1,ISER) WITH TVSYMM(NCOEFF,NCOEFF)
  DO IVAR = 1, NVARS
    OVERLAY DIAGVAR(1+(IVAR-1)*NLAGS) WITH OVERDIAG(NLAGS)
    EWISE OVERDIAG(I) = CROSS*EXP(-1.*WT*WEIGHT(ISER,IVAR))/ $
                     (I*SAMVAR(IVAR))
  END DO IVAR
  MATRIX XXSYMM = DIAG(DIAGVAR) * SCALE(SAMVAR(ISER))
  IEVAL IPT = (ISER-1)*NLAGS
  DO L = 1, NLAGS
    EVAL XXSYMM(IPT+L,IPT+L) = XXSYMM(IPT+L,IPT+L) * OWN/CROSS
  END DO L
  DO IEQN = 1, NVARS
  MAT SUMCON = CONST(0.0)
  OVERLAY SUMCON(1+(IEQN-1)*NLAGS) WITH OVERDIAG(NLAGS)
  MAT OVERDIAG = CONST(SUM*SQRT(SAMVAR(IEQN)/SAMVAR(ISER)))
  MAT TMPCON = XXSYMM*SUMCON
  MAT FAC = TR(TMPCON)*SUMCON
```

```
   EVAL FACTOR = FAC(1)
   MAT XXSYMM= XXSYMM - (THPCON*TR(THPCON))*SCALE(1./(1.+FACTOR))
   END DO IEQN
   MATRIX XXSYMM=XXSYMM*SCALE(TITE)
   MATRIX TVSYMM=XXSYMM*SCALE(TVAR)
   EVAL TVSYMM(NCOEFF,NCOEFF)=TVSYMM(NCOEFF,NCOEFF)/CON
END DO ISER
MAT COEFSAV = COEFSAV * SCALE(1.-DECAY)
LIST IEQN = 1 TO NENDOG
FORECAST NENDOG 1 CALBEG+1+NLAGS
CARDS IEQN IEQN CALBEG+1+NLAGS
DO IVAR = 1,NVARS+1
  SET IVAR CALBEG+1 CALBEG+1+NLAGS = X(T,IVAR)*BEGWT
END DO IVAR
KALMAN(STARTUP=CALBEG+1+NLAGS)
*
*** RESET LIKEHD TO ZERO SO THAT DUMMY OBSERVATIONS PROCESSED ABOVE IN
*** SETTING THE PRIOR DON'T AFFECT THE LIKELIHOOD STATISITICS
MAT LIKEHD = CONST(0.0)
*
THEIL(SETUP) NENDOG FH ENDDATA
# 1 TO NENDOG
KALMAN(STARTUP=DBEG+NLAGS)
DO DATE = DBEG+NLAGS+1,ENDDATA
  MATRIX COEFF = SCALE(DECAY) * COEFF + COEFSAV
  DO IEQN = 1,NENDOG
    EVAL COEFF(NCOEFF,IEQN) = (COEFF(NCOEFF,IEQN) $
                              -COEFSAV(NCOEFF,IEQN))/DECAY
  END DO IEQN
  MATRIX XXXOVER = SCALE(DECAY**2) * XXXOVER
*** TO PRINT COEFFS AND OTHER STATS AT ENDDATA, USE NEXT LINES
*KALMAN(SINGLE) 7 0 0 (ENDDATA.EQ.DATE)
*** TO SURPRESS PRINTING OF COEFFICIENTS, ETC, USE NEXT LINE
  KALMAN(SINGLE) 7 0 0
 IF DATE.GE.BEGTHEIL.AND.DATE.LT.ENDDATA
   THEIL DATE+1
END DO DATE
IF FOREC.EQ.1
BEGIN
  VCV(MATRIX=VCV4) DBEG+NLAGS ENDDATA
  # 7 8 9 10
END
***
EVAL LIKLYSUM = 0.0
DO I = 1,NENDOG
EVAL LIKLYSUM=LIKLYSUM-(LIKEHD(1,I)+(ENDDATA-DBEG+1-NLAGS) $
                                  *LOG(LIKEHD(2,I)))
END DO I
MAT LIKEHD = CONST(0.0)
EVAL LIKLYSUM
*
WRITE(UNIT=COPY,NOSKIP,FOR='(F9.6,1X,F9.6,1X,F10.9,1X,F12.4,1X,F9.5)') $
          CROSS OWN TVAR SUM TITE
WRITE(UNIT=COPY,NOSKIP,FOR='(1X,F9.6,1X,F11.3,1X,F9.7,1X,F9.5,1X)') $
          NT CON DECAY BEGWT
WRITE(UNIT=COPY,NOSKIP,FOR='(F9.5,1X,F13.6,1X)') MWT LIKLYSUM
*
*** COMPUTE FORECAST STATISTICS AND UPDATE BEST HYPERPARAMETERS
*
THEIL(DUMP,REPLACE)
FETCH FF = THEIL
```

```
EVAL METRIC = -LIKLYSUM
NOTE * CURRENT VALUE OF THE METRIC IS
EVAL METRIC
IF METRIC.LT.OMETRIC
  BEGIN
  EVAL VCROSS(1) = CROSS
  EVAL VOWN(1) = OWN
  EVAL VTVAR(1) = TVAR
  EVAL VSUM(1) = SUM
  EVAL VTITE(1) = TITE
  EVAL VWT(1) = WT
  EVAL VCON(1) = CON
  EVAL VDECAY(1) = DECAY
  EVAL VBEGWT(1) = BEGWT
  EVAL VMNWT(1) = MNWT
  EVAL OMETRIC = METRIC
  END
EVAL METRIC = 0.0
IF FOREC.EQ.1
BEGIN FOREC
SET STOCKS (87,4) (87,4) = LOG(246.27)
SET TBILL (87,4) (87,4) = 5.4
SMODIFY(BLOCK=2) 4 VCV4
# 1 5 1
# 2 6 2
# 3 7 3
# 4 8 4
FORECAST(PRINT) 2 1 (87,4)
# 7 GNP (87,4)
# 8 EMPYMNS (87,4)
FORECAST(PRINT) 4 12 (88,1)
# 1 7 (88,1)
# 2 8 (88,1)
# 3 9 (88,1)
# 4 10 (88,1)
FORECAST(PRINT) 4 13 (87,4)
# 1 1 (87,4)
# 2 2 (87,4)
# 3 3 (87,4)
# 4 4 (87,4)
EXP 1 DBEG CALEND 1 DBEG
EXP 3 DBEG CALEND 3 DBEG
EXP 4 DBEG CALEND 4 DBEG
EXP 7 DBEG CALEND 7 DBEG
EXP 9 DBEG CALEND 9 DBEG
EXP 10 DBEG CALEND 10 DBEG
PRINT(DATES) (86,1) (90,4) 1 TO 4 7 TO 10
DO I = 1,10
  DO J = (87,4),(90,4)
    IF X(J-1,I).NE.(0.0)
    EVAL X(J-ENDDATA,I) = (((X(J,I)/X(J-1,I))**4)-1.0)*100.
  END DO J
END DO I
PRINT(DATES) 1 14 1 TO 4 7 TO 10
END FOREC
*
MAT COEFF = CONST(0.0)
MAT CVECT1 = CONST(0.0)
MAT CVECT2 = CONST(0.0)
MAT OVERTMP = CONST(0.0)
MAT DIAGVAR = CONST(0.0)
```

```
MAT XXOVER = CONST(0.0)
MAT XXSYMM = CONST(0.0)
MAT TVOVER = CONST(0.0)
MAT TVSYMM = CONST(0.0)
MAT OVERDIAG = CONST(0.0)
MAT SUMCON = CONST(0.0)
MAT TMPCON = CONST(0.0)
MAT FAC = CONST(0.0)
MAT COEFSAV = CONST(0.0)
MAT FF = CONST(0.0)
MAT XX1 = CONST(0.0)
MAT XX2 = CONST(0.0)
MAT XX3 = CONST(0.0)
MAT XX4 = CONST(0.0)
MAT TV1 = CONST(0.0)
MAT TV2 = CONST(0.0)
MAT TV3 = CONST(0.0)
MAT TV4 = CONST(0.0)
*
  END DO LOOP
END DO HYPL
NOTE THE LOOP HAS ENDED AND THE PROGRAM SHOULD NOW END NORMALLY
END
```