TOWARD MORE ACCURATE MACROECONOMIC FORECASTS FROM VECTOR AUTOREGRESSIONS

Roy H. Webb

Several recent articles have used vector autoregressive (VAR) models to forecast national and regional economic variables.[1] Although the models were small and in many cases the statistical techniques were relatively simple, the forecasts produced were of comparable accuracy to predictions made by forecasting services using much larger models, more elaborate statistical techniques, and incorporating the judgment of many analysts.[2] This article extends the earlier work by first showing a method for improving VAR forecasts. That method is illustrated in conjunction with the VAR model introduced by Webb [1984]. Briefly, the source of improvement is to determine lag lengths in the model by a clear-cut statistical procedure, rather than by the typical practice of specifying an arbitrary length a priori. The effect is to significantly reduce the number of estimated coefficients relative to an unrestricted VAR.

The article is organized as follows. First, a rationale for using VAR models for forecasting is presented. Next is a discussion of lag length selection. Empirical results from the proposed method of lag length selection are presented in the following section. Finally, those results are compared with two other methods for improving a VAR model's predictive accuracy.

*An earlier version of this paper was presented to the Western Economic Association Conference, Anaheim, California, July 3, 1985. The author is indebted to John Connaughton, Susan Dolman, Michael Dotsey, Thomas M. Humphrey, Anatoli Kuprianov, William Lupoletti, Yash P. Mehra, and Lee Ohanian for helpful comments. Eric Hill provided valuable research assistance. The views and opinions expressed in this paper are those of the author, and should not be attributed to other individuals in the Federal Reserve System.

[1] For example, see Litterman [1984b] for national forecasts, and Kuprianov and Lupoletti [1984] for regional forecasts.

[2] The assertion that VAR models can be competitive with major forecasting services has been made by Litterman [1984a], based on four years of actual forecasts. That assertion has also been made by Lupoletti and Webb [1984], based on fifteen years of simulated forecasts.

Why Use VARs?

Much of the recent interest in atheoretical methods, including VAR models, reflects a growing disenchantment with conventional structural macromodels. In part, that disenchantment is based on conventional models' spurious endogeneity/exogeneity distinctions, their ad hoc treatment of expectations, a perceived lack of correspondence between model equations and the original motivating theory, and their need for continuous ad hoc adjustments in order to produce satisfactory results.[3]

A particularly appealing motivation for using VAR models has been presented by Hakkio and Morris [1984]. They view a VAR as a reduced form that provides a flexible approximation to the reduced form of any model included in a wide variety of structural models. As such, they present empirical evidence that a VAR model can be dramatically superior to a misspecified structural model. Therefore, critics who believe that conventional macromodels are grossly misspecified have room to believe that a simple VAR model might better approximate the reduced form that would be derived from a model that reflected the true structure of the economy.

Statistical Lag Length Selection

VAR models estimate future values of a set of variables from their own past values. For example, consider one of the equations from a VAR model:

$$X_{v,t} = \alpha + \sum_{i=1}^{k} \sum_{j=1}^{m_i} \beta_{i,j} X_{i,t-j} + \epsilon_{v,t} \qquad (1)$$

where X is a vector of k variables, v is an integer between 1 and k, t is an integer that indexes time, α is a constant term, $\beta_{i,j}$ is a coefficient (α and the βs are estimated by ordinary least squares), m_i is the

[3] Sims [1980] presents a particularly forceful and lucid critique of conventional modeling strategies.

lag length for variable i, and $\epsilon_{v,t}$ is an error term. In other words, the current value of each variable is predicted by lagged values of itself and all other variables in X. It is apparent that the value chosen for each lag length is of some importance. Too large a value means that too many coefficients are estimated, resulting in a loss of precision in the estimates. But too small a value means that important lagged terms are omitted, thereby producing biased coefficient estimates. Both imprecision and bias will increase forecast error variance.

Nevertheless, there is no generally accepted procedure for choosing a lag length. Many authors simply present a number with no explanation for how it was chosen. Others use traditional hypothesis tests, such as F-tests or likelihood ratio tests: to compare alternative specifications. Those tests, however, were designed for testing well-defined alternatives derived from a priori theory. Since the choice of lag lengths in a VAR model does not involve theory-based dichotomies, the use of classical hypothesis tests to determine the lag length is questionable.

In addition, the prevailing custom is to treat all variables identically, thereby using one common lag length for each independent variable in each equation. It is possible, however, that identical treatment could lead to the common lag length being too long in some cases while being too short in others.

Traditional methods of choosing the lag lengths in VAR models therefore seem to leave room for improvement. The strategy examined below is to consider each lag separately, and to use a statistical procedure appropriate for exploratory data analysis rather than hypothesis testing. For each equation, lag lengths are chosen to minimize the Akaike information criterion, or AIC, which was originally proposed for selecting the order of a univariate autoregression. The AIC is a function for which the value depends on the number of estimated coefficients, and can be written as

$$\mathrm{AIC}(m_1, \ldots, m_k) = (N-p) \log \sigma^2 + 2p \quad (2)$$

where k again is the number of variables in the VAR model, each m_i is the lag length for the i'th variable in the equation, p is the number of estimated coefficients, N is the number of observations used to estimate the equation, and σ^2 is the maximum likelihood estimate of the residual variance.

The intuition behind equation (2) is straightforward, and reflects the tradeoff that exists with respect to adding a coefficient to a statistical model. To the extent that the additional coefficient improves the in-sample fit, the residual variance declines. By itself, that would generally[4] cause the AIC to decline by lowering the first term in equation (2). However, estimating an additional coefficient with a fixed number of observations tends to reduce the precision of all the coefficient estimates. That is reflected in the second term, which imposes a penalty for each additional coefficient that is estimated. The minimum AIC therefore reflects a balance between the two opposing factors.

Although the minimum AIC can be determined by inspection in small models, a problem arises as the size of a model increases. Consider a five variable VAR, with possible lag lengths from one to twelve for each variable. Each equation would have 12^5, or 248,832, possible combinations of variables. Rather than attempting to examine each possibility, an alternative is to start with short lag lengths and add coefficients as long as the AIC declines. A difficulty arises, however, since with typical macroeconomic data, the AIC does not decline smoothly as the number of coefficients rises. Thus it is easy to reach one of several local minima without finding the specification that yields the global minimum AIC. This paper uses an extensive search procedure that is described in the Appendix. That procedure is somewhat different from one proposed by Fackler [1985], who addressed a related problem: using Akaike's final prediction error to determine the lag lengths in a VAR model designed for indicating causality.

In addition to its intuitive appeal, there is evidence that the AIC has been used successfully in other settings. Most notably, Meese and Geweke [1984] investigated the performance of several methods of choosing lag lengths for univariate autoregressions that were used to predict 150 macroeconomic time series. They found that the AIC produced the best forecasts more often than any other method studied.

Empirical Results

The relative forecasting performance of several models is examined by studying simulated forecasts over a fifteen-year period. The models include an unrestricted, five-variable VAR model (UVAR5), the corresponding model specified by the AIC method (AVAR5), an unrestricted six-variable VAR model (UVAR6), the corresponding model specified by the

[4] In a few cases the term (N-p) log σ^2 can increase due to an increase in p.

BOX

AN ILLUSTRATIVE VAR MODEL

This box employs a simple two-variable VAR model to illustrate the techniques discussed in the text. The variables are the T-bill rate (RTB) and the percentage change in the monetary base (DB). With a common lag length of one, that model would consist of two equations :

$$RTB_t = \alpha_1 + \beta_{11}RTB_{t-1} + \beta_{12}DB_{t-1} \qquad (B1)$$

$$DB_t = \alpha_2 + \beta_{21}RTB_{t-1} + \beta_{22}DB_{t-1} \qquad (B2)$$

where the α'sand ß's are estimated coefficients. Current observations of RTB and DB can be used to forecast future values by inserting the current values into the right sides of equations B1 and B2 and then calculating a one-quarter-ahead forecast for each variable. Those values, in turn, can be inserted into the right side of each equation and a two-quarter-ahead forecast prepared for each equation. The same process can be repeated as many times as desired: in this way, a forecast can be produced for as many steps ahead as desired.

In practice, longer lag lengths are usually necessary for accurate forecasts. A generalization of the model presented above that allows for longer lags can be written

$$RTB_t = \alpha_1 + \sum_{i=1}^{m_{11}} \beta_{11i}RTB_{t-i} + \sum_{i=1}^{m_{12}} \beta_{12i}DB_{t-i} \qquad (B3)$$

$$DB_t = \alpha_2 + \sum_{i=1}^{m_{21}} \beta_{21i}RTB_{t-i} + \sum_{i=1}^{m_{22}} \beta_{22i}DB_{t-i} \qquad (B4)$$

where the m's represent the lag length for each variable in each equation. In order to choose specific values for each m, suppose that lag lengths between one and four are under consideration for equation B3. Since there are two variables on the right side and four possible lag lengths, there are 4^2 possible choices. Based on data from 1952:2 to 1969:4, the sixteen regression equations were estimated, values of the Akaike Information Criterion (AIC) were calculated, and the results are displayed below.

			Lag length for the monetary base, DB			
			1	2	3	4
AIC VALUES	Lag length for the T-bill rate, RTB	1	—119	—118	—113	—108
		2	—125	—122	—118	—113
		3	—127*	—123	—118	—113
		4	—123	—119	—114	—109

The starred value represents the minimum AIC for the lag lengths that were examined. Therefore, in equation B3, m_{11} would equal three and m_{12} would equal one.

Another example is the construction of a "counter" variable (C). Suppose one wished to construct a counter for the T-bill rate over a two-year period for which the values are shown below. One could first construct the "indicator" variable (I) below, which would indicate the direction of change of the T-bill rate by letting 1 represent an increase, -1 represent a decline, and 0 represent no change. The next step would be to let the counter variable in a particular quarter equal the cumulative sum of the indicator variable up to that point.

Variable	Quarter								
	0	1	2	3	4	5	6	7	8
RTB	7	8	9	8	8	7	8	9	10
I		1	1	—1	0	—1	1	1	1
C		1	2	1	1	0	1	2	3

A counter variable thus constructed could be added as an independent variable in either of the equations above.

AIC method (AVAR6) and a set of univariate autoregressive forecasts (AR).[5]

Table I contains forecast results for three key variables: the 90-day Treasury bill rate, the growth rate of real GNP, and the growth rate of the GNP implicit price deflator. Each model's coefficients were estimated using quarterly data from 1952:2 to 1969:4. Forecasts were then constructed for 1970:1 through 1971:4. Each model's coefficients were then reestimated, using data from 1952:2 to 1970:1, and forecasts were constructed from 1970:2 to 1972:1. That procedure was repeated until the model was reestimated and forecasts were prepared for every quarter through 19841:3. Thus for each model a series of out-of-sample forecasts was generated: 60 one-quarter-ahead forecasts, 59 two-quarter-ahead forecasts, and so forth, up to 53 eight-quarter-ahead forecasts. Those forecasts were then compared with actual data. The root-mean-squared-error, or RMSE, is given by each entry in the first three columns of Table I.

In order to more easily compare models, the final two columns of the table include (admittedly crude) summary statistics for each model's performance. The first measure is simply the sum of the RMSEs for each variable at each horizon indicated. Lower values, of course, indicate increasing accuracy. The other measure is the sum of points awarded for the relative performance of each forecast. For each variable at each horizon, three points are awarded for the most accurate forecast, two for the second best, and one for the third best. (Points are split for ties.) In this case, higher point totals represent: better relative forecasts.

It is useful to initially consider the first three models listed. The summary measures indicate that the UVAR5 and AR models are rather evenly matched. Considering only those two models, it appears that the benefits from multivariate interaction in UVAR5 are almost exactly negated by the burden of estimating too many coefficients. It is

[5] The lag lengths for the univariate autoregressions were simply the own-lag lengths from the AVAR5 model.

Table I

FORECAST ERRORS, 1970 TO 1984

Model	Horizon	Interest Rote	Real GNP	Implicit Deflator	Sum 1	Sum 2
UVAR5	1	1.20	4.93	2.04		
	4	2.76	2.56	1.91		
	8	4.19	2.41	2.46	24.26	6
AVAR5	1	1.20	4.61	1.89		
	4	2.47	3.01	1.96		
	8	3.57	2.25	2.29	23.25	15
AR	1	1.24	4.62	1.76		
	4	2.62	3.17	1.96		
	8	4.13	2.29	2.54	24.33	8
UVAR6	1	1.22	5.37	2.07		
	4	2.99	2.92	1.89		
	8	4.31	2.52	2.34	25.63	5.5
AVAR6	1	1.20	4.76	1.89		
	4	2.44	2.96	1.89		
	8	3.57	1.92	2.20	22.83	19.5

Note: Column 1 contains the names of the models used to generate forecasts. Column 2 is the forecast horizon-the length of the forecast, measured in quarters. Columns 3 through 5 contain the RMSE's of post-sample forecasts. The interest rote is the 90-day Treasury bill rate. Real GNP and the implicit deflator are percentage changes at annual rates. For those two variables, 4 and 8 quarter changes are the average change over the particular period. Column 6 contains a summary measure of model performance, namely the sum of the RMSE's for each variable at each horizon. Column 7 contains another summary measure, a point total that assigns three points for the most accurate forecast for each variable at each horizon, two points for the second most accurate forecast, and one point for the third most accurate forecast.

also interesting to note the AR model's best relative performance was at the one-quarter interval. That result is intuitively plausible, since one might expect the benefits of multivariate interaction to be greatest at longer intervals.[6]

Comparing the UVAR5 and AVAR5 models illustrates the improved accuracy attainable from estimating fewer coefficients. As can be seen in Table II, which contains the specification of the models, the AVAR5 model[7] contains only 75 coefficients. In contrast, the UVAR5 model contains 155 coefficients -31 per equation. The substantial reduction in the number of coefficients suggests a possible avenue for further improvement: the smaller number of coefficients in AVAR5 might leave enough room for another variable to be included.

As Table I indicates, forecasting accuracy at various horizons for the three variables of interest was improved by adding a sixth variable. Finding that sixth variable required a fair amount of search, however.[8] It quickly became evident that an additional variable would fail to improve the accuracy of forecasts if any one of three conditions held : (1) the additional variable did not appreciably augment the explanatory power of in-sample regressions for which the dependent variable was one of the three key variables; (2) when the additional variable was the dependent variable, the in-sample fit of its regression equation was poor; or (3) the additional variable itself could not be predicted accurately in post-sample forecasts.

The third condition is worth emphasizing, since it may not be obvious to the casual user of VAR forecasts. For example, in searching for a sixth variable for AVAR6, preliminary regressions (with GNP, the deflator, or the interest rate as the dependent variable) fit better within the sample period when either a stock price index or the foreign exchange value of the dollar was included as an explanatory variable. Attempts to predict those two variables were unsuccessful, however, with forecasts at all horizons having a Theil U-statistic substantially greater than one. (That statistic indicates that simply using the last observation of the stock index or exchange rate as the forecast would have been more accurate than the model's prediction.) Not sur-

Table II

LAG LENGTHS FOR EACH EQUATION IN EACH MODEL

Model	Dependent Variable	Explanatory Variables: RTB	DX	DP	CU	DB	GX
UVAR5	RTB	6	6	6	6	6	
	DX	6	6	6	6	6	
	DP	6	6	6	6	6	
	CU	6	6	6	6	6	
	DB	6	6	6	6	6	
AVAR5	RTB	4	0	4	0	1	
	DX	1	15	0	7	4	
	DP	0	1	4	5	2	
	CU	2	0	0	5	1	
	DB	3	4	3	0	4	
AR	RTB	4	0	0			
	DX	0	15	0			
	DP	0	0	4			
UVAR6	RTB	6	6	6	6	6	6
	DX	6	6	6	6	6	6
	DP	6	6	6	6	6	6
	CU	6	6	6	6	6	6
	DB	6	6	6	6	6	6
	GX	6	6	6	6	6	6
AVAR6	RTB	4	0	4	0	1	0
	DX	1	9	0	1	4	7
	DP	0	1	4	5	2	0
	CU	2	0	0	5	4	5
	DB	2	1	1	0	4	5
	GX	2	0	0	0	0	4

Note: The entries in this table represent the lag lengths for each variable in each equation in each model. An entry of zero indicates that the variable is excluded from that equation. A positive number indicates the number of lags that were included, beginning with the first lagged value. A constant term was also included in each equation. The variables are as follows: RTB, 90-day Treasury bill rate; DX, real GNP, first difference in logs times 400; DP, GNP implicit price deflator, first difference in logs times 400; CU, capacity utilization rate, manufacturing; DB, monetary base, St. Louis version, first difference in logs times 400; and GX, federal government expenditure as a percentage of GNP.

[6] In Lupoletti and Webb, a similar comparison between Charles Nelson's actual ARIMA forecasts and the UVAR5 simulated forecasts found the ARIMA forecasts uniformly superior at a one-quarter horizon, the VAR forecasts uniformly superior at a four-quarter horizon, and mixed results at a two-quarter horizon.

[7] The lag lengths in both the AVAR5 and AVAR6 models were chosen by using data from 1952:2 to 1969:4. The post-sample forecasts that were prepared over a fifteen-year interval were all derived from that single setting of lag lengths for each model. It appears likely that periodic respecification of the models would have yielded more accurate forecasts.

[8] Variables examined included the foreign exchange value of the dollar, the NYSE composite price index deflated by the PCE deflator, manufacturers' unfilled orders deflated by the producer price index, business fixed investment and personal savings (both expressed as a percentage of GNP), and growth rates of employment and federal debt.

prisingly, a model containing such a poorly predicted variable produced less accurate forecasts of the other variables at four- and eight-quarter horizons, since the poor forecasts of one variable added noise to other forecasts.[9]

In brief, evidence presented in this section suggests that specifying the lag lengths in VAR models by using the AIC can improve the accuracy of forecasts. The benefits are twofold. First, there is a substantial reduction in the number of estimated coefficients. That reduction allows the remaining coefficients to be estimated more accurately. At the same time, the coefficients extracting the least information from the data are the ones removed. The second benefit is that additional variables can be added once the profligate parameterization of an unrestricted VAR model is reduced. These additional variables may contain information that is not contained in the original data.

Other Techniques

The most widely used method for reducing the parameterization of VAR models was proposed by Litterman.[10] In essence, his method involves imposing prior beliefs concerning some statistical properties of the data. Those beliefs, often referred to as Bayesian priors, include such ideas as (1) macroeconomic data can be accurately described as random walks around a trend, and (2) a variable's own lags are better predictors of that variable's future values than are lags of other variables.

One example of Bayesian restrictions was imposed on UVAR6, with the results given in Table III under the heading BVAR6.[11] The comparison between AVAR6 and BVAR6 is of particular interest, since each model takes a different approach toward effectively restricting the parameterization of a VAR

[9] Of course, movements of a price determined in a competitive financial market should not be able to be predicted with such publicly available information as lagged variables. If so, that would indicate the persistence of unexploited profit opportunities.

[10] For example, see Litterman [1979] and Doan, Litterman, and Sims [1983].

[11] The exact setting is given by the RATS statement

SPECIFY(TIGHT=.1,DECAY=.9).5

which imposes a particular pattern on the lags in the model, a particular relationship between own lags and lags of other variables, and the standard deviation of the prior. No experimentation was conducted to find the setting that worked best within the sample period-it is hoped that the results are indicative of the benefits of using priors for real-time forecasting. See Doan and Litterman [1984] for more discussion of the exact meaning of the restrictions that were imposed.

In addition, it should be noted that Litterman's priors are most oft-n used for variables expressed in levels, whereas three variables in UVAR6 are expressed as percentage changes. No experimentation was conducted to determine if the form of the priors should be changed in such a case.

Table III

FORECAST ERRORS, 1970 TO 1984

Model	Horizon	Interest Rate	Real GNP	Implicit Deflator	Sum 1	Sum 2
AVAR6	1	1.20	4.76	1.89		
	4	2.44	2.96	1.89		
	8	3.57	1.92	2.20	22.83	8.5
BVAR6	1	1.20	4.44	1.75		
	4	2.48	2.45	1.94		
	8	3.74	1.60	2.61	22.21	8.5
CVAR6	1	1.19	4.71	1.91		
	4	2.26	1.96	3.09		
	8	2.92	2.20	2.56	22.80	10

Note: Column 1 contains the names of the models used to generate forecasts. Column 2 is the forecast horizon-the length of the forecast, measured in quarters. Columns 3 through 5 contain the RMSE's of post-sample forecasts. The interest rate is the 90-day Treasury bill rate. Real GNP and the implicit deflator are percentage changes at annual rates. For those two variables, 4 and 8 quarter changes ore the average change over the particular period. Column 6 contains a summary measure of model performance, namely the sum of the RMSE's for each variable at each horizon. Column 7 contains another summary measure, a point total that assigns the most accurate forecast for each variable at each horizon two points, and the second most accurate forecast one point.

model. Both methods are atheoretical from an economic point of view. The Bayesian method, however, relies on a priori statistical restrictions, whereas the method used to construct AVAR6 lets the data determine the form of the model.

A different approach to improving forecast accuracy was given by Neftci, who proposed a method for incorporating the stage of the business cycle into VAR models. He first constructed an "indicator" variable I for the unemployment rate U by letting I_t equal one if the unemployment rate in quarter t is higher than in the previous quarter, letting I_t equal minus one if the unemployment rate declined in quarter t, and letting I_t equal zero if the unemployment rate was unchanged. He then constructed a "counter" variable C by setting

$$C_t = \sum_{j=1}^{t} I_j \qquad (3)$$

The counter variable (lagged one quarter in order to avoid adding contemporaneous information) can then be added to the equations of a VAR model. Neftci found that the counter significantly improved the explanatory power of the equation explaining the unemployment rate.

Table III shows what happens when one employs Neftci's method to construct a counter variable for the capacity utilization rate and then adds that variable to equations in AVAR6. The resulting model is labeled CVAR6. Since the counter is a nonlinear transformation of the capacity utilization rate, it can add information that would not be picked up by OLS regressions containing the capacity utilization variable. It is possible that the additional information helps to incorporate the stage of the business cycle.

As Table III indicates, there are mixed results in comparing AVAR6 and the two alternatives. The two summary measures give different orderings of the three models. Looking at individual variables, AVAR6 was most accurate for predicting the implicit deflator, BVAR6 was most accurate for predicting real GNP, and CVAR6 was most accurate for predicting the T-bill rate. Without additional information it is difficult to assert with any confidence that one model is likely to outperform the others in the near future.

Conclusion

The results in this article document the improved forecast accuracy that can be obtained by restricting the parameterization of a VAR model. Although the gains are consistent, they are not dramatically large. Initial experiments with other techniques of improving forecast accuracy did not yield consistently large additional improvements.

Researchers interested in improving atheoretical forecasts may continue to investigate methods of restricting the parameterization of VARs. Further work may also examine the benefits of combining forecasts. For example, Lupoletti and Webb have documented small but consistent improvements in accuracy from combining dissimilar forecasts. At some point, however, it will be appropriate to ask if we are near the boundary of forecast accuracy, given the limited information in historic macroeconomic time series.

APPENDIX

SETTING LAG LENGTHS IN VECTOR AUTOREGRESSIONS

Consider the following equation from a vector autoregression

$$X_{v,t} = \alpha + \sum_{i=1}^{k} \sum_{j=1}^{m_i} \beta_{i,j} X_{i,t-j} + \epsilon_{v,t} \qquad (1)$$

where X is a vector of k variables, v is an integer between 1 and k, t indexes time, α is a constant term, $\beta_{i,j}$ is a coefficient, m_i is the lag length for variable i, and $\epsilon_{v,t}$ is an error term. The problem addressed in this Appendix is choosing values for the m_i's. The choice can be made by attempting to minimize the Akaike Information Criterion (AIC)

$$AIC(m_i, \ldots, m_k) = (n-p) \log \sigma^2 + 2p \qquad (2)$$

where n is the number of observations, σ^2 is the maximum likelihood estimate of the residual variance, and p is the number of estimated coefficients—α and the $\beta_{i,j}$'s.

THE PROBLEM

The AIC has been primarily used by other authors to determine the lag length in univariate autoregressions. For that task it is easy to find the minimum AIC by inspection of a small number of alternatives. For the multivariate case, finding the minimum can be more difficult. For k variables and a maximum lag of L periods, there are Lk possibilities. It can quickly become infeasible to compute the AIC for all potential alternatives.

Accordingly, some strategy for examining a subset of alternatives is necessary. Designing such a strategy is complicated by two characteristics that were observed with macroeconomic data. First, the AIC often has many local minima. Therefore, from an arbitrary starting point it is likely that a sequence of lag-length selections will fail to converge to the set of choices that yields the global minimum of the AIC. Second, the partial derivative of the AIC with respect to a particular lag length depends on the other lag lengths. Therefore it is possible that lengthening a particular lag will lower the AIC even though a shorter lag length belongs to the set that minimizes the AIC.

The strategy for selecting lag lengths in this paper has five elements: (1) Choose a starting specification. A specification is defined as a particular value (possibly zero) for the lag length of each variable that might enter the equation. (2) Lengthen the lag for each variable by one period. Consider adding a term (that is, lengthening the lag length for one variable by one period) if it lowers the AIC more than any other term examined. (3) Look several steps ahead, in order to avoid converging to a local minimum. (4) When steps (2) and (3) fail to find a lower AIC after several attempts, stop adding terms. (5) Examine the final values in each lag, to see if removing a term lowers the AIC. Each element of the strategy is discussed below. An objective of the strategy is to minimize the role of judgment in finding a specification, in addition to finding a specification that is likely to have an AIC in a reasonably small neighborhood of the global minimum.

1. Starting Specification Experimentation revealed that the choice of the starting specification would often affect the final specification. Since univariate autoregressions with a lag length of four often forecast macroeconomic data fairly well, all specification searches began with an own lag of four, a constant, and no other variables.

2. Adding Terms Alternatives to the starting specification included adding one period to the own lag, or adding one period for an additional variable. The process of lengthening each lag in turn by one period, while holding other lags constant, is the first step of the adding procedure. After looking at the effects of lengthening each lag, the term that lowered the AIC by the largest amount was added to the equation, unless it led to a cul-de-sac (as described in the next paragraph).

3. Look Ahead It was observed in some cases that, although adding one term did not lower the AIC, adding more than one did. Therefore a three-step look-ahead procedure was built into the search. That is, even if adding one term failed to lower the AIC, two additional attempts were made at lengthening that lag. Even with the three-step look-ahead, however, it was still possible to take the wrong path and reach a cul-de-sac. Therefore, when the AIC failed to decline, six more attempts were made to add terms. In each attempt one term was added, even if the AIC rose. In many cases the additional search would successfully bypass a local minimum and find an even lower value for the AIC.

4. Stop Adding When the AIC failed to decline after six rounds, no further attempts were made to add more terms. In most equations that endpoint represented a seemingly reasonable specification. For real GNP, however, the process added many variables while only reducing the AIC by a small amount. The result was an equation that appeared overparameterized. Therefore, a limit of thirty-one coefficients (the number of coefficients in each equation of the unrestricted five-variable VAR) was imposed on the GNP equation.

5. Subtracting Variables The four preceding steps produced specifications that would occasionally include values at the end of lags with suspiciously low t-statistics in regression equations. Therefore attempts were made to remove those particular terms and to recalculate the AIC. This step resulted in a lower AIC in a few cases. To mechanize the procedure, the final lagged value for each variable was removed and the AIC recalculated. If the AIC declined, the term that lowered the AIC by the greatest amount was removed and the procedure repeated. This procedure is necessary since the adding process could include a term that would make redundant a term added earlier.

CONCLUSION

The efficacy of this strategy in approaching the global minimum of the AIC is unknown. Further investigation may employ Monte Carlo studies in order to compare this strategy with other approaches to selecting lag lengths. Intuitively, this strategy has the appeal of avoiding certain pitfalls by including techniques for bypassing local minima and removing redundant variables.

The other objective, minimizing the role of judgment in specifying equations, can be more readily assessed. Judgment is used in setting the following values: the initial specification, the number of look-ahead steps, the number of repetitions of the adding procedure attempted before stopping, and the maximum number of estimated coefficients per equation. Compared to other methods of specifying equations, however, that is a small amount of judgment. The procedure is therefore compatible with the atheoretical spirit that has motivated many authors to use VAR models.

References

Akaike, Hirotugu. "A New Look at Statistical Model Identification." *IEEE Transactions on Automatic Control* AC19-6 (December 1974) : 716-23.

Doan, Thomas, and Robert Litterman. *Regression Analysis of Time Series User's Manual*. Version 4.30. Minneapolis : VAR Econometrics, 1984, section 9.4.

________ and Christopher A. Sims. "Fore&sting and Conditional Projection Using Realistic Prior Distributions." NBER Working Paper 1202, September 1983.

Fackler, James S. "An Empirical Analysis of the Markets for Goods, Money, and Credit." *Journal of Money, Credit, and Banking* 17 (February 1985) : 28-42.

Hakkio, Craig S., and Charles S. Morris. "Vector Autoregressions : A User's Guide." Research Working Paper 84/10. Federal Reserve Bank of Kansas City, November 1984.

Kuprianov, Anatoli, and William Lupoletti. "The Economic Outlook for Fifth District States in 1984 : Forecasts from Vector Autoregression Models." *Economic Review*, Federal Reserve Bank of Richmond 70 (January/February 1984) : 12-23.

Litterman, Robert B. "Techniques of Forecasting Using Vector Autoregressions." Working Paper 115, Federal Reserve Bank of Minneapolis, November 1979.

________ "Forecasting with Bayesian Vector Autoregressions : Four Years of Experience." Working Paper 259. Federal Reserve Bank of Minneapolis, August 1984.

________ "Above-Average National Growth in 1985 and 1986. *Quarterly Review,* Federal Reserve Bank of Minneapolis 8 (Fall 1984) : 3-7.

Lupoletti, William M., and Roy H. Webb. "Defining and Improving the Accuracy of Macroeconomic Forecasts : Contributions from a VAR Model." Working Paper 84-6. Federal Reserve Bank of Richmond, 1984. *Journal of Business,* forthcoming.

Meese, Richard, and John Geweke. "A Comparison of Autoregressive Univariate Forecasting Procedures for Macroeconomic Time Series." *Journal of Business and Economic Statistics* 2 (July 1984) : 191-200.

Neftci, Salih N. "Is There a Cyclical Time Unit?" Presented to the Carnegie-Rochester Conference on Public Policy, University of Rochester, April 1985.

Sims, Christopher A. "Macroeconomics and Reality." *Econometrica* 48 (January 1980) : l-48.

Webb, Roy H. "Vector Autoregressions as a Tool for Forecast Evaluation." *Economic Review,* Federal Reserve Bank of Richmond 70 (January/February 1984) : 3-11.