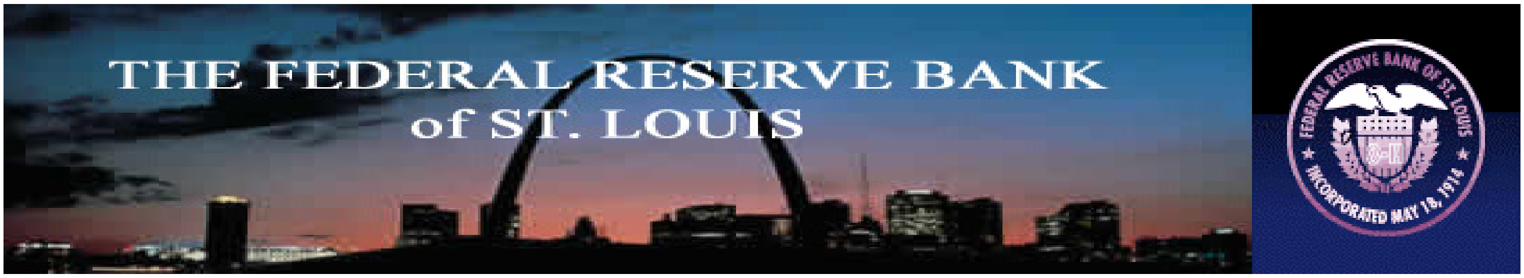

WORKING PAPER SERIES

Policy Inference Using VAR Models: The Effects of Alternative Lag Structures

R. W. Hafer
and
Richard G. Sheehan

Working Paper 1986-009B
http://research.stlouisfed.org/wp/1986/1986-009.pdf

1986
Revised February 1987

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
411 Locust Street
St. Louis, MO 63102

Revised February 1987

POLICY INFERENCE USING VAR MODELS: THE EFFECTS OF ALTERNATIVE LAG STRUCTURES

R. W. Hafer* and Richard G. Sheehan**

Federal Reserve Bank of St. Louis

86-009

* Visiting Associate Professor, Department of Economics, Washington University--St. Louis. On leave from the Federal Reserve Bank of St. Louis.

** Economist, Research and Public Information Department, Federal Reserve Bank of St. Louis.

We would like to thank Stuart Allen, Sandy Batten, James Fackler, Benjamin Friedman, Scott Hein, James Kennedy, Stefan Mittnik, Dan Thornton, Roy Webb and Arnold Zellner for useful comments and criticisms on earlier drafts of this paper. The usual caveat applies. We also thank Tom Pollmann for his research assistance.

1. INTRODUCTION

Macroeconomics has a long tradition of examining the response of different macroeconomic variables to shocks in alternative policy measures. Much recent work in this area has been based on the empirical regularities between variables in vector autoregressive (VAR) models.[1/] The VAR results in turn have been used to generate specific policy recommendations. Recent examples using VAR analysis in this way include: Sims' (1980b, 1982) results that point to a dominant role of interest rates over money in explaining the postwar variance in U.S. production;[2/] B. Friedman's (1983) analysis that shows credit to be as important as money in determining economic activity; and Chowdhury, Fackler and McMillin's (1986) use of a VAR model to argue that monetary impulses dominate fiscal policy changes in explaining investment spending.

In contrast to such VAR-based analyses, of which the foregoing are only a small sample, our purpose in this paper is to investigate the general sensitivity of policy recommendations derived from VAR models. While previous research has studied whether policy recommendations change based on the set of variables included in the model, our concern centers solely on the effects of alternative dynamic structures. Although a few authors have employed different statistically based lag structure determination procedures, we are not aware of any systematic analysis of a given multivariate VAR model to changes in the lag structure. In this sense, our study is in the spirit of Batten and Thornton (1984), who analyzed the effects of different lag structures on the "St. Louis" reduced-form income equation.

To test the sensitivity of VAR-based policy conclusions, we employ six different lag selection criteria, ranging from an often used ad hoc fixed lag structure to ones that are determined by Bayesian rules. The next section presents a brief description of the different lag length selection procedures used. Section 3 discusses the estimation of a VAR model applying the alternative procedures to quarterly macro data for the 1960-85 period. To test the sensitivity of policy across models, variance decompositions are compared in section 4. Our evidence indicates that policy conclusions drawn from a VAR model likely are specific to a particular dynamic specification. Conclusions close the paper section 5.

2. LAG LENGTH SELECTION PROCEDURES

Specification of the VAR model to be estimated requires prior determination of the included variables as well as a procedure for determining the appropriate lag lengths. Following Sims (1982), we use a relatively simple four variable model that includes the money stock, an interest rate, the price level and real output. Other variables are assumed to have no impact.

Our concern is possible differences in inference based on alternative lag structures. While many previous researchers have opted to impose equal lag lengths on each of the variables in their models,[3/] the resulting parameter estimates may suffer from bias or inefficiency, depending on whether the model is under or over parameterized. Six different approaches are used in this study to determine the lag specification. Following Batten and Thornton (1984), five are based on statistical criteria, with the exact criteria listed in table 1. We also employ one non-statistical procedure, arbitrarily fixing the maximum lag on all variables at four.[4/]

One statistical approach used to determine lag length is the standard F-test. The F-test imposes zero restrictions on lags beyond some point and tests this restricted model against a more general lag structure. The F-statistic is calculated sequentially until adding another lag does not statistically improve the fit of the model. Two of the selection procedures are based on a mean squared error criterion: Mallows' (1973) C_p statistic and Akaike's (1970) Final Prediction Error (FPE) criterion. Both trade off bias from selecting lag lengths that are too short against a loss in efficiency caused by selecting lag lengths that are too long. The FPE procedure initially was suggested by Hsiao (1979, 1981) and has been used by McMillin and Fackler (1984) and Chowdhury, Fackler and McMillin (1986) in estimating VAR models. The last two approaches use Bayesian rules of lag choice. One is Schwarz' (1978) Bayesian Information Criterion (BIC) and the other, due to Geweke and Meese (1981), is called the Bayesian Estimation Criterion (BEC). As noted by Geweke and Meese, both the BIC and BEC criteria tend to underfit in small samples, a function of the fact that the Bayesian criteria place more weight on efficiency in the bias-efficiency trade-off.[5/]

3. LAG LENGTH SELECTION RESULTS

The six lag length selection procedures are used to determine the structure of our four-variable VAR system. Because our purpose is to test the sensitivity of policy inferences from a given VAR model to alternative lag lengths, we do not experiment with changing the variables included in the model or the estimation period. Thus, the estimated VAR models differ only in terms of the lag lengths selected.

The variables included in our model are the money stock (M1), real GNP (in 1982 dollars), the implicit GNP price deflator (1982 = 100) and the three-month Treasury bill rate (hereafter T-bill). All variables except the T-bill are seasonally adjusted, quarterly values and are transformed into stationary series by taking first-differences in the logarithms; the T-bill is the first difference in its quarterly average. The estimation period is I/1960 to IV/1985.[6/]

Application of the various statistical criteria requires the necessarily arbitrary selection of a maximum allowed lag length. Geweke (1978) has argued that specifying a maximum lag that is too short may impose unwarranted zero restrictions. To simultaneously keep the lost degrees of freedom manageable and allow a reasonable lag set for the procedures to choose from, the maximum lag length was set at 12.[7/] The results obtained from imposing the lag length selection procedures on the data are reported in table 2.

The results dramatically indicate that different statistical procedures produce a variety of "optimal" lag structures. For example, the lags found on money (DM1) when inflation (DP) is the dependent variable range from five using the F-test criterion to zero based on the C_p, BIC and BEC procedures. The influence of real GNP (DRGNP) on the T-bill rate (DTB) also varies substantially, from long lags (10 or more) in the F, C_p and FPE determined models to zero lags in the Bayesian models.[8/]

The variety of lag lengths for each variable is reduced somewhat by comparing the C_p and FPE results and the BIC and BEC outcomes. Lag lengths within these two groups generally are similar. In only three instances do the C_p and FPE results differ. Likewise, the BIC and BEC criteria produce different lags in only three cases. Given the different bias-efficiency tradeoff between these two groups, however, differences in lag selection between, say, the FPE and BEC outnumber similar choices by about two to one. While this outcome may not be surprising, it reinforces our concern that the choice of lag lengths in previous VAR analyses may account for the reported divergence in results and subsequent alternative policy recommendations. In the next section we investigate the sensitivity of VAR policy inferences to changes in the estimated lag structure.

4. VARIANCE DECOMPOSITION RESULTS

We examine the sensitivity of the VAR policy results by estimating each of the VAR models using the five lag structures in table 2, as well as a model that arbitrarily imposes four lags on each variable. To compare the outcomes, we use the variance decompositions for a 20-quarter horizon.[9/]

To cast the results in a style similar to previous analyses, we report the variance decompositions based on the presumption that money is the "policy variable." Our purpose is not to discover whether money or, say, the interest rate better explains the behavior of other variables and therefore <u>should be</u> the policymakers' choice variable. Instead, presuming that M1 is the policy variable allows us to focus exclusively on the sensitivity of policy implications to lag length selection. Our conclusions are not sensitive to the choice of policy variable. Following previous research, we report the outcome for two orderings. Results obtained from ordering M1 first are reported in table 3, while table 4 presents the outcomes when M1 is ordered last.[10/]

The results in table 3 are striking. Looking down the column under DM1, the impacts of money on prices, output and interest rates vary substantially across the alternative lag structures. For example, shocks in M1 explain from 18 percent of the variation in inflation using the FPE lag structure to 0 percent of the variation when the BIC and BEC lags are used. Because much recent debate has focused on the relative roles of money and interest rates in explaining the behavior of real output, the results for M1 and the T-bill rate in explaining real output is of interest.[11/] From table 3 we find that shocks to the T-bill dominate M1 shocks in explaining the variation of output for three lag structures: the fixed-4 (20 percent vs. 6 percent), the F (41 percent vs. 13 percent) and the FPE (17 percent vs. 3 percent). The result using the C_p lag structure is close, with the edge going to the T-bill rate (13 percent vs. 10 percent). If one uses the BIC criterion, however, shocks to M1 explain more of the variation in GNP than do interest rates (12 percent vs. 0 percent). In contrast, the BEC results suggest that money (3 percent) and interest rates (0 percent) both play no apparent role in explaining the behavior of real output. In terms of accounting for output growth based on the use of just one lag length selection procedure, the results in table 3 provide some support for the hypothesis that (a) interest rates dominate money, (b) money dominates interest rates or (c) neither money or interest rates play a significant role in explaining variations in real output growth.

The results in table 3 also indicate that the effect of M1 shocks on inflation, output and the interest rate varies dramatically across the different lag structures. While the magnitude of the divergence in impact depends on the dependent variable used, the differences are large enough to preclude a firm statement about M1's role without prior specification of the lags. The appropriate monetary policy recommendations derived from these VAR models remain a matter of belief rather than being based on unambiguous empirical evidence.

The variance decompositions based on a M1-last ordering are presented in table 4. Although there are some instances where changing M1's ordering greatly influences its role in explaining the variation of the dependent variable, the changes from table 3 often are relatively small. In addition, the reordering of M1 does not remove the uncertainty about its relative effectiveness in explaining variation in the other variables. Again contrasting M1 with the T-bill, the variance decompositions show M1 shocks dominating T-bill shocks in explaining inflation for the fixed-4 and FPE lag structures and near equality for the other lag structures. Turning to the results for output, again we discover that across the six alternative lag structures, interest rate shocks dominate money shocks in explaining output three times, money dominates once and twice the outcomes are close. As in table 3, the failure to reject a hypothesis about the relative importance of money or interest rate shocks in explaining output behavior may hinge only on the choice of the lag structure used. Thus, policy recommendations derived from VAR estimates without due regard to the choice of lag length must be viewed with some skepticism.

5. CONCLUSIONS

Previous research on the usefulness of VAR models for policy analysis largely have focused on the variables included. While some have recognized that the choice of the dynamic structure is important, there has been no systematic study of the effects of alternative lag structures on policy recommendations derived from a VAR model. This paper attempts to fill that void.

The conclusions derived from the evidence presented here can be stated succinctly. First, because no hard rule exists for the selection of lags in a multivariate setting, VAR modelers must explicitly concern themselves with their choice of dynamic structure. Second, we have shown that policy recommendations obtained from a given VAR model are quite sensitive to different lag structures. Our evidence does not preclude the possibility of selecting one policy recommendation over another purely because of different dynamic structures. VAR modelers should check the sensitivity of their results not only to alternative orderings for the variance decompositions but also, more fundamentally, to the choice of lags. While this outcome is undoubtedly shared by other empirical models, it nonetheless raises another problem that limits the usefulness of VAR models for macroeconomic policy analysis.

FOOTNOTES

1/ The popularity of VAR models stems, in part, from their supposed lack of *a priori* theory in constructing the model. This feature of VAR models has been addressed by Cooley and LeRoy (1985), who argue that interpreting VAR models as structural representations imposes "theoretical" identifying restrictions on the model that generally are not supported. See also related discussions by Leamer (1985) and Braun and Mittnik (1985).

2/ King (1983) investigates Sims' findings and concludes that Sims' conclusions are derived from a misspecified model. Sims' results are based on a model estimated in log levels which, King argues, violates the assumption of stationarity in the data. King's reestimation of the model in first-differences of logarithms reverses Sims' conclusion that interest rates dominate money in explaining the behavior of output during the postwar period.

3/ For example, Fischer (1981) fixes lags to be three quarters for each variable, Sims (1980b) fixes lags at four for each variable, based on a test between four or eight, Friedman (1983) fixes the lags to be eight quarters and King (1983) fixes the lags at six quarters for each variable.

4/ The following discussion draws from Batten and Thornton (1984).

5/ This feature of the Bayesian procedures was noted in tests of the St. Louis equation by Batten and Thornton (1984) and in their tests of the money-income relationship

[Thornton and Batten (1985)]. Geweke and Meese (1981) also have shown using Monte Carlo experiments that the probability of underfitting the model is quite large (50 percent) in small samples.

6/ Use of the 1960-85 sample period may bother some readers concerned about the effects of estimating the models across the pre- and post-1979 periods. While such a concern is valid if we were attempting to specifically establish some policy recommendation, the force of such criticism is mitigated under current circumstances. Again, we are not concerned about any one specific policy outcome, but only about the sensitivity of a representative policy recommendation derived from a VAR model.

7/ Batten and Thornton (1984) found that, except possibly for the F-test, there was little sensitivity in the chosen lags when the maximum lag was extended beyond 12.

8/ The shorter lag-length models are nested within the larger and are subject to testing against the larger model. Thornton and Batten (1985) have noted, however, that application of a standard F or χ^2 test for lag length restrictions imposes a different criteria that confounds the bias-efficiency trade-off characterizing each lag selection procedure. Consequently, we treat the different criteria's lag selection as if they are the "true" lag structures.

9/ VAR models that impose equal lag lengths across the variables are estimated using OLS. Because other models use various lag lengths in different equations, the estimation procedure is GLS. The estimated equations are used to generate both impulse response functions and variance decompositions. To conserve space, only the variance decompositions are reported below. Our basic conclusion, that the choice of lag length selection procedure may substantially influence policy inference, follows both from the alternative variance decompositions and the impulse response functions. In addition, since our quantitative results are not sensitive to horizon length, we report only the outcomes for the 20-quarter horizon.

10/ To conserve space, the correlation matrices among the innovations are not reported. It should be noted, however, that these correlations are all relatively low.

11/ See, among others, Sims (1980a, b), King (1983) and Litterman and Weiss (1985).

REFERENCES

Akaike, H. "Statistical Predictor Identification," Annals of the Institute of Statistical Mathematics (No. 2, 1970), pp. 203-17.

Batten, D. S., and D. L. Thornton. "How Robust are the Policy Conclusions of the St. Louis Equation?: Some Further Evidence," Federal Reserve Bank of St. Louis Review (June/July 1984), pp. 26-32.

Braun, P. A., and S. Mittnik. "Structural Analysis with Vector Autoregressive Models: Some Experimental Evidence," Working Paper 84, Department of Economics, Washington University (December 1985).

Chowdhury, A. R., J. S. Fackler and W. D. McMillin. "Monetary Policy, Fiscal Policy, and Investment Spending: An Empirical Analysis," Southern Economic Journal (January 1986), pp. 794-806.

Cooley, T. F., and S. F. LeRoy. "Atheoretical Macroeconometrics: A Critique," Journal of Monetary Economics (November 1985), pp. 283-308.

Fackler, J. S., and S. C. Kreiger. "An Application of Vector Time Series Techniques to Macroeconomic Forecasting," Journal of Business and Economic Statistics (January 1986), pp. 71-80.

Fischer, S. "Relative Shocks, Relative Price Variability, and Inflation," Brookings Papers on Economic Activity (No. 2, 1981), pp. 381-431.

Friedman, B. "The Roles of Money and Credit in Macroeconomic Analysis," in J. Tobin, ed., Macroeconomics, Prices and

Quantities: Essays in Memory of Arthur M. Okun (Washington: The Brookings Institution, 1983).

Geweke, J. "Testing the Exogeneity Specification in the Complete Dynamic Simultaneous Equation Model," _Journal of Econometrics_ (1978), pp. 163-85.

Geweke, J., and R. Meese. "Estimating Regression Models of Finite But Unknown Order," _International Economic Review_ (February 1981), pp. 55-70.

Gordon, R. J., and S. R. King. "The Output Cost of Disinflation in Traditional and Vector Autoregressive Models," _Brookings Papers on Economic Activity_ (No. 1, 1982), pp. 205-42.

Hsiao, C. "Autoregressive Modeling of Canadian Money and Income Data," _Journal of the American Statistical Association_ (September 1979), pp. 533-60.

______. "Autoregressive Modeling and Money Income Causality Detection," _Journal of Monetary Economics_ (January 1981), pp. 85-106.

King, S. R. "Interest Rates and the Influence of Money on Output and Prices," unpublished manuscript, Stanford University (August 1983).

Leamer, E. E. "Vector Autoregressions for Causal Inference?" _Carnegie-Rochester Conference Series on Public Policy: Understanding Monetary Regimes_ (Vol. 22, Spring 1985), pp. 255-304.

Litterman, R. B., and L. Weiss. "Money, Real Interest Rates, and Output: A Reinterpretation of Post-War U.S. Data," _Econometrica_ (January 1985), pp. 129-56.

Mallows, C. L. "Some Comments on C_p," Technometrics (November 1973), pp. 661-75.

McMillin, W. D., and J. S. Fackler. "Monetary vs. Credit Aggregates: An Evaluation of Monetary Policy Targets," Southern Economic Journal (January 1984), pp. 711-23.

Schwarz, G. "Estimating the Dimension of a Model," The Annals of Statistics (March 1978), pp. 461-64.

Sims, C. "Policy Analysis with Econometric Models," Brookings Papers on Economic Analysis (No. 1, 1982), pp. 107-52.

______. "Macroeconomics and Reality," Econometrica (January 1980a), pp. 1-48.

______. "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered," American Economic Review, Papers and Proceedings (May 1980b), pp. 250-57.

Thornton, D. L., and D. S. Batten. "Lag-Length Selection and Tests of Granger Causality Between Money and Income," Journal of Money, Credit and Banking (May 1985), pp. 164-78.

Table 1
Estimation Criteria for Lag Length Selection

Criterion	Procedure	Definition of estimation criterion
F-test	min j such that $F(j) > F_{0.05}$, $j = 0, 1, \ldots, L$	$F(j) = \frac{(SSR_{L-j-1} - SSR_L)(T-L-1)}{SSR_L(j+1)}$
C_p	min $C_p(i)$ $i = 0, 1, \ldots, L$	$C_p(i) = \frac{SSR_i(T-L-1)}{SSR_L} - T + 2(i+1)$
FPE	min FPE(i) $i = 0, 1, \ldots, L$	$FPE(i) = \frac{SSR_i}{T}\left[\frac{T+i+1}{T-i-1}\right]$
BIC	min BIC(i) $i = 0, 1, \ldots, L$	$BIC(i) = \ln\left(\frac{SSR_i}{T-i-1}\right) + \frac{(i+1)\ln T}{T}$
BEC	min BEC(i) $i = 0, 1, \ldots, L$	$BEC(i) = \frac{SSR_i}{T-i-1} + \frac{SSR_L(i+1)}{T-L-1}\left(\frac{\ln T}{T-L-1}\right)$

where $j \equiv$ number of lags constrained to equal zero,

$T \equiv$ sample size,

$L \equiv$ maximum allowed lag length,

$i \equiv L-j =$ number of non-zero lags and

$SSR_k \equiv$ residual sum of squares from the model estimated with k lags.

Table 2
Alternative Lag Structures
Period: I/1960 - IV/1985

Dependent variable	Selection procedure	Independent variable/lag			
		DP	DM1	DRGNP	DTB
DP	F	0	5	5	12
	Cp	3	0	0	1
	FPE	3	1	0	1
	BIC	2	0	0	1
	BEC	2	0	0	1
DM1	F	7	1	12	12
	Cp	1	9	3	1
	FPE	1	9	3	1
	BIC	0	1	0	1
	BEC	0	1	0	1
DRGNP	F	12	12	12	12
	Cp	3	2	0	6
	FPE	3	2	0	7
	BIC	1	1	0	0
	BEC	1	0	0	0
DTB	F	1	0	12	9
	Cp	0	7	10	3
	FPE	0	12	10	3
	BIC	0	4	0	3
	BEC	0	1	0	0

Notes: All variables except the interest rate are measured as logarithmic changes. The interest rate is measured as the simple first-difference. P is the implicit GNP deflator (1982=100), M1 is the narrow money stock, RGNP is real GNP (1982 dollars) and TB is the three-month Treasury bill rate.

Table 3
Variance Decompositions (20-quarter horizon)
Alternative Lag Structures
Ordering: M1 First

Variance in:	Selection procedure	Explained by (percent): DM1	DP	DRGNP	DTB
DM1	FIX-4	53	8	5	33
	F	47	7	4	41
	Cp	86	2	6	5
	FPE	88	0	1	11
	BIC	97	0	0	3
	BEC	96	0	0	4
DP	FIX-4	3	67	2	29
	F	16	48	8	29
	Cp	1	99	0	0
	FPE	18	62	12	7
	BIC	0	99	0	1
	BEC	0	99	0	1
DRGNP	FIX-4	6	14	60	20
	F	13	9	37	41
	Cp	10	24	53	13
	FPE	3	9	71	17
	BIC	12	17	70	0
	BEC	3	11	87	0
DTB	FIX-4	18	2	2	77
	F	8	3	8	80
	Cp	17	4	11	68
	FPE	2	8	5	86
	BIC	25	0	1	74
	BEC	2	0	0	98

Notes: All variables except the interest rate are measured as logarithmic changes. The interest rate is measured as the simple first-difference. P is the implicit GNP deflator (1982=100), M1 is the narrow money stock, RGNP is real GNP (1982 dollars) and TB is the three-month Treasury bill rate.

Table 4
Variance Decompositions (20-quarter horizon)
Alternative Lag Structures
Ordering: M1 Last

Variance in:	Selection procedure	Explained by (percent): DP	DRGNP	DTB	DM1
DP	FIX-4	67	1	2	9
	F	49	8	20	22
	Cp	100	0	0	0
	FPE	89	0	1	11
	BIC	99	0	1	0
	BEC	99	0	1	0
DRGNP	FIX-4	14	62	17	7
	F	8	37	43	11
	Cp	23	55	12	10
	FPE	18	61	12	8
	BIC	17	73	1	9
	BEC	11	89	0	0
DTB	FIX-4	2	3	87	8
	F	3	8	86	3
	Cp	5	10	71	15
	FPE	3	9	71	17
	BIC	0	3	86	11
	BEC	0	0	100	0
DM1	FIX-4	8	6	30	57
	F	8	6	42	43
	Cp	2	7	5	86
	FPE	2	8	5	85
	BIC	0	3	7	90
	BEC	0	3	4	93

Notes: All variables except the interest rate are measured as logarithmic changes. The interest rate is measured as the simple first-difference. P is the implicit GNP deflator (1982=100), M1 is the narrow money stock, RGNP is real GNP (1982 dollars) and TB is the three-month Treasury bill rate.