Federal Reserve Bank of Minneapolis
Research Department Staff Report 107

February 1987

VECTOR AUTOREGRESSIONS AND REALITY

David E. Runkle*

Federal Reserve Bank of Minneapolis
and Brown University

ABSTRACT

The statistical significance of variance decompositions and impulse reponse functions for unrestricted vector autoregressions is questionable. Most previous studies are suspect because they have not provided confidence intervals for variance decompositions and impulse response functions. Here two methods of computing such intervals are developed, one using a normal approximation, the other using bootstrapped resampling. An example from Sims' work illustrates the importance of computing these confidence intervals. In the example, the 95 percent confidence intervals for variance decompositions span up to 66 percentage points at the usual forecasting horizon.

*Thanks to the National Science Foundation for its support of this research through an NSF Graduate Fellowship. Thanks also to J. Hausman, G. Lamson, N. G. Mankiw, J. Poterba, D. Quah, J. Rotemberg, and R. Solow for comments on an earlier draft and to K. Rolfe for editorial assistance.

1. Introduction

Explaining the relationships among money, interest rates, prices, and output is one of the most important challenges in macroeconomics. Traditionally, economists have tried to explain those relationships using structural models which impose a priori restrictions on the intercorrelations of the data. Recently, Sims (1980a) has championed a different approach to understanding these economic relationships. He claims that unrestricted vector autoregressions (VARs) provide a better understanding of macroeconomic relationships than structural models because structural models use "incredible" identifying restrictions (Sims 1980a, p.1).

Vector autoregressions have brought with them their own terminology and their own folklore. Granger-causality, variance decompositions, innovation accounting, and impulse response functions fill the places in this methodology that parameter estimates, identifying restrictions, and hypothesis testing do in traditional economic modeling.

In this paper I try to determine whether unrestricted vector autoregressions can in fact help economists understand the empirical relationships among money, interest rates, prices, and output. I do not examine the theoretical validity of using vector autoregressions to understand macroeconomic questions. That issue has been addressed by others, including Cooley and LeRoy (1985). I also do not question the usefulness of Bayesian vector autoregressions. (They have been defended by, for example, Litterman 1986.)

What I do is re-examine the inferences that Sims (1980b) has drawn about these macroeconomic relationships. I ask two types of questions about his results: Do they depend on incorrectly or arbitrarily chosen data? And are they statistically significant?

My answers suggest that unrestricted vector autoregressions often do not tell much about interesting macroeconomic questions. Sims' results are sensitive to data selection in unreasonable ways. And the confidence intervals for their variance decompositions and impulse response functions are often so large that little useful inference can rely on them.

This paper is organized as follows: Section 2 sets up the theoretical framework for an empirical evaluation of Sims' results. Section 3 examines Sims' particular results within that framework. Section 4 presents empirical results using different specifications. Section 5 concludes.

2. Theory

Computing Variance Decompositions and Impulse Response Functions

Before examining empirical claims that vector autoregressions explain the relationships among macroeconomic variables, I will review how variance decompositions and impulse response functions are computed.

Consider the autoregressive representation

$$x_t = b(L)x_t + e_t \tag{1}$$

where x_t is a stationary stochastic vector process, L is the lag operator, and e_t is the vector of innovations to x at time t. If such a representation exists, the roots of $\det\,[I - b(z)] = 0$ have modulus greater than one. This insures that $[I - b(z)]$ is invertible.

Although VAR estimation is based on this autoregressive representation, most interpretations of VARs are based on the vector moving average representation

$$x_t = n_t + a(L)e_t, \qquad E(e_t) = 0 \tag{2}$$

$$E(e_t e_{t-k}^{'}) = \begin{cases} W, & |k| = 0 \\ 0, & |k| \neq 0 \end{cases}$$

where n_t is perfectly predictable and the matrix of coefficients of $a(L)$ at lag zero is the identity matrix. The Wold decomposition theorem shows that the vector of errors e_t is the forecast error of the autoregression given information available at $t-1$ if the roots of $a(z)$ lie outside the unit circle.

To get from the moving average representation to impulse response functions and variance decompositions requires a normalization. To quantify the cumulative response of an element of x_t—say, industrial production—to an unpredicted innovation in some component of e_t, the components of e_t must be orthogonal. Since the sample covariance matrix W is unlikely to be nearly diagonal, the covariance of the residuals must be arbitrarily divided in some way so that the errors themselves are orthogonal. The usual convention is to adopt some particular ordering and allocate any correlation between the residuals of any two elements to the variable that comes first in the ordering.

The variance decomposition is simply a function of the moving average representation. The variance decomposition of the k-step-ahead forecast is the proportion of the total forecast variance of one component of x_{t+k}—say, industrial production—due to shocks to the moving average representation of another variable.

Because variance decompositions and impulse responses are functions of the underlying parameters of the autoregressive representation, users of VARs could compute asymptotic standard errors for the estimates of decompositions and response functions. But because these estimates are complicated functions of the autoregressive parameters, this has not been done. Some authors (for example, Sims 1980b and Fischer 1981) have estimated empirical confidence intervals for impulse response functions based on Bayesian methods, but they are rarely used in the VAR literature. Also, no author has reported even empirical confidence intervals for variance decompositions. Supplying impulse response functions or variance decompositions without confidence intervals is tantamount to using regression coefficients without t-statistics.

A simple example shows why reporting confidence intervals for variance decompositions and impulse response functions is important. Suppose that y_t has the univariate autoregressive representation

$$y_t = ay_{t-1} + e_t \tag{3}$$

where $|a| < 1$ and e_t is white noise. The moving average representation of (3) is

$$y_t = 1/(1 - aL)e_t \tag{4}$$

where L is the lag operator. The impulse response function is identical to the moving average representation. The response of y_{t+2} to a shock in e_t is a^2. According to standard asymptotic theory, if $t^{1/2}(\hat{a} - a)$ is distributed as $N(0, s^2V^{-1})$, then $t^{1/2}[g(\hat{a}) - g(a)]$ is distributed as $N(0, s^2GV^{-1}G')$, where s^2 is the variance of e_t, V is the probability limit of $y'y/n$, and $G = dg(a)/da$. Hence, a standard deviation confidence interval for the n^{th} term in the impulse response function is

$$\hat{a}^n \pm 2n\hat{a}^{n-1}[\hat{s}^2(y'y)^{-1}]^{1/2}. \tag{5}$$

Note that this confidence interval implies that the asymptotic t-statistic of the n^{th} term in the impulse response function goes to zero at the rate a/n. After only a few periods, the size of the confidence interval grows dramatically compared to the size of the coefficient. The current shock e_t gets multiplied by a^n, so its importance goes to zero.

An extension of this logic to the vector case suggests why the standard errors for VAR variance decompositions and impulse response functions also should be computed. Those decompositions and response functions are simply nonlinear functions of the autoregressive parameters and their covariance matrix. If the individual coefficients in the unconstrained vector autoregressions are insignificant, their large standard errors imply large and growing standard errors for the estimates of variance decompositions and impulse response functions. Typically, tests that constrain all the coefficients of one or more variables to be zero in one part of a VAR do not reject that null hypothesis. Therefore, when those high variances are nonlinearly expanded, the result can be huge confidence bounds.

Computing Confidence Intervals

The standard errors for the variance decompositions of vector autoregressions are computed here using two different methods: first, using a normal approximation to the distribution of the parameters of the variance decomposition; second, using Efron's

(1982) bootstrap method to generate confidence intervals based on the empirical distribution of the residuals from the vector autoregression. Because variance decompositions have been reported more often than impulse response functions, I concentrate on them here. But the same methods can be used to compute confidence intervals for impulse response functions.

The underlying logic for using the normal approximation estimator is quite conventional. The parameters of the variance decomposition are functions of the autoregressive parameters of the VAR. The covariance matrix of those autoregressive parameters is $[S \otimes (X'X)^{-1}]$, where S is the covariance matrix of the disturbance terms in the vector autoregression. Although a VAR is efficiently estimated using OLS, the covariance matrix for the autoregressive parameters is not the same as that for OLS because of the cross-equation covariances.

Let $g(b)$ be the function that transforms the autoregressive parameters into the parameters of the variance decomposition. If $t^{1/2}(\hat{b} - b)$ is distributed as $N[0, (S \otimes V^{-1})]$, where $V = \text{plim } X'X/t$, then $t^{1/2}[g(\hat{b}) - g(b)]$ is distributed as $N[0, G(S \otimes V^{-1})G']$, where $G = dg(b)/db$. This is the normal approximation used to form confidence intervals of the variance decompositions. Deriving the analytic form of $g(b)$ is very hard, but doing that is not necessary to form the confidence intervals. Numerical differentiation can be used to estimate G without such an analytic representation.

Because normal approximations to nonlinear functions often perform poorly in small samples, I also use bootstrapping to generate confidence intervals for variance decompositions. The basic insight behind the bootstrap is that since the estimated residuals of the model are a representative sample of the true disturbances, the order in which the disturbances occur should not matter. This means that the distribution of the estimator can be determined by generating many artificial observations from the actual data and the estimated residuals.

To determine confidence intervals for the variance decompositions of a vector autoregression (using my earlier notation), start by estimating the equation $X_{t+1} = b(L)X_t + e_{t+1}$ and saving the estimated residuals and parameters. Then take the first m observations (where m is the number of lags in the VAR) as initial conditions, and generate an artificial realization of X_{m+1} by randomly selecting one of the disturbance terms—say, $\hat{e}_r$— and forming $X_m + \hat{e}_r$. Repetition of this process will generate a complete series of observations for X. To get a bootstrap estimate of the confidence intervals of the variance decompositions, generate a large number of series of observations—say, 1,000. After each series is generated, estimate all the parameters of interest, including the variance decompositions. Store these for future analysis. When this is done for all 1,000 replications, compute empirical confidence intervals for single parameters or composite hypotheses. (Below I report 95 percent empirical confidence intervals.)

Using the normal approximation for computing confidence intervals for variance decompositions has one problem. Although the bootstrap simultaneously imposes the two constraints that the variance decomposition parameters each lie between 0 and 100 percent and the parameters for each equation add up to 100 percent, the normal approximation does not. It imposes the adding up constraint, but not the truncation at 0 and 100. In practice, this does not seem to have much of an effect. (I discuss this fact below, along with the empirical results.)

3. Testing Sims' Results

Sims (1980b) tries to use VARs to test whether monetary policy has a significant effect on output. If money is not neutral, monetarist theory says that either changes in the rate of growth of the money supply or unanticipated changes in money will affect subsequent output.

The source of this monetary nonneutrality, however, differs with the preferred version of monetarist theory: traditional or rational expectations. Traditional monetarists have emphasized the effect of the rate of monetary growth on output. Friedman and Schwartz (1963), for example, provide evidence that for much of U.S. history changes in the rate of growth of money have preceded changes in output. Adherents of this kind of monetarism do not typically explore the theoretical justification for this form of nonneutrality; they seem to find the empirical regularity (of movements in output following movements in money) sufficient. Rational expectations monetarists have, instead, focused on the price effect of unanticipated changes in the money supply. This version of monetarism claims that a surprise change in money affects output because people are fooled into thinking that the resulting change in the price they are paid for their goods is a change in relative prices.

Sims suggests that both types of monetarism can be tested by examining whether a significant proportion of the variance in output over some period is due to money. He thinks vector autoregressions can show whether or not this is true. If either form of monetarism is correct, Sims claims, the variance decomposition of output should show that money explains a large percentage of output variation.

Sims uses a four-variable vector autoregression to test monetarism. He uses U.S. monthly data for 1947-78 on the logs of the U.S. commercial paper rate, M1, the wholesale price index, and industrial production. Sims' results of a variance decomposition of that VAR with the ordering interest rates, money, prices, and output are in my Table 1. They suggest that money has little effect on output: Only 4 percent of the variance in output is explained by innovations in money. The evidence seems to suggest instead that innovations in nominal interest rates are the driving force behind movements in both output and money. Innovations in the commercial paper rate explain 30 percent of the variance in industrial production and 56 percent of the variance in M1.

While these estimates of variance decompositions are suggestive, Sims provides no indication of their distribution. I re-estimated Sims' equations using quarterly data for the same time period, then computed 95 percent confidence intervals using both the normal approximation and bootstrapping. The bootstrap estimates are based on a sample of 1,000 replications.

The results of neither of these methods confirm Sims' claim that innovations in nominal interest rates explain much of the variation in money and output. Figures 1 and 2 illustrate this. They display the point estimates of the percentage of the variance in money (Figure 1) and output (Figure 2) due to innovations in the interest rate along with the empirical bootstrap 95 percent confidence intervals for the estimates. As both figures indicate, the confidence intervals grow rapidly, but seem to stabilize, as do the point estimates, within about two years. After 16 quarters (the standard forecast length for variance decompositions) the confidence intervals are about 50 percentage points wide. The lower bound in Figure 2 is close to zero, which means that the hypothesis that interest rate innovations have almost no effect on output cannot be rejected.

Because the point estimates for bootstrapping must lie between 0 and 100 percent, asking whether innovations in one variable do not cause any of the variation in another variable is not appropriate. A better question is whether those innovations explain less than some small portion—say, 10 percent—of the variance in another variable. For example, using the bootstrap, I cannot reject at the 95 percent confidence level the hypothesis that the sum of the percentage of the variance in industrial production explained by innovations in the nominal variables (money and prices) is less than 10 percent.

For comparison, the confidence intervals for the 16-quarter-ahead variance decomposition using both methods are shown in Table 2. (Programs to compute both kinds of confidence intervals are available from me.) Both estimates for the parameters that Sims emphasizes are huge. This questions much inference based on variance decompositions. For example, the confidence intervals show that a negligible percentage of the variance in M1 is attributable to innovations in industrial production. This seems to suggest that monetary growth is not influenced by fluctuations in output.

The similarity of the confidence intervals for the two methods is remarkable. They only seem to be very different when the point estimate is close to zero. That is understandable. Because the normal approximation does not take into account the skewness of the parameter distribution at such a point, the normal confidence intervals include negative numbers. But since the parameter is of no practical significance in such cases, this error is inconsequential.

Since the normal approximation works so well (and is so much cheaper than bootstrapping), I use it for all other estimates in this paper. I adopt the convention of truncating the lower confidence interval. If the estimated lower bound is negative, it is reported as zero. (In general, bootstrapping probably should be preferred, but the computational time it takes is considerable. Each bootstrap estimate computed for this paper on a Prime 850 Computer used 8 hours of CPU time, while each normal approximation used about 3 minutes. As computing power becomes cheaper, bootstrap estimates will probably become more commonplace.)

Since Sims' (1980b) results originally came from monthly data, I re-estimated his original specification using monthly data to compute the confidence intervals for variance decompositions at the 48-month-horizon, the one Sims uses. The results are in Table 3. They are similar to those in Table 2. The confidence intervals on the parameters Sims emphasizes, those in Figures 1 and 2, are very wide. I also cannot reject at the 95 percent level either the hypothesis that innovations to money and prices together explain less than 5 percent of the variance in industrial production or the hypothesis that innovations to interest and money together explain less than 5 percent of that variance. I do reject the hypothesis that all three together explain less than 5 percent of that variance.

4. Changing Sims' Specification

So far I have examined only results using Sims' specification. Supporters of Sims' results might still claim that because the point estimates of the variance decompositions are still the best estimates of the effect of innovations in one variable on another, his conclusions should still be accepted. But this position is untenable if both the point estimates and the confidence intervals change markedly with minor changes in the specification. I will here consider two changes in Sims' specification: using U.S. Treasury bill interest rates instead of U.S. commercial paper interest rates and using a trend instead of no trend.

Since most empirical macroeconomic analysis implicitly assumes that the results do not depend on which interest rate is used, substituting Treasury bill yields for commercial paper yields should not make a difference. In fact, as long as no trend is included, it does not. Tables 4 and 5 report the variance decompositions of VARs using both commercial paper and Treasury bill yields along with data for money, prices, and output. A comparison of these and earlier tables reveals that many of the variance decompositions are very sensitive to the addition of trends. This has been noted by King (1982), but he does not report any confidence intervals. The addition of a trend makes money seem more important in explaining the variance of output; in fact, in Table 5, variance in innovations in money appears to explain more of the variance in output than do innovations in interest rates. But the large confidence intervals of the interesting components of the variance decompositions suggest that regardless of which point estimate considered, this analysis provides very little support for drawing strong conclusions about the effect of innovations in these time series.

Economic theory may suggest reasons for including or excluding a trend in a particular regression. When it does not, the sensitivity of the coefficients to that change should be noted. One possible reason for this sensitivity is that explosive roots may be present without trends, but trends eliminate them. This would question inference based on the VARs without trends. However, note that the inclusion of trends does not noticeably affect the width of the confidence intervals for variance decompositions.

These results should not be surprising. Since the variance decompositions are merely functions of the variance of the 16-quarter-ahead forecast, if the forecasting error is large, then so will be the confidence interval for the variance decomposition. The forecasting performance of unrestricted vector autoregressions has been shown to be poor after about one year. (See, for example, Fair 1979.) Therefore, attempts to estimate the long-run impact of an innovation in one variable are subject to enormous error.

5. Conclusion

The evidence presented here suggests that drawing strong conclusions about the interrelationship of interest rates, money, prices, and output from unrestricted vector autoregressions is more difficult than some seem to think. The standard errors of the variance decompositions in Sims' (1980b) specification are too large to easily make inferences about them. Alternative specifications change the parameter estimates, but the standard errors are still huge. In other words, vector autoregressions may let the data speak for themselves, but the data are not talking very loudly.

Table 1

SIMS' RESULTS

Percentage of Variance in Four Variables
Explained by Innovations in Each of Them

(48-Month Horizon, 1947-78, No Trends)

	By Innovations in			
Variable Explained	CP	M1	WPI	IP
Commercial Paper Rate	50	19	4	28
M1	56	42	1	1
Wholesale Price Index	2	32	60	6
Industrial Production	30	4	14	52

Source: Sims 1980b, p. 253

Table 2

MY RESULTS, USING QUARTERLY DATA

Percentage of Variance in Four Variables
Explained by Innovations in Each of Them:
Point Estimates and Two 95% Confidence Intervals

(16-Quarter Horizon, 1947-78, No Trends)

	By Innovations in			
Variable Explained	CP	M1	WPI	IP
Commercial Paper Rate				
Point Estimate	51	20	4	24
Normal Interval	(31-71)	(0-41)	(-7-14)	(4-44)
Bootstrap Interval	(31-68)	(3-9)	(1-24)	(7-44)
M1				
Point Estimate	55	44	1	1
Normal Interval	(17-93)	(5-83)	(-4-6)	(-3-4)
Bootstrap Interval	(17-83)	(9-76)	(0-18)	(0-16)
Wholesale Price Index				
Point Estimate	2	34	58	6
Normal Interval	(-5-9)	(-1-68)	(22-93)	(-8-20)
Bootstrap Interval	(0-21)	(4-55)	(26-82)	(1-37)
Industrial Production				
Point Estimate	32	7	12	48
Normal Interval	(2-62)	(-8-21)	(-11-36)	(13-83)
Bootstrap Interval	(8-59)	(1-31)	(1-41)	(18-74)

Table 3

MY RESULTS, USING MONTHLY DATA

Percentage of Variance in Four Variables
Explained by Innovations in Each of Them:
Point Estimates and Normal 95% Confidence Intervals

(48-Month Horizon, 1947-78, No Trends)

	By Innovations in			
Variable Explained	CP	M1	WPI	IP
Commercial Paper Rate				
Point Estimate	50	16	3	30
Confidence Interval	(29-71)	(0-36)	(0-52)	(9-52)
M1				
Point Estimate	50	49	0	1
Confidence Interval	(11-89)	(10-88)	(0-1)	(0-5)
Wholesale Price Index				
Point Estimate	1	33	58	7
Confidence Interval	(0-6)	(0-68)	(23-92)	(0-17)
Industrial Production				
Point Estimate	29	5	12	54
Confidence Interval	(0-58)	(0-17)	(0-35)	(18-88)

Tables 4 and 5

RESULTS OF CHANGING SIMS' SPECIFICATION

Percentage of Variance in Four Variables
Explained by Innovations in Each of Them:
Point Estimates and Normal 95% Confidence Intervals

(16-Quarter Horizon, 1947-78)

Table 4

Using a Trend

	By Innovations in			
Variable Explained	CP	M1	WPI	IP
Commercial Paper Rate				
Point Estimate	59	29	1	11
Confidence Interval	(36-81)	(6-52)	(0-6)	(0-25)
M1				
Point Estimate	54	44	1	1
Confidence Interval	(14-94)	(5-85)	(0-3)	(0-3)
Wholesale Price Index				
Point Estimate	3	35	60	3
Confidence Interval	(0-13)	(0-72)	(22-98)	(0-10)
Industrial Production				
Point Estimate	34	22	11	33
Confidence Interval	(10-57)	(4-41)	(0-29)	(15-60)

Table 5

Using a Trend and Treasury Bill Rates

	By Innovations in			
Variable Explained	T-Bill	M1	WPI	IP
Treasury Bill Rate				
Point Estimate	73	19	1	7
Confidence Interval	(51-95)	(0-40)	(0-4)	(0-17)
M1				
Point Estimate	36	62	1	1
Confidence Interval	(0-72)	(25-99)	(0-10)	(0-5)
Wholesale Price Index				
Point Estimate	1	40	54	5
Confidence Interval	(0-6)	(2-77)	(17-91)	(0-19)
Industrial Production				
Point Estimate	27	28	12	32
Confidence Interval	(4-51)	(8-49)	(0-32)	(15-50)

Figures 1 and 2

PERCENTAGE OF VARIANCE IN MONEY AND OUTPUT
EXPLAINED BY INNOVATIONS IN THE COMMERCIAL PAPER RATE

——— Point Estimates

------- Bootstrap 95% Confidence Intervals

Figure 1 M1

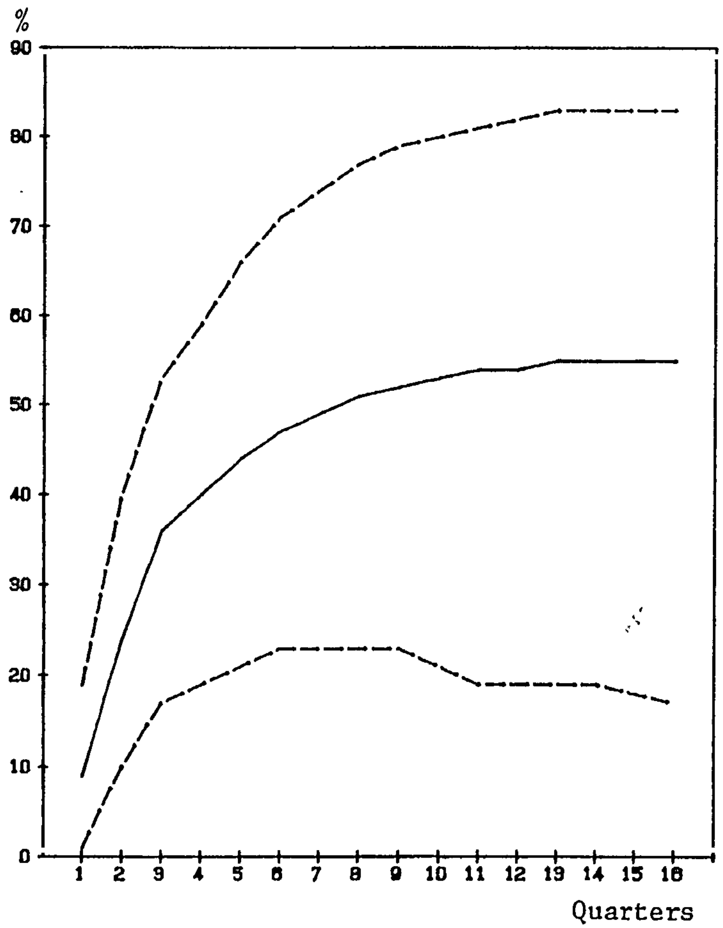
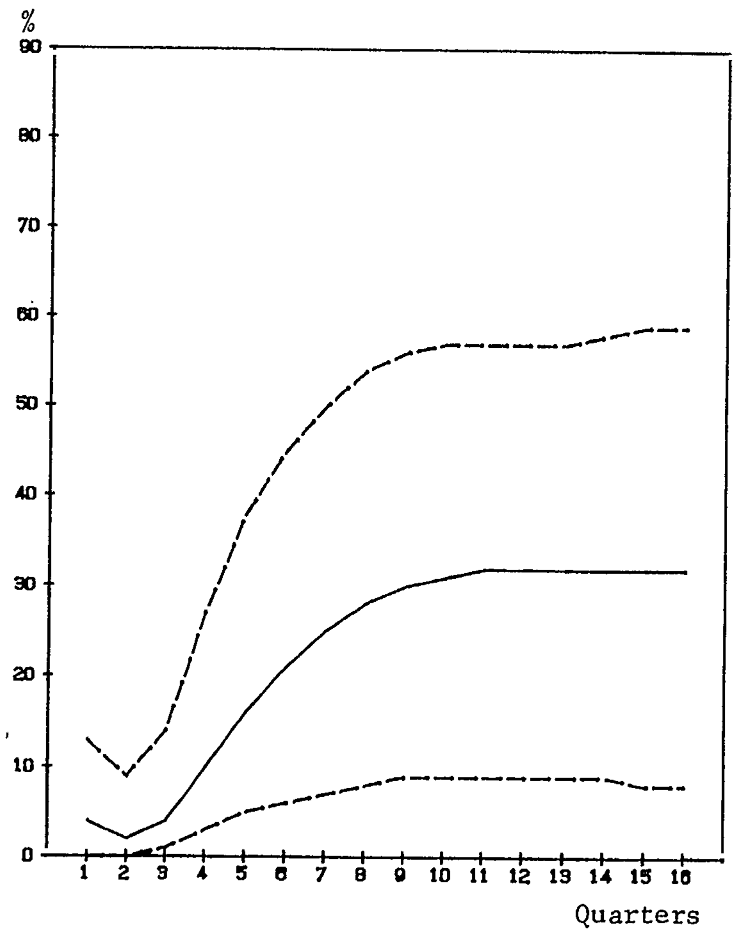


Figure 2 INDUSTRIAL PRODUCTION



References

Cooley, T.F., and LeRoy, S.F.(1985), "Atheoretical Macroeconometrics: A Critique," *Journal of Monetary Economics*, 16, 283-308.

Efron, B.(1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, Philadelphia: Society for Applied and Industrial Mathematics.

Fair, R.C.(1979), "An Analysis of the Accuracy of Four Macroeconomic Models," *Journal of Political Economy*, 87, 701-18.

Fischer, S.(1981), "Relative Shocks, Relative Price Variability, and Inflation," *Brookings Papers on Economic Activity*, 381-431.

Friedman, M., and Schwartz, A.J.(1963), *A Monetary History of the United States, 1867-1960*, Princeton: Princeton University Press.

King, S.R.(1982), "Interest Rates and the Transmission of Monetary Fluctuations to Output and Prices," unpublished manuscript, Northwestern University.

Litterman, R.(1986), "Forecasting With Bayesian Vector Autoregressions—Five Years of Experience," *Journal of Business & Economic Statistics*, 4, 25-38.

Sims, C.A.(1980a), "Macroeconomics and Reality," *Econometrica*, 48, 1-48.

______________(1980b), "A Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered," *American Economic Review*, 70, 250-57.