Journal of Economic Perspectives—Volume 15, Number 4—Fall 2001—Pages 101–115

Vector Autoregressions

James H. Stock and Mark W. Watson

Macroeconometricians do four things: describe and summarize macroeconomic data, make macroeconomic forecasts, quantify what we do or do not know about the true structure of the macroeconomy, and advise (and sometimes become) macroeconomic policymakers. In the 1970s, these four tasks—data description, forecasting, structural inference and policy analysis—were performed using a variety of techniques. These ranged from large models with hundreds of equations to single-equation models that focused on interactions of a few variables to simple univariate time series models involving only a single variable. But after the macroeconomic chaos of the 1970s, none of these approaches appeared especially trustworthy.

Two decades ago, Christopher Sims (1980) provided a new macroeconometric framework that held great promise: vector autoregressions (VARs). A univariate autoregression is a single-equation, single-variable linear model in which the current value of a variable is explained by its own lagged values. A VAR is an n-equation, n-variable linear model in which each variable is in turn explained by its own lagged values, plus current and past values of the remaining $n - 1$ variables. This simple framework provides a systematic way to capture rich dynamics in multiple time series, and the statistical toolkit that came with VARs was easy to use and to interpret. As Sims (1980) and others argued in a series of influential early papers, VARs held out the promise of providing a coherent and credible approach to data description, forecasting, structural inference and policy analysis.

In this article, we assess how well VARs have addressed these four macroecono-

■ *James H. Stock is the Roy E. Larsen Professor of Political Economy, John F. Kennedy School of Government, Harvard University, Cambridge, Massachusetts. Mark W. Watson is Professor of Economics and Public Affairs, Department of Economics and Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, New Jersey. Both authors are Research Associates, National Bureau of Economic Research, Cambridge, Massachusetts.*

metric tasks.[1] Our answer is "it depends." In data description and forecasting, VARs have proven to be powerful and reliable tools that are now, rightly, in everyday use. Structural inference and policy analysis are, however, inherently more difficult because they require differentiating between correlation and causation; this is the "identification problem," in the jargon of econometrics. This problem cannot be solved by a purely statistical tool, even a powerful one like a VAR. Rather, economic theory or institutional knowledge is required to solve the identification (causation versus correlation) problem.

A Peek Inside the VAR Toolkit

What, precisely, is the effect of a 100-basis-point hike in the federal funds interest rate on the rate of inflation one year hence? How big an interest rate cut is needed to offset an expected half percentage point rise in the unemployment rate? How well does the Phillips curve predict inflation? What fraction of the variation in inflation in the past 40 years is due to monetary policy as opposed to external shocks?

Many macroeconomists like to think they know the answers to these and similar questions, perhaps with a modest range of uncertainty. In the next two sections, we take a quantitative look at these and related questions using several three-variable VARs estimated using quarterly U.S. data on the rate of price inflation (π_t), the unemployment rate (u_t) and the interest rate (R_t, specifically, the federal funds rate) from 1960:I–2000:IV.[2] First, we construct and examine these models as a way to display the VAR toolkit; criticisms are reserved for the next section.

VARs come in three varieties: reduced form, recursive and structural.

A *reduced form VAR* expresses each variable as a linear function of its own past values, the past values of all other variables being considered and a serially uncorrelated error term. Thus, in our example, the VAR involves three equations: current unemployment as a function of past values of unemployment, inflation and the interest rate; inflation as a function of past values of inflation, unemployment and the interest rate; and similarly for the interest rate equation. Each equation is estimated by ordinary least squares regression. The number of lagged values to include in each equation can be determined by a number of different methods, and we will use four lags in our examples.[3] The error terms in these regressions are the "surprise" movements in the variables after taking its past values into account. If the different variables are correlated with each other—as they typically are in

[1] Readers interested in more detail than provided in this brief tutorial should see Hamilton's (1994) textbook or Watson's (1994) survey article.

[2] The inflation data are computed as $\pi_t = 400ln(P_t/P_{t-1})$, where P_t is the chain-weighted GDP price index and u_t is the civilian unemployment rate. Quarterly data on u_t and R_t are formed by taking quarterly averages of their monthly values.

[3] Frequently, the Akaike (AIC) or Bayes (BIC) information criteria are used; for a discussion, see Lütkepohl (1993, chapter 4).

macroeconomic applications—then the error terms in the reduced form model will also be correlated across equations.

A *recursive VAR* constructs the error terms in each regression equation to be uncorrelated with the error in the preceding equations. This is done by judiciously including some contemporaneous values as regressors. Consider a three-variable VAR, ordered as 1) inflation, 2) the unemployment rate, and 3) the interest rate. In the first equation of the corresponding recursive VAR, inflation is the dependent variable, and the regressors are lagged values of all three variables. In the second equation, the unemployment rate is the dependent variable, and the regressors are lags of all three variables *plus* the current value of the inflation rate. The interest rate is the dependent variable in the third equation, and the regressors are lags of all three variables, the current value of the inflation rate *plus* the current value of the unemployment rate. Estimation of each equation by ordinary least squares produces residuals that are uncorrelated across equations.[4] Evidently, the results depend on the order of the variables: changing the order changes the VAR equations, coefficients, and residuals, and there are $n!$ recursive VARs representing all possible orderings.

A *structural VAR* uses economic theory to sort out the contemporaneous links among the variables (Bernanke, 1986; Blanchard and Watson, 1986; Sims, 1986). Structural VARs require "identifying assumptions" that allow correlations to be interpreted causally. These identifying assumptions can involve the entire VAR, so that all of the causal links in the model are spelled out, or just a single equation, so that only a specific causal link is identified. This produces instrumental variables that permit the contemporaneous links to be estimated using instrumental variables regression. The number of structural VARs is limited only by the inventiveness of the researcher.

In our three-variable example, we consider two related structural VARs. Each incorporates a different assumption that identifies the causal influence of monetary policy on unemployment, inflation and interest rates. The first relies on a version of the "Taylor rule," in which the Federal Reserve is modeled as setting the interest rate based on past rates of inflation and unemployment.[5] In this system, the Fed sets the federal funds rate R according to the rule

$$R_t = r^* + 1.5(\bar{\pi}_t - \pi^*) - 1.25(\bar{u}_t - u^*) + \textit{ lagged values of } R, \pi, u + \varepsilon_t,$$

where r^* is the desired real rate of interest, $\bar{\pi}_t$ and $\bar{u}_t$ are the average values of inflation and unemployment rate over the past four quarters, π^* and u^* are the target values of inflation and unemployment, and ε_t is the error in the equation. This relationship becomes the interest rate equation in the structural VAR.

[4] In the jargon of VARs, this algorithm for estimating the recursive VAR coefficients is equivalent to estimating the reduced form, then computing the Cholesky factorization of the reduced form VAR covariance matrix; see Lütkepohl (1993, chapter 2).

[5] Taylor's (1993) original rule used the output gap instead of the unemployment rate. Our version uses Okun's Law (with a coefficient of 2.5) to replace the output gap with unemployment rate.

The equation error, ε_t, can be thought of as a monetary policy "shock," since it represents the extent to which actual interest rates deviate from this Taylor rule. This shock can be estimated by a regression with $R_t - 1.5\ \bar{\pi}_t + 1.25\ \bar{u}_t$ as the dependent variable, and a constant and lags of interest rates, unemployment and inflation on the right-hand side.

The Taylor rule is "backward looking" in the sense that the Fed reacts to past information ($\bar{\pi}_t$ and $\bar{u}_t$ are averages of the past four quarters of inflation and unemployment), and several researchers have argued that Fed behavior is more appropriately described by forward-looking behavior. Because of this, we consider another variant of the model in which the Fed reacts to forecasts of inflation and unemployment four quarters in the future. This Taylor rule has the same form as the rule above, but with $\bar{\pi}_t$ and $\bar{u}_t$ replaced by four-quarter ahead forecasts computed from the reduced form VAR.

Putting the Three-Variable VAR Through Its Paces

The different versions of the inflation-unemployment-interest rate VAR are put through their paces by applying them to the four macroeconometric tasks. First, the reduced form VAR and a recursive VAR are used to summarize the comovements of these three series. Second, the reduced form VAR is used to forecast the variables, and its performance is assessed against some alternative benchmark models. Third, the two different structural VARs are used to estimate the effect of a policy-induced surprise move in the federal funds interest rate on future rates of inflation and unemployment. Finally, we discuss how the structural VAR could be used for policy analysis.

Data Description

Standard practice in VAR analysis is to report results from Granger-causality tests, impulse responses and forecast error variance decompositions. These statistics are computed automatically (or nearly so) by many econometrics packages (RATS, Eviews, TSP and others). Because of the complicated dynamics in the VAR, these statistics are more informative than are the estimated VAR regression coefficients or R^2 statistics, which typically go unreported.

Granger-causality statistics examine whether lagged values of one variable help to predict another variable. For example, if the unemployment rate does not help predict inflation, then the coefficients on the lags of unemployment will all be zero in the reduced-form inflation equation. Panel A of Table 1 summarizes the Granger-causality results for the three-variable VAR. It shows the p-values associated with the F-statistics for testing whether the relevant sets of coefficients are zero. The unemployment rate helps to predict inflation at the 5 percent significance level (the p-value is 0.02, or 2 percent), but the federal funds interest rate does not (the p-value is 0.27). Inflation does not help to predict the unemployment rate, but the federal funds rate does. Both inflation and the unemployment rates help predict the federal funds interest rate.

Table 1
VAR Descriptive Statistics for (π, u, R)

A. Granger-Causality Tests

Regressor	*Dependent Variable in Regression*		
	π	u	R
π	0.00	0.31	0.00
u	0.02	0.00	0.00
R	0.27	0.01	0.00

B. Variance Decompositions from the Recursive VAR Ordered as π, u, R

B.i. Variance Decomposition of π

Forecast Horizon	*Forecast Standard Error*	*Variance Decomposition (Percentage Points)*		
		π	u	R
1	0.96	100	0	0
4	1.34	88	10	2
8	1.75	82	17	1
12	1.97	82	16	2

B.ii. Variance Decomposition of u

Forecast Horizon	*Forecast Standard Error*	*Variance Decomposition (Percentage Points)*		
		π	u	R
1	0.23	1	99	0
4	0.64	0	98	2
8	0.79	7	82	11
12	0.92	16	66	18

B.iii. Variance Decomposition of R

Forecast Horizon	*Forecast Standard Error*	*Variance Decomposition (Percentage Points)*		
		π	u	R
1	0.85	2	19	79
4	1.84	9	50	41
8	2.44	12	60	28
12	2.63	16	59	25

Notes: π denotes the rate of price inflation, u denotes the unemployment rate and R denotes the Federal Funds interest rate. The entries in Panel A show the p-values for F-tests that lags of the variable in the row labeled *Regressor* do not enter the reduced form equation for the column variable labeled *Dependent Variable.* The results were computed from a VAR with four lags and a constant term over the 1960:I–2000:IV sample period.

Impulse responses trace out the response of current and future values of each of the variables to a one-unit increase in the current value of one of the VAR errors, assuming that this error returns to zero in subsequent periods and that all other errors are equal to zero. The implied thought experiment of changing one error while holding the others constant makes most sense when the errors are uncorrelated across equations, so impulse responses are typically calculated for recursive and structural VARs.

The impulse responses for the recursive VAR, ordered π_t, u_t, R_t, are plotted in Figure 1. The first row shows the effect of an unexpected 1 percentage point increase in inflation on all three variables, as it works through the recursive VAR system with the coefficients estimated from actual data. The second row shows the effect of an unexpected increase of 1 percentage point in the unemployment rate, and the third row shows the corresponding effect for the interest rate. Also plotted are ± 1 standard error bands, which yield an approximate 66 percent confidence interval for each of the impulse responses. These estimated impulse responses show patterns of persistent common variation. For example, an unexpected rise in inflation slowly fades away over 24 quarters and is associated with a persistent increase in unemployment and interest rates.

The *forecast error decomposition* is the percentage of the variance of the error made in forecasting a variable (say, inflation) due to a specific shock (say, the error term in the unemployment equation) at a given horizon (like two years). Thus, the forecast error decomposition is like a partial R^2 for the forecast error, by forecast horizon. These are shown in Panel B of Table 1 for the recursive VAR. They suggest considerable interaction among the variables. For example, at the 12-quarter horizon, 75 percent of the error in the forecast of the federal funds interest rate is attributed to the inflation and unemployment shocks in the recursive VAR.

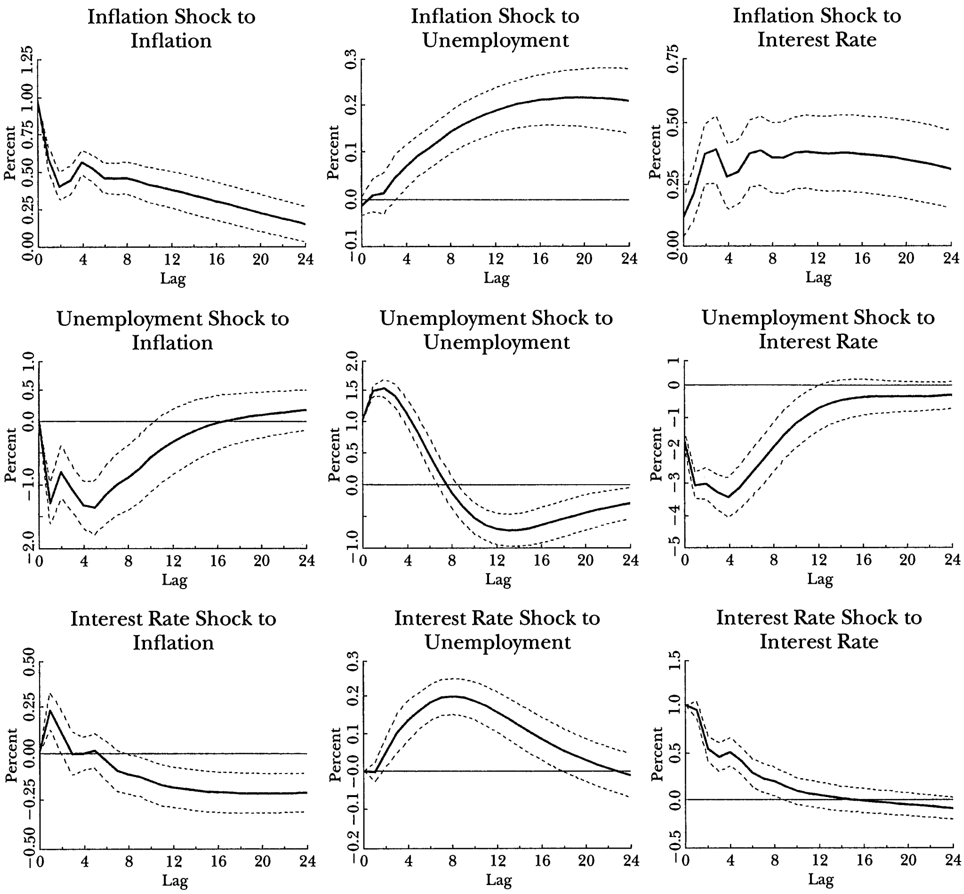
Forecasting

Multistep-ahead forecasts, computed by iterating forward the reduced form VAR, are assessed in Table 2. Because the ultimate test of a forecasting model is its out-of-sample performance, Table 2 focuses on pseudo out-of-sample forecasts over the period from 1985:I to 2000:IV. It examines forecast horizons of two quarters, four quarters and eight quarters. The forecast h steps ahead is computed by estimating the VAR through a given quarter, making the forecast h steps ahead, reestimating the VAR through the next quarter, making the next forecast and so on through the forecast period.[6]

As a comparison, pseudo out-of-sample forecasts were also computed for a univariate autoregression with four lags—that is, a regression of the variable on lags

[6] Forecasts like these are often referred to as pseudo or "simulated" out-of-sample forecasts to emphasize that they simulate how these forecasts would have been computed in real time, although, of course, this exercise is conducted retrospectively, not in real time. Our experiment deviates slightly from what would have been computed in real time because we use the current data, which includes later revisions made to the inflation and unemployment data by statistical agencies, rather than the data available in real time.

Figure 1
Impulse Responses in the Inflation-Unemployment-Interest Rate Recursive VAR



of its own past values—and for a random walk (or "no change") forecast. Inflation rate forecasts were made for the average value of inflation over the forecast period, while forecasts for the unemployment rate and interest rate were made for the final quarter of the forecast period. Table 2 shows the root mean square forecast error for each of the forecasting methods. (The mean squared forecast error is computed as the average squared value of the forecast error over the 1985–2000 out-of-sample period, and the resulting square root is the root mean squared forecast error reported in the table.) Table 2 indicates that the VAR either does no worse than or improves upon the univariate autoregression and that both improve upon the random walk forecast.

Structural Inference

What is the effect on the rates of inflation and unemployment of a surprise 100 basis-point increase in the federal funds interest rate? Translated into VAR jargon,

Table 2
Root Mean Squared Errors of Simulated Out-Of-Sample Forecasts, 1985:1–2000:IV

Forecast Horizon	Inflation Rate			Unemployment Rate			Interest Rate		
	RW	AR	VAR	RW	AR	VAR	RW	AR	VAR
2 quarters	0.82	0.70	0.68	0.34	0.28	0.29	0.79	0.77	0.68
4 quarters	0.73	0.65	0.63	0.62	0.52	0.53	1.36	1.25	1.07
8 quarters	0.75	0.75	0.75	1.12	0.95	0.78	2.18	1.92	1.70

Notes: Entries are the root mean squared error of forecasts computed recursively for univariate and vector autoregressions (each with four lags) and a random walk ("no change") model. Results for the random walk and univariate autoregressions are shown in columns labeled RW and AR, respectively. Each model was estimated using data from 1960:I through the beginning of the forecast period. Forecasts for the inflation rate are for the average value of inflation over the period. Forecasts for the unemployment rate and interest rate are for the final quarter of the forecast period.

this question becomes: What are the impulse responses of the rates of inflation and unemployment to the monetary policy shock in a structural VAR?

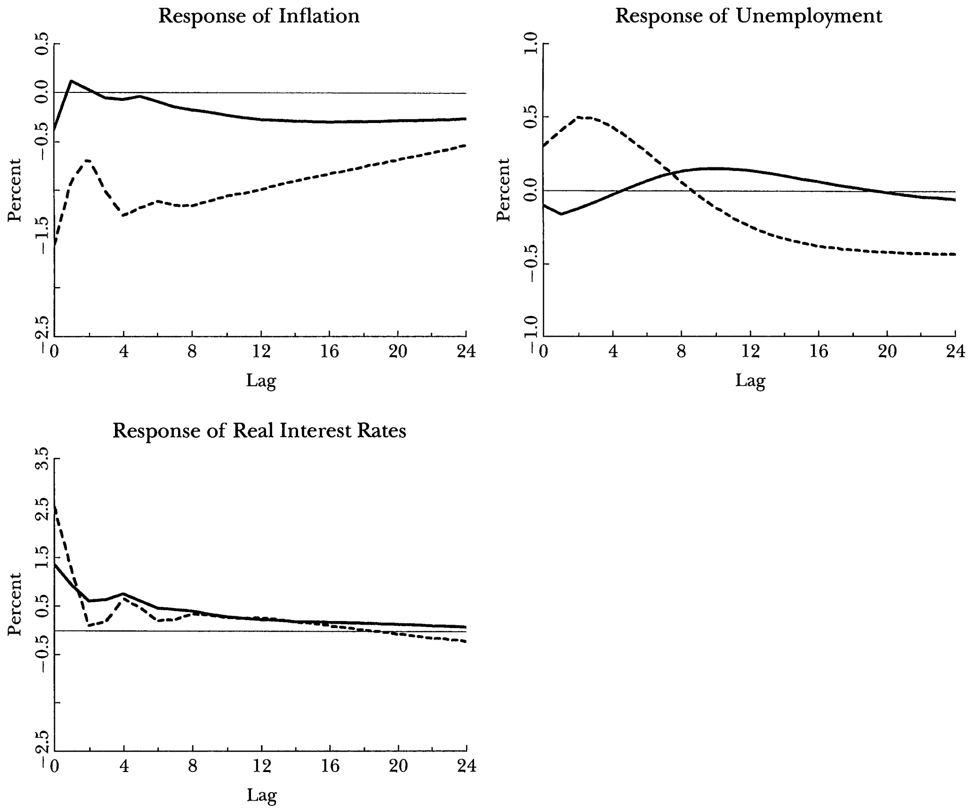
The solid line in Figure 2 plots the impulse responses computed from our model with the backward-looking Taylor rule. It shows the inflation, unemployment and real interest rate ($R_t - \pi_t$) responses to a 1 percentage point shock in the nominal federal funds rate. The initial rate hike results in the real interest rate exceeding 50 basis points for six quarters. Although inflation is eventually reduced by approximately 0.3 percentage points, the lags are long, and most of the action occurs in the third year after the contraction. Similarly, the rate of unemployment rises by approximately 0.2 percentage points, but most of the economic slowdown is in the third year after the rate hike.

How sensitive are these results to the specific identifying assumption used in this structural VAR—that the Fed follows the backward-looking Taylor rule? As it happens, very sensitive. The dashed line in Figure 2 plots the impulse responses computed from the structural VAR with the forward-looking Taylor rule. The impulse responses in real interest rates are broadly similar under either rule. However, in the forward-looking model the monetary shock produces a 0.5 percentage point increase in the unemployment rate within a year, and the rate of inflation drops sharply at first, fluctuates, then leaves a net decline of 0.5 percentage points after six years. Under the backward-looking rule, this 100 basis-point rate hike produces a mild economic slowdown and a modest decline in inflation several years hence; under the forward-looking rule, by this same action the Fed wins a major victory against inflation at the cost of a swift and sharp recession.

Policy Analysis

In principle, our small structural VAR can be used to analyze two types of policies: surprise monetary policy interventions and changing the policy rule, like shifting from a Taylor rule (with weight on both unemployment and inflation) to an explicit inflation targeting rule.

Figure 2
Impulse Responses of Monetary Policy Shocks for Different Taylor Rule Identifying Assumptions



Notes: The solid line is computed with the backward-looking Taylor rule; the dashed line, with the forward-looking Taylor rule.

If the intervention is an unexpected movement in the federal funds interest rate, then the estimated effect of this policy on future rates of inflation and unemployment is summarized by the impulse response functions plotted in Figure 2. This might seem a somewhat odd policy, but the same mechanics can be used to evaluate a more realistic intervention, such as raising the federal funds rate by 50 basis points and sustaining this increase for one year. This policy can be engineered in a VAR by using the right sequence of monetary policy innovations to hold the federal funds interest rate at this sustained level for four quarters, taking into account that in the VAR, actions on interest rates in earlier quarters affect those in later quarters (Sims, 1982; Waggoner and Zha, 1999).

Analysis of the second type of policy—a shift in the monetary rule itself—is more complicated. One way to evaluate a new policy rule candidate is to ask what would be the effect of monetary and nonmonetary shocks on the economy under the new rule. Since this question involves all the structural disturbances, answering it requires a complete macroeconomic model of the simultaneous determination of all the variables, and this means that all of the causal links in the structural VAR must be specified. In this case, policy analysis is carried out as follows: a structural VAR is estimated in which all the equations are identified, then a new model is formed by replacing the monetary policy rule. Comparing the impulse responses in the two models shows how the change in policy has altered the effects of monetary and nonmonetary shocks on the variables in the model.

How Well Do VARs Perform the Four Tasks?

We now turn to an assessment of VARs in performing the four macroeconometric tasks, highlighting both successes and shortcomings.

Data Description

Because VARs involve current and lagged values of multiple time series, they capture comovements that cannot be detected in univariate or bivariate models. Standard VAR summary statistics like Granger-causality tests, impulse response functions and variance decompositions are well-accepted and widely used methods for portraying these comovements. These summary statistics are useful because they provide targets for theoretical macroeconomic models. For example, a theoretical model that implied that interest rates should Granger-cause inflation but unemployment should not would be inconsistent with the evidence in Table 1.

Of course, the VAR methods outlined here have some limitations. One is that the standard methods of statistical inference (such as computing standard errors for impulse responses) may give misleading results if some of the variables are highly persistent.[7] Another limitation is that, without modification, standard VARs miss nonlinearities, conditional heteroskedasticity and drifts or breaks in parameters.

Forecasting

Small VARs like our three-variable system have become a benchmark against which new forecasting systems are judged. But while useful as a benchmark, small VARs of two or three variables are often unstable and thus poor predictors of the future (Stock and Watson, 1996).

State-of-the-art VAR forecasting systems contain more than three variables and allow for time-varying parameters to capture important drifts in coefficients (Sims, 1993). However, adding variables to the VAR creates complications, because the number of VAR parameters increases as the square of the number of variables: a nine-variable, four-lag VAR has 333 unknown coefficients (including the inter-

[7] Bootstrap methods provide some improvements (Kilian, 1999) for inference about impulse responses, but treatments of this problem that are fully satisfactory theoretically are elusive (Stock, 1997; Wright, 2000).

cepts). Unfortunately, macroeconomic time series data cannot provide reliable estimates of all these coefficients without further restrictions.

One way to control the number of parameters in large VAR models is to impose a common structure on the coefficients, for example using Bayesian methods, an approach pioneered by Litterman (1986) (six variables) and Sims (1993) (nine variables). These efforts have paid off, and these forecasting systems have solid real-time track records (McNees, 1990; Zarnowitz and Braun, 1993).

Structural Inference

In our three-variable VAR in the previous section, the estimated effects of a monetary policy shock on the rates of inflation and unemployment (summarized by the impulse responses in Figure 2) depend on the details of the presumed monetary policy rule followed by the Federal Reserve. Even modest changes in the assumed rule resulted in substantial changes in these impulse responses. In other words, the estimates of the structural impulse responses hinge on detailed institutional knowledge of how the Fed sets interest rates.[8]

Of course, the observation that results depend on assumptions is hardly new. The operative question is whether the assumptions made in VAR models are any more compelling than in other econometric models. This is a matter of heated debate and is thoughtfully discussed by Leeper, Sims and Zha (1996), Christiano, Eichenbaum and Evans (1999), Cochrane (1998), Rudebusch (1998) and Sims (1998). Below are three important criticisms of structural VAR modeling.[9]

First, what really makes up the VAR "shocks?" In large part, these shocks, like those in conventional regression, reflect factors omitted from the model. If these factors are correlated with the included variables, then the VAR estimates will contain omitted variable bias. For example, officials at the Federal Reserve might scoff at the idea that they mechanically followed a Taylor rule, or any other fixed-coefficient mechanical rule involving only a few variables; rather, they suggest that their decisions are based on a subtle analysis of very many macroeconomic factors, both quantitative and qualitative. These considerations, when omitted from the VAR, end up in the error term and (incorrectly) become part of the estimated historical "shock" used to estimate an impulse response. A concrete example of this in the VAR literature involves the "price puzzle." Early VARs showed an odd result: inflation tended to increase following monetary policy tightening. One explanation for this (Sims, 1992) was that the Fed was looking forward when it set interest rates and that simple VARs omitted variables that could be used to predict future inflation. When these omitted variables intimated an increase in inflation, the Fed tended to increase interest rates. Thus, these VAR interest rate shocks presaged

[8] In addition, the institutional knowledge embodied in our three-variable VAR is rather naïve; for example, the Taylor rule was designed to summarize policy in the Greenspan era, not the full sample in our paper.

[9] This list hits only the highlights; other issues include the problem of "weak instruments" discussed in Pagan and Robertson (1998) and the problem of noninvertible representations discussed in Hansen and Sargent (1991) and Lippi and Reichlin (1993).

increases in inflation. Because of omitted variables, the VAR mistakenly labeled these increases in interest rates as monetary shocks, which led to biased impulse responses. Indeed, Sims's explanation of the price puzzle has led to the practice of including commodity prices in VARs to attempt to control for predicted future inflation.

Second, policy rules change over time, and formal statistical tests reveal widespread instability in low-dimensional VARs (Stock and Watson, 1996). Constant parameter structural VARs that miss this instability are improperly identified. For example, several researchers have documented instability in monetary policy rules (for example, Bernanke and Blinder, 1992; Bernanke and Mihov, 1998; Clarida, Gali and Gertler, 2000; Boivin, 2000), and this suggests misspecification in constant coefficient VAR models (like our three-variable example) that are estimated over long sample periods.

Third, the timing conventions in VARs do not necessarily reflect real-time data availability, and this undercuts the common method of identifying restrictions based on timing assumptions. For example, a common assumption made in structural VARs is that variables like output and inflation are sticky and do not respond "within the period" to monetary policy shocks. This seems plausible over the period of a single day, but becomes less plausible over a month or quarter.

In this discussion, we have carefully distinguished between recursive and structural VARs: recursive VARs use an arbitrary mechanical method to model contemporaneous correlation in the variables, while structural VARs use economic theory to associate these correlations with causal relationships. Unfortunately, in the empirical literature the distinction is often murky. It is tempting to develop economic "theories" that, conveniently, lead to a particular recursive ordering of the variables, so that their "structural" VAR simplifies to a recursive VAR, a structure called a "Wold causal chain." We think researchers yield to this temptation far too often. Such cobbled-together theories, even if superficially plausible, often fall apart on deeper inspection. Rarely does it add value to repackage a recursive VAR and sell it as structural.

Despite these criticisms, we think it is possible to have credible identifying assumptions in a VAR. One approach is to exploit detailed institutional knowledge. An example of this is the study by Blanchard and Perotti (1999) of the macroeconomic effects of fiscal policy. They argue that the tax code and spending rules impose tight constraints on the way that taxes and spending vary within the quarter, and they use these constraints to identify the exogenous changes in taxes and spending necessary for causal analysis. Another example is Bernanke and Mihov (1998), who use a model of the reserves market to identify monetary policy shocks. A different approach to identification is to use long-run restrictions to identify shocks; for example, King, Plosser, Stock and Watson (1991) use the long-run neutrality of money to identify monetary shocks. However, assumptions based on the infinite future raise questions of their own (Faust and Leeper, 1997).

A constructive approach is to recognize explicitly the uncertainty in the assumptions that underlie structural VAR analysis and see what inferences, or range of inferences, still can be made. For example, Faust (1998) and Uhlig (1999) discuss inference methods that can be applied using only inequality restrictions on the theoretical impulse responses (for example, monetary contractions do not cause booms).

Policy Analysis

Two types of policies can be analyzed using a VAR: one-off innovations, in which the same rule is maintained; and changes in the policy rule. The estimated effect of one-off innovations is a function of the impulse responses to a policy innovation, and potential pitfalls associated with these have already been discussed.

Things are more difficult if one wants to estimate the effect of changing policy rules. If the true structural equations involve expectations (say, an expectational Phillips curve), then the expectations will depend on the policy rule; thus, in general, all the VAR coefficients will depend on the rule. This is just a version of the Lucas (1976) critique. The practical importance of the Lucas critique for this type of VAR policy analysis is a matter of debate.

After Twenty Years of VARs

VARs are powerful tools for describing data and for generating reliable multivariate benchmark forecasts. Technical work remains, most notably extending VARs to higher dimensions and richer nonlinear structures. Even without these important extensions, however, VARs have made lasting contributions to the macroeconometrician's toolkit for tackling these two tasks.

Whether 20 years of VARs have produced lasting contributions to structural inference and policy analysis is more debatable. Structural VARs can capture rich dynamic properties of multiple time series, but their structural implications are only as sound as their identification schemes. While there are some examples of thoughtful treatments of identification in VARs, far too often in the VAR literature the central issue of identification is handled by ignoring it. In some fields of economics, such as labor economics and public finance, identification can be obtained credibly using natural experiments that permit some exogenous variation to be teased out of a relationship otherwise fraught with endogeneity and omitted variables bias. Unfortunately, these kinds of natural experiments are rare in macroeconomics.

Although VARs have limitations when it comes to structural inference and policy analysis, so do the alternatives. Calibrated dynamic stochastic general equilibrium macroeconomic models are explicit about causal links and expectations and provide an intellectually coherent framework for policy analysis. But the current generation of these models do not fit the data well. At the other extreme, simple single-equation models, for example, regressions of inflation against lagged interest rates, are easy to estimate and sometimes can produce good forecasts. But if it is difficult to distinguish correlation and causality in a VAR, it is even more so in single-equation models, which can, in any event, be viewed as one equation pulled from a larger VAR. Used wisely and based on economic reasoning and institutional detail, VARs both can fit the data and, at their best, can provide sensible estimates of some causal connections. Developing and melding good theory and institutional detail with flexible statistical methods like VARs should keep macroeconomists busy well into the new century.

■ *We thank Jean Boivin, Olivier Blanchard, John Cochrane, Charles Evans, Ken Kuttner, Eric Leeper, Glenn Rudebusch, Chris Sims, John Taylor, Tao Zha and the editors for useful suggestions. This research was funded by NSF grant SBR-9730489.*

References

Bernanke, Ben S. 1986. "Alternative Explanations of the Money-Income Correlation." *Carnegie-Rochester Conference Series on Public Policy.* Autumn, 25, pp. 49–99.

Bernanke, Ben S. and Alan Blinder. 1992. "The Federal Funds Rate and the Channels of Monetary Transmission." *American Economic Review.* September, 82:4, pp. 901–21.

Bernanke, Ben S. and Ilian Mihov. 1998. "Measuring Monetary Policy." *Quarterly Journal of Economics.* August, 113:3, pp. 869–902.

Blanchard, Olivier J. and Roberto Perotti. 1999. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." NBER Working Paper No. 2769, July.

Blanchard, Olivier J. and Mark W. Watson. 1986. "Are Business Cycles All Alike?" in *The American Business Cycle: Continuity and Change.* R.J. Gordon, ed. Chicago: University of Chicago Press, pp. 123–56.

Boivin, Jean. 2000. "The Fed's Conduct of Monetary Policy: Has it Changed and Does it Matter?" Manuscript, Columbia University, December.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans. 1997. "Sticky Price and Limited Participation Models: A Comparison." *European Economic Review.* June, 41:6, pp. 1201–49.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans. 1999. "Monetary Policy Shocks: What Have We Learned and To What End?" in *Handbook of Macroeconomics, Volume 1A.* John B. Taylor and Michael Woodford, eds. Amsterdam: Elsevier Science Ltd., pp. 65–148.

Clarida, Richard, Jordi Gali and Mark Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature.* December, 37:4, pp. 1661–707.

Clarida, Richard, Jordi Gali and Mark Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics.* February, 115:1, pp. 147–80.

Cochrane, John H. 1998. "What Do the VARs Mean?: Measuring the Output Effects of Monetary Policy." *Journal of Monetary Economics.* 41:2, pp. 277–300.

Faust, Jon. 1998. "The Robustness of Identified VAR Conclusions About Money." *Carnegie-Rochester Conference Series on Public Policy.* December, 49, pp. 207–44.

Faust, Jon and Eric M. Leeper. 1997. "When Do Long-Run Identifying Restrictions Give Reliable Results?" *Journal of Business and Economic Statistics.* July, 15:3, pp. 345–53.

Granger, Clive W.J. and Paul Newbold. 1977. *Forecasting Economic Time Series, First Edition.* New York: Academic Press.

Hamilton, James D. 1994. *Time Series Analysis.* Princeton, N.J.: Princeton University Press.

Hansen, Lars P. and Thomas J. Sargent. 1991. "Two Problems in Interpreting Vector Autoregressions," in *Rational Expectations Econometrics.* Lars P. Hansen and Thomas J. Sargent, eds. Boulder: Westview, pp. 77–119.

Kilian, Lutz. 1999. "Finite-Sample Properties of Percentile and Percentile-*t* Bootstrap Confidence Intervals for Impulse Responses." *Review of Economics and Statistics.* November, 81:4, pp. 652–60.

King, Robert G. et al. 1991. "Stochastic Trends and Economic Fluctuations." *American Economic Review.* 81:4, pp. 819–40.

Leeper, Eric M., Christopher A. Sims and Tao Zha. 1996. "What Does Monetary Policy Do?" *Brookings Papers on Economic Activity.* 2, pp. 1–63.

Lippi, Marco and Lucrezia Reichlin. 1993. "The Dynamic Effects of Supply and Demand Disturbances: Comment." *American Economic Review.* June, 83:3, pp. 644–52.

Litterman, Robert B. 1986. "Forecasting With Bayesian Vector Autoregressions: Five Years of Experience." *Journal of Business and Economic Statistics.* January, 4:1, pp. 25–38.

Lucas, Robert E., Jr. 1976. "Economic Policy Evaluation: A Critique." *Journal of Monetary Economics.* 1:2, pp. 19–46.

Lütkepohl, Helmut. 1993. *Introduction to Multiple Time Series Analysis, Second Edition.* Berlin: Springer-Verlag.

McNees, Stephen K. 1990. "The Role of Judgment in Macroeconomic Forecasting Accuracy." *International Journal of Forecasting.* October, 6:3, pp. 287–99.

Pagan, Adrian R. and John C. Robertson. 1998. "Structural Models of the Liquidity Effect." *Review of Economics and Statistics.* May, 80:2, pp. 202–17.

Rudebusch, Glenn D. 1998. "Do Measures of Monetary Policy in a VAR Make Sense?" *International Economic Review.* November, 39:4, pp. 907–31.

Sims, Christopher A. 1980. "Macroeconomics and Reality." *Econometrica.* January, 48:1, pp. 1–48.

Sims, Christopher A. 1982. "Policy Analysis With Econometric Models." *Brookings Papers on Economic Activity.* 1, pp. 107–52.

Sims, Christopher A. 1986. "Are Forecasting Models Usable for Policy Analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review.* Winter, 10:1, pp. 2–16.

Sims, Christopher A. 1992, "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy." *European Economic Review.* June, 36:5, pp. 975–1011.

Sims, Christopher A. 1993. "A Nine Variable Probabilistic Macroeconomic Forecasting Model," in *NBER Studies in Business: Business Cycles, Indicators, and Forecasting, Volume 28.* James H. Stock and Mark W. Watson, eds. Chicago: University of Chicago Press, pp. 179–214.

Sims, Christopher A. 1998. "Comment on Glenn Rudebusch's 'Do Measures of Monetary Policy in a VAR Make Sense?' (with reply)." *International Economic Review.* November, 39:4, pp. 933–48.

Sims, Christopher A. and Tao Zha. 1995. "Does Monetary Policy Generate Recessions?" Manuscript, Federal Reserve Bank of Atlanta.

Stock, James H. 1997. "Cointegration, Long-Run Comovements, and Long-Horizon Forecasting," in *Advances in Econometrics: Proceedings of the Seventh World Congress of the Econometric Society, Volume III.* David Kreps and Kenneth F. Wallis, eds. Cambridge: Cambridge University Press, pp. 34–60.

Stock, James H. and Mark W. Watson. 1996. "Evidence on Structural Instability in Macroeconomic Time Series Relations." *Journal of Business and Economic Statistics.* January, 14:1, pp. 11–30.

Taylor, John B. 1993. "Discretion Versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy.* December, 39, pp. 195–214.

Uhlig, Harald. 1999. "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." Manuscript, CentER, Tilburg University.

Waggoner, Daniel F. and Tao Zha. 1999. "Conditional Forecasts in Dynamic Multivariate Models." *Review of Economics and Statistics.* November, 81:4, pp. 639–51.

Watson, Mark W. 1994. "Vector Autoregressions and Cointegration," in *Handbook of Econometrics, Volume IV.* Robert Engle and Daniel McFadden, eds. Amsterdam: Elsevier Science Ltd., pp. 2844–915.

Wright, Jonathan H. 2000. "Confidence Intervals for Univariate Impulse Responses with a Near Unit Root." *Journal of Business and Economic Statistics.* July, 18:3, pp. 368–73.

Zarnowitz, Victor and Phillip Braun. 1993. "Twenty-Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance," in *NBER Studies in Business Cycles: Business Cycles, Indicators, and Forecasting, Volume 28.* James H. Stock and Mark W. Watson, eds. Chicago: University of Chicago Press, pp. 11–94.