Comment on Glenn Rudebusch's "Do Measures of Monetary Policy in a Var Make Sense?"

Author(s): Christopher A. Sims

Source: *International Economic Review*, Nov., 1998, Vol. 39, No. 4, Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance (Nov., 1998), pp. 933-941

Published by: Wiley for the Economics Department of the University of Pennsylvania and Institute of Social and Economic Research, Osaka University

Stable URL: https://www.jstor.org/stable/2527345

INTERNATIONAL ECONOMIC REVIEW
Vol. 39, No. 4, November 1998

COMMENT ON GLENN RUDEBUSCH'S "DO MEASURES OF MONETARY POLICY IN A VAR MAKE SENSE?"*

BY CHRISTOPHER A. SIMS†

Yale University, U.S.A.

GENERALITIES

The identified VAR literature has in the last few years begun producing consistent, somewhat controversial results. The results challenge conventional wisdom about the nature of monetary policy and its effects, and the methods rest on careful attention to simultaneity problems. The literature accordingly has begun attracting vigorous criticism, both from economists comforable with one version or another of conventional wisdom and from economists comfortable with the currently fashionable view that macroeconomics, properly executed, never requires thinking about more than one regression equation at a time.

Before addressing the criticisms put forth in Glenn Rudebusch's paper, let me describe broadly what the literature has discovered.

(1) Most variation in monetary policy instruments is accounted for by responses of policy to the state of the economy, not by random disturbances to policy behavior.
(2) Responses of real variables to monetary policy shifts are estimated as modest or nil, depending on the specification.
(3) Monetary policy has historically increased interest rates in response to nonpolicy shocks that create inflationary pressure by more than it would have under a policy of fixing the monetary stock.
(4) A reasonable picture of the effects of monetary policy shifts emerges only under identifying assumptions of delay in the reaction of certain 'sluggish' private sector variables to monetary policy shifts.

Most papers in the literature have not emphasized these conclusions, probably because it has until recently been regarded as a substantial accomplishment just to display a model that is consistent with multivariate time series data and implies reasonable effects for monetary policy shifts. The small real effects have particularly not been emphasized, since it is apparently sometimes thought that finding substantial real effects for monetary policy changes helps make identified VAR analyses credible. That these conclusions are nonetheless borne out in the literature is argued in somewhat more detail in my paper with Tao Zha, "Does Monetary Policy

* Manuscript received January 1998.
† E-mail: *sims@econ.yale.edu*

Generate Recessions?" available at ftp://ftp/econ.yale.edu/pub/sims/mpolicy. They are confirmed for a number of European countries and Japan in work by Soyoung Kim, (1995).

There is a view, which Milton Friedman used to restate regularly some years ago, that erratic variation in monetary policy is the primary source of business cycle fluctuations, with each postwar U.S. business cycle largely explainable via the pattern of monetary policy variations preceding it. Friedman used to defend this view via statistical analysis that took the time path of a monetary aggregate as a sufficient statistic for the time path of monetary policy. The recent VAR literature decisively undercuts this way at looking at history, and as far as I know there are no attacks on the VAR literature, including this one, that explicitly put forth a quantitative model that contradicts the VAR literature on this score.

SPECIFICS

Many of the paper's complaints are what I would call 'quibbles.' They critize modeling choices about aggregation, time aggregation, time variation, and linearity that are an essential part of any quantitative macroeconomic modeling. Of course there is some general distaste for quantitative macroeconomic modeling these days, precisely because such choices are necessary.[1] But to me at least, these criticisms, though in some cases they deserve discussion, are not very interesting, so I postpone them to the latter part of these notes.

There is one component of Rudebusch's scattershot attack on identified VAR modeling that raises legitimate and interesting research questions.

Let's Look at the Policy Shock Time Series. In Section 4.3, Rudebusch makes the most interesting point in his paper. His own VAR-style policy equations has residuals that have a correlation of only 0.12 with those of another VAR study by Bernanke and Mihov (1995). This result may not be representative of the literature as a whole, but it deserves wider examination. Zha and I, for example, find that the monetary policy shocks for our two specifications (one using the Treasury Bill Rate and M2, the other replacing these variables with the Federal Funds Rate and Total Reserves) have a correlation of only 0.29. Examination of the policy shock time series in our two specifications show that they disagree substantially on when the biggest policy shifts occurred. I expect that when this issue is more widely examined it will be confirmed that despite their broad agreement on the nature of the response of the economy to monetary policy shocks, identified VAR studies disagree rather substantially about the history of policy disturbances.

Further study of these disagreements about policy shock time series will be important in resolving the disagreements about the size and existence of real effects

[1] In this version of the paper, the conclusions adopt a constructive tone, pointing to some recent variants of the VAR modeling approach that incorporate nonstochastic breaks in the specification, or 'regime shifts,' as promising. However, I still think it fair to summarize the paper's overall tone as implying deep skepticism of quantitative macroeconomic policy modeling in general, since it criticizes VARs for generic problems of quantitative modeling.

of monetary policy shocks. As the paper points out, for some purposes this disagreement is not important. The conclusions listed at the beginning of this comment are all robust to this issue, for example, though the paper is right to observe that while this issue is unresolved one must be skeptical about the models' time sequences of monetary policy tightness and looseness. However, it is important to understand the implications of disagreement about the policy shock time series, since even in this version of the paper it is mistaken in some respects on this score.

How Models Can Disagree on Policy Shocks, While Agreeing on their Effects. Multivariate models can be hard to grasp intuitively. An appealing but misleading way to think about identified VARs is that they 'identify' a policy shock time series, then find the effects of policy essentially by regressing everything else on the policy shock time series. This would imply that if two identified VARs that find very different policy shock time series agree on the effects of policy shocks on other variables, this is only an accident (or the result of data mining). This interpretation of VARs did apply to some of the earliest, rational expectations monetarist, approaches to estimating the effects of monetary policy. Those models implied that the innovation in a monetary aggregate was determined entirely by policy behavior. Some researchers, Robert Barro for instance, estimated models that literally constructed money innovations and regressed other variables on the money innovations to find the effects of policy.

But the modern identified VAR literature is not doing this. It works with multivariate models that can be understood only in a simultaneous equations framework. Suppose we were dealing with the simplest supply and demand simultaneous equations model. Everyone understands that if both equations in such a model are to be identified, then each equation must have a 'shifter' variable that does not appear in the other equation—for example, weather as a supply shifter, and income as a demand shifter. The effect of a shock to supply (a change in its disturbance term, which moves the supply curve up and down or left and right, depending on the normalization) is a movement along the demand curve. It is accurately estimated if we have an accurate estimate of the demand curve. This requires that there be substantial variation in the weather, the supply shifter. On the other hand, the accuracy of the estimates of the supply equation shocks depends on the accuracy of the estimates of the coefficients in the supply equation. These depend on the amount of variation in income, the demand shifter. Thus in this simple example it is clearly not true that estimates of the responses to supply shocks are 'only as good as' the estimates of the supply shocks themselves. In fact, as the amount of variation in weather grows, while the amount of variation in income shrinks, the accuracy of the estimated supply shocks can get arbitrarily bad while the accuracy of the estimate of the response to supply shocks becomes arbitrarily accurate.

Tao Zha and I have checked our own model to see if this is what is happening. We find that this simple story is not adequate. We construct the posterior distribution of the correlation of the true shock series with our point estimate of the shock series. The distribution shows almost no probability on correlations below 0.8, and most probability concentrated on correlations above 0.9. Plots of typical random draws from the posterior distribution of shock time series show close agreement around the dates of the largest shocks. It is interesting to note, though, that the 1988–1994 period that covers most of Rudebusch's data set is a period of small shocks, and within this period correlations are substantially weaker.

Our conclusion is, then, that the 0.29 correlation between the policy shocks of our two specifications is much lower than can be accounted for by each model's own uncertainty about the policy shock series. It comes from a 'statistically significant' difference in specifications.

Could it be then that two models have different specifications, implying quite different policy shock time series, yet both accurately estimate the same response to policy shocks? Consider again our simple supply and demand model. Suppose there are two supply shifters, weather and insect density. Suppose one model includes the weather variable, but omits, and thus relegates to the error term, insect density. The other model does the reverse. So long as both supply shifters are legitimate exogenous variables, uncorrelated with the disturbance term in the demand equation, both models can lead to accurate estimates of the demand equation, because each offers one legitimate instrumental variable for that equation. But of course, since each model includes the other's supply shifter in the 'supply shock,' there is no limit to how different their estimated supply shock time series might appear.

Tao Zha and I have made the point that there is a common characteristic across all the apparently various VAR studies that have found responses to monetary policy that appear reasonable, avoiding the price puzzle. Each of these studies, in one way or another, assumes that there are one or more 'sluggish' private sector variables, usually including GDP, industrial production, or investment, that do not respond within the time unit to monetary policy variables. This variable or group of variables therefore qualifies as instruments for the money supply sector of the model, while lagged monetary policy variables—monetary aggregates and interest rates—become legitimate instruments for the nonpolicy equations describing the determination of the sluggish variables. Just as in the simple supply and demand model, the agreement among models in the exclusion of contemporaneous values of monetary policy instruments from the equations determining certain key private sector aggregates leads to agreement on the dynamics by which those variables are determined. The models disagree much more on which variables they include in the policy reaction function, and therefore also disagree more on the time series of policy shocks. Several of the models in the literature, for example, have a block recursive structure that entails that the group of sluggish variables coincide with the group of variables that enter contemporaneously in the reaction function. On the other hand Tao Zha and I, and Soyoung Kim, exclude contemporaneous GDP and GDP deflator from the reaction function, while still treating these variables as sluggish.

Note that on this interpretation, the conclusion that the contribution of policy shocks to business cycle variation is modest remains robust, contrary to the claim in this paper by Rudebusch. If different studies are allocating different components of the endogenous response of policy to the error term, then they are all exaggerating the effects of policy shocks on other variables. A conclusion that policy shock effects are small will only be strengthened by modeling improvements that explain more of policy as systematic.

Federal Funds Futures. The paper demonstrates at some length that forecasts of the federal funds rate implicit in futures market prices are more accurate, and little correlated with, those obtained from a reduced form prediction equation for the federal funds rate similar to those in some VAR's. This point is of some interest, but it does not have the implications the paper claims for it.

It is a main point of the literature the paper criticizes that there is no reason in principle to assume that unforecastable changes in the federal funds rate are policy shocks. The literature uses a variety of identification schemes, and as the paper points out, comes up with widely varying estimates of the history of policy disturbances. The 'VAR policy shocks' that Rudebusch constructs for this paper are, as he points out, nearly identical with one-step-ahead prediction errors for the federal funds rate from a VAR. It is clear then that they are not nearly identical to the policy shocks in some other identified VAR's. But even if they were, the fact that futures market predictions produce better forecasts than the VAR variable list the paper considers does not imply that the futures market prediction errors are good measures of policy shifts. The Federal Reserve can and does respond systematically within days or hours to current information on the private economy. Forecast errors for policy instruments therefore inevitably confound forecast errors for private sector variables that influence policy with forecast errors for policy itself. So if what futures markets can predict is policy, but not the private sector variables that systematically influence policy, then the futures market forecast errors, though smaller, would be clearly worse as measures of policy shocks than the VAR errors.

In other words, much of the paper's discussion of the results for federal funds future starts from a false premise. That doesn't prevent it from being of some interest. Any variable that can substantially improve reduced form VAR forecasts is an interesting candidate for inclusion in an identified VAR analysis. Note, though, that if a model is being used to interpret behavior, not just for forecasting, it is not generally a good idea to include every variable that improves reduced form forecasts. There are a number of possible choices for a short interest rate, for example, and the models have generally chosen a single short rate, though the choice has ranged over the federal funds rate, commercial paper rates, treasury bill rates, and (usually in other countries) call rates. Short interest rates do not move perfectly together, but they are tightly related. Few people would be surprised to find out that over short horizons treasury bill rates help predict commercial paper rates or vice versa, yet this does not lead to many structural models containing both.

The reason for this is that in some sense, for structural modeling purposes, short interest rates are all measures of a single concept. If one includes multiple short rates while still hoping to obtain behavioral interpretations of the results, it is necessary to carefully model the frictions, institutional factors, and so forth, that lead to discrepancy among the rates, so as to distinguish the movement in the underlying short-rate concept from the factors that distinguish the rates. Commonly such detailed modeling appears not to offer insight or accuracy commensurate with the effort it requires, so modelers settle on a single measure of short rates.

Federal funds rates futures are not short rates, but it turns out that the extra forecasting power they provide is a subset of that available by adding additional short rates to a VAR. Rudebusch kindly provided me with the data set he used for the paper, extended to include also futures prices from the middle of the month preceding the contract month as well as the end-of-month data used in the paper itself. It is important to keep track of the precise dating, here, because results are strongly affected by it. The series Rudebusch uses has a half-month timing advantage over the monthly average data used in most VAR's. It can therefore appear spuriously to have extra information just because of its timing advantage. Table 1 shows results from a regression of the federal funds rate (FFR) on lags of itself, of the treasury bill rate (TB3MS), and of the discount rate (DISCRT), as well as on the futures price prediction as of 15 days before the start of the current month (FFF15).

Clearly there is no forecasting power in the futures market forecasts that is not already available in lagged funds rate data, lagged treasury bill rate data, and lagged discount rate data. I found similar results with other combinations of interest rates, and find puzzling the claim in the paper that results are essentially unchanged when the 15-day lagged FFF is substituted for the unlagged FFF.

The same regression run with the futures market forecasts from the end of the preceding month, instead of the middle of the preceding month, makes the futures market variable very significant, confirming that time aggregation issues are critical to interpreting results here. However even the end-of-preceding-month federal funds future variable leaves substantial predictive power for the lagged interest rates. Since this casts doubt on Rudebusch's strong claims of market efficiency, I display the results in Table 2, though they are not very relevant to the main points of this comment.

This is not to say that in modeling the behavior of the Federal Reserve, which is certainly aware of Tbill rates and the discount rate without delay, we should not use

TABLE 1

Dependent Variable FFR: Estimation by Least Squares
Monthly Data From 88:11 To 95:03

Usable Observations	77	Degrees of Freedom	69
Centered R**2	0.996339	R Bar**2	0.995967
Uncentered R**2	0.999481	$T \times R$**2	76.960
Mean of Dependent Variable		5.7343662338	
Standard Error of Dependent Variable		2.3468022835	
Standard Error of Estimate		0.1490275966	
Sum of Squared Residuals		1.5324364930	
Regression $F(7,69)$		2682.5155	
Significance Level of F		0.00000000	
Durbin–Watson Statistic		2.033023	
$Q(19-0)$		15.789015	
Significance Level of Q		0.67130548	

Variable	Coeff.	Std. Error	T-Stat.	Signif.
1. Constant	−0.005249077	0.071294538	−0.07363	0.94152162
2. FFF15{1}	0.097391950	0.176886917	0.55059	0.58369385
3. FFR{1}	0.686652699	0.180151603	3.81153	0.00029683
4. FFR{2}	−0.038487564	0.137030961	−0.28087	0.77965240
5. TB3MS{1}	0.756340771	0.192798902	3.92295	0.00020403
6. TB3MS{2}	−0.252331336	0.155386818	−1.62389	0.10985900
7. DISCRT{1}	−0.406015930	0.151887938	−2.67313	0.00937140
8. DISCRT{2}	0.149247075	0.144783362	1.03083	0.30622024

TABLE 2

Dependent Variable FFR: Estimation by Least Squares	
Monthly Data From 88:11 To 95:03	
Usable Observations 77	Degrees of Freedom 69
Centered $R^{**}2$ 0.996785	R Bar**2 0.996459
Uncentered $R^{**}2$ 0.999544	$T \times R^{**}2$ 76.965
Mean of Dependent Variable	5.7343662338
Standard Error of Dependent Variable	2.3468022835
Standard Error of Estimate	0.1396545143
Sum of Squared Residuals	1.3457334526
Regression $F(7,69)$	3056.0472
Significance Level of F	0.00000000
Durbin–Watson Statistic	1.990067
$Q(19-0)$	15.272533
Significance Level of Q	0.70512557

Variable	Coeff.	Std. Error	T-Stat.	Signif.
1. Constant	−0.003869406	0.066606552	−0.05809	0.95384209
2. FFF1{1}	0.456011245	0.144797876	3.14930	0.00241910
3. FFR{1}	0.453104681	0.171897568	2.63590	0.01035617
4. FFR{2}	−0.069977581	0.127293211	−0.54974	0.58427593
5. TB3MS{1}	0.406371377	0.178547413	2.27599	0.02595467
6. TB3MS{2}	−0.078785963	0.153419346	−0.51353	0.60921877
7. DISCRT{1}	−0.325033483	0.141919120	−2.29027	0.02506690
8. DISCRT{2}	0.151405023	0.135602940	1.11653	0.26806756

data on them. But clearly using them raises fresh modeling problems. The Federal Reserve may respond to market forces that push Tbill rates away from the funds rate, but the Tbill market may also anticipate Federal policy moves. No single-equation approach will unravel this.

QUIBBLES

There is a more complete version of these comments available on the Internet at http://www.econ.yale.edu/~sims. Due to space limitations, this version just briefly summarizes my views on the topics below.

Time-Invariant Linear Structure. No doubt the structure of the economy and the Federal Reserve's reaction function are nonlinear and time varying. I have produced some evidence of this myself (Sims 1993). This does not mean we must necessarily use models that allow for time variation and nonlinearity in order to get accurate results. Indeed, allowing for time variation and nonlinearity of the wrong kinds can be misleading or can lead to models that are of little use. Particularly in modeling policy, assuming that there are nonstochastic 'break points' at which exogenous shifts in behavior occur is almost surely a worse distortion of reality than assuming constancy. The best evidence is that nonlinearity and time variation are of modest quantitative significance.

Information Sets. There are a lot of variables that might influence policy behavior that do not appear in at least any one VAR model. Many of the variables

Rudebusch lists have been tried in such models, and have not proved to be of major importance.

Use of Final Revised Data. Since my paper (Sims, 1986) that unwittingly resolved the 'price puzzle' before its time, I have always used specifications of policy behavior that assume that the authorities react immediately to the variables they can observe without delay (commodity prices, monetary aggregates, and financial variables), and only with a delay to variables that they can observe only with delay, such as GDP and the GDP deflator. I think it is true that specifications pretending that authorities can react instantly to such data can lead to important distortions. I am therefore sympathetic to this complaint of Rudebusch's. On the other hand, it is not clear that this criticism is quantitatively important.

Long Distributed Lags. Rudebusch takes the fact that VAR modelers find significant coefficients on variables with long lags to imply that 'the Fed reacts systematically to old information.' This is just a mistake. The length of lags in a behavioral relation has nothing to do with whether it reflects reaction to 'old information.'

CONCLUSION: CONSTRUCTIVE CRITICISM

The Rudebusch paper does not offer much constructive criticism. It is more of an attack than a critique. So I close by offering constructive criticism of it, by trying to point it in the direction of leaving aside the more obvious and weaker points it has embraced and expanding its treatment of the points that are more penetrating. I also offer a briefly summary of my own views of where the identified VAR literature is deficient and where it should be heading.

The issues of time invariance, linearity, and variable selection are universal in macroeconomic modeling. VAR modelers have inevitably thought about them already, as have their critics. To make criticism along these lines telling, the paper would have to show that a particular way of introducing nonlinearity, allowing for time variation, or changing the usual variable set, not yet considered in the literature, results in sharp changes in results. The paper does not begin this task, so its criticisms on these lines do not bite.

That financial market variables, and in particular interest rates other than the funds rate, have predictive value not captured in the usual list of VAR variables is a point worth making. It does not prove that existing VAR studies are worthless, but it suggests that there might be returns to more detailed modeling of monetary policy and money supply. The one-short-interest-rate convention may have outlived its usefulness. Rudebusch has already in this paper estimated one-sixth (one equation) of a VAR. I urge him to take the plunge and estimate a full scale dynamic model, in which policy shocks are identified and the interaction of funds and discount rate setting with private-sector based disturbances to market rates is modeled explicitly. It appears from this paper as if he would not like to think of himself as a VAR modeler. In that case I urge him to present us with another approach to dynamic modeling that can progress in the same direction.

As for my own views of what the literature lacks, they are partly implicit in Leeper et al., (1996) which models the joint behavior of multiple monetary aggregates (one 'controllable' like total reserves or nonborrowed reserves, one 'transactions' like M1 or M2) and multiple interest rates. Such a model has to be bigger than existing identified VAR models, and hence requires some new numerical and statistical approaches, but proves to be feasible.

The restriction of identified VAR modeling to handling only either just-identified models or over-identified models that restrict only contemporaneous coefficients is artificial. It is time for some move in the direction of relaxing this computationally based constraint.

Ultimately the identified VAR literature and the dynamic stochastic general equilibrium model (DSGE) literature should converge. I see no satisfactory way to handle parameter drift and nonlinearity in identified dynamic models that does not require more complete behavioral interpretations than are used in the identified VAR literature. Leeper and Sims (1994) and Jinill Kim (1995) represent steps in that direction. They use models that are nonlinear in parameters and that model explicitly time variation in interpretable parameters.

REFERENCES

Bernanke, B. and I. Mihov, "Measuring Monetary Policy," Presented at the CEPR/Federal Reserve Bank of San Francisco Monetary Policy Conference, March 1995.

Kim, J., "Effects of Monetary Policy in a Stochastic Equilibrium Model with Real and Nominal Rigidities," manuscript, 1995, Yale University.

Kim, S., "Does Monetary Policy Matter in the G-6 Countries? Using Common Identifying Assumptions about Monetary Policy across Countries," manuscript, 1995, Yale University.

Leeper, E. and C. Sims, "Toward a Modern Macro Model Usable for Policy Analysis," (with Eric Leeper), *NBER Macroeconomics Annual* (1994), 81–117.

———, C. Sims and T. Zha, "What Does Monetary Policy Do?" *Brookings Papers on Economic Activity*, 2 (1996).

Sims, C., "A 9 Variable Probabilistic Macroeconomic Forecasting Model," *Business Cycles, Indicators, and Forecasting*, James H. Stock and Mark W. Watson, eds., NBER Studies in Business Cycles 28, (1993), 179–214.

———, "Are Forecasting Models Usable for Policy Analysis?" Minneapolis Federal Reserve Bank *Quarterly Review*, Winter (1986), 2–16.