



Commentary

Lee E. Ohanian

In “Model Fit and Model Selection,” Narayana Kocherlakota (2007) warns econometricians and the users of econometric analyses that macroeconomic models that fit the data well—as measured by a high R^2 and/or low residual variance—may not be very useful for policy advice because key parameters may not be identified. As an alternative, Kocherlakota provides a Bayesian approach that recognizes the significant challenge of identifying all parameters in a fully specified general equilibrium model and that also treats uncertainty about parameter values in a consistent fashion.

Kocherlakota’s agnostic approach means that the range of uncertainty associated with the conditional forecasts (policy advice) generated by the macroeconomic models used by central banks and other policymaking agencies is probably much larger than recognized by macroeconomic practitioners. This also suggests that policymakers, who use the forecasts from these models as an input into policymaking, should also modify their priors and recognize the considerable uncertainty in conditional forecasts.

There is substantial evidence that supports Kocherlakota’s recommendation, and there is also substantial evidence that this practice—or other practices that explicitly recognize the degree of uncertainty in modeling the economy—is not followed by macroeconomic model builders. Nor is agnosticism followed by policymakers regarding the structure of the economy. Instead, current model-building practice focuses largely

on models that feature a Phillips curve, and recent monetary policy decisions also appear to focus on the Phillips curve. I first discuss Kocherlakota’s analysis of fit and identification and then discuss the broader issues of agnosticism in choosing among alternative theoretical frameworks for evaluating policy and the role of agnosticism in making monetary policy.

AGNOSTICISM IN MODELING

Kocherlakota reminds macroeconomic modelers and the users of these models that a model with a good fit may not be a useful tool for conditional forecasting and policy analysis. To summarize practitioners’ current view about fit and its implications for policy analysis, Kocherlakota quotes from recent influential work by Smets and Wouters (2003), who suggest that a useful model for policy analysis is one that includes enough shocks to fit the data well. Smets and Wouters’s view is quite representative of macroeconomic modeling strategies used today.

To illustrate why a model that fits well may be not be useful for conditional forecasting, Kocherlakota constructs a model economy in which a model fits the data perfectly. He then shows that despite the perfect fit, the model cannot accurately forecast the impact of a tax cut on the economy because the elasticity of labor supply is unidentified. The reason that the perfect-fitting model provides a poor conditional forecast is

Lee E. Ohanian is a professor of economics at the University of California at Los Angeles, a visiting scholar at the Federal Reserve Bank of Minneapolis, and a research associate of the National Bureau of Economic Research.

Federal Reserve Bank of St. Louis *Review*, July/August 2007, 89(4), pp. 361-69.

© 2007, The Federal Reserve Bank of St. Louis. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

because it includes a non-testable identifying restriction that is false.

Kocherlakota suggests an alternative procedure that completely discards the non-testable identifying restriction. As an alternative, Kocherlakota recommends using auxiliary information to construct a range of values that a parameter can take. This is not standard Bayesian analysis, however, in that the standard approach commits the investigator to a single prior. Instead, Kocherlakota's procedure considers many priors over the parameter of interest. This delivers a collection of posteriors, which are restricted only in that they have support between the minimum and maximum values specified for the parameter. In principle, this approach is very sensible, as it explicitly and systematically allows the researcher to conduct a sensitivity analysis.

From a practical perspective, however, this procedure may be difficult to apply. To see this, note that in Kocherlakota's example of the impact of a tax cut, there is a single parameter. In this case, the application of the agnostic procedure yields a collection of posteriors over this single parameter, with support over a minimum and maximum value. This one-dimensional case is fairly simple to implement and to investigate. However, in a high-dimensional setting, the investigator must specify *many* priors over *several* parameters. Specifying multidimensional priors can be difficult; in practice, specifications are often chosen for computational ease, but in this case the prior is particularly important because the effect of the prior does not wash out as the sample size becomes arbitrarily large, as in standard analysis with full identification. Understanding how various multidimensional priors affect the analysis is very much an open, and difficult, question. Moreover, understanding how to distill and interpret the information from a collection of posteriors is also an open question. Making progress on these fronts seems necessary to successfully apply the agnostic, Bayesian approach in any rich model that includes many parameters and/or shocks.

Kocherlakota's agnostic approach presumes that there is an inherent identification problem in macroeconomics that is not easy to resolve. Is

the identification problem in macroeconomics as difficult as suggested by Kocherlakota? Unfortunately, there is no easy answer to this question; the profession has wrestled with identification in aggregate economics since the work of Tinbergen (1937), Haavelmo (1944), the Cowles Commission (Koopmans, 1950), Liu (1960), Sims (1980), and it continues today (Canova and Sala, 2006).

The identification debate in macroeconomics has suggested many different resolutions to the problem. One must understand the differences: Sims (1980) viewed the identification challenge in macroeconomics a sufficiently tall order to fill as to recommend relatively unrestricted vector autoregression (VAR) models that achieved identification by imposing a sufficient number of lags in the VAR to generate white noise innovations and then impose a Wold causal ordering on the innovation covariance matrix. In contrast, Kocherlakota recommends a very different approach, in which the behavioral equations of the model are tightly restricted by theory, but only minimal restrictions are imposed on the structure of the shock processes. Regarding the relative merits of these two different approaches, identification achieved through restrictions on shock processes are often difficult to justify because economic theory typically does not shed much light on the correct stochastic specification of shocks. Moreover, evaluating the identification of shocks is difficult, as identifying shocks almost always requires strong non-testable restrictions.

Some economists argue that shocks should be uncorrelated, and that this apparently innocuous assumption can go a long ways toward achieving identification. But we have several observations that shocks can be correlated. For example, there were several scientific, productivity breakthroughs in World War II that were largely the consequence of the large wartime government spending shock. Similarly, World War II monetary shocks were due to fiscal spending requirements that induce the need for seignorage finance. The deregulation of financial markets over the past 30 years has led to significant technological change in financial intermediation. The Great Depression led to enormous changes in economic regulation and government management of the economy. These

are just a few examples that indicate that achieving identification through orthogonality assumptions can be at variance with the data.

Identification will always be a difficult issue in macroeconomics, as maintained identifying restrictions are by definition not testable and are almost always open to debate. But what researchers can do is not make the identification problem any more difficult than it needs to be, and here Kocherlakota's recommendations are particularly valuable. The focus on fit, as exemplified by Smets and Wouters, tends to increase the difficulty of the identification problem. This is because increasing the richness of the model—by including more shocks—makes identification harder by requiring more restrictions to be placed on the shock process. From this perspective, relatively simple models may be easier to identify than densely parameterized models.

It is puzzling that the profession needs to be reminded of the “fallacy of fit” (Kocherlakota's words). The profession learned this dictum the hard way in the 1970s, when the apparently well-fitting large-scale macroeconomic models broke down, particularly the Phillips curve (the inflation-unemployment relationship), which was a central component of these models. Specifically, the 1970s witnessed both unemployment and inflation rising to levels far outside their fitted historical relationship. At the same time Charles Nelson (1972) showed that atheoretic, low-order univariate ARMA models of macroeconomic time series—that typically were characterized by a worse fit than the large-scale models—dominated the large-scale models in forecasting competitions. Further improvements in forecasting were generated by pseudo-Bayesian VARs, which imposed random-walk priors on time series to reduce the problem of overparameterization that is inherent in VARS. Bayesian VARs are used for forecasting at several research agencies and commercial banks, including the Minneapolis Fed and the Richmond Fed. All of these events led to traditional large-scale econometric models playing a much smaller role in central bank research and in policymaking.

So why did central bank researchers and policymakers return to macroeconomic models

after these failures? One reason stems from the fact that current models feature a much deeper structure than their large-scale predecessors. Today's models include dynamically maximizing households, maximizing firms, an internally consistent set of expectations, and a precise definition of equilibrium. All of these advances were absent from earlier models, and it is believed by many that these improvements would allow macroeconomic models to successfully confront the Lucas critique. Nevertheless, it is critical to distinguish between a model with a deep structure that in principle can be used for conditional forecasting and a model in which the parameters are reasonably identified. The first feature doesn't imply the second; ironically, specifying rich, fully articulated general equilibrium models will tend to make the identification problem even more difficult.

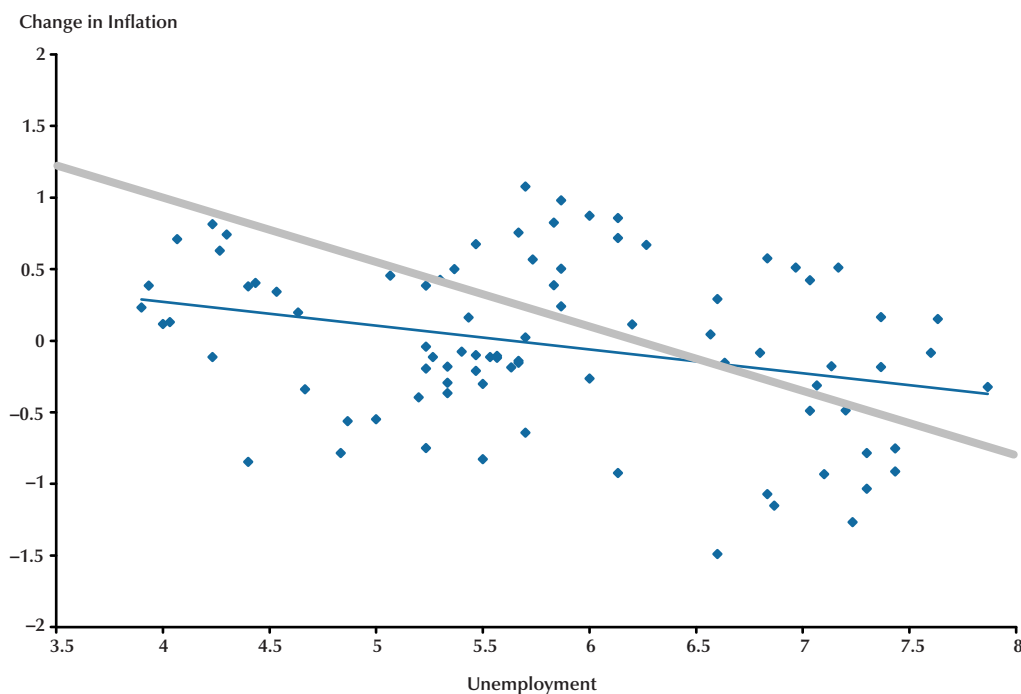
Applied economists face a difficult trade-off in specifying macroeconomic models. Simple models may in principle be easier to identify, conditional on correct specification, but simple models will tend to be false, and thus may not be useful for parameter estimation, at least from a classical perspective. Richer models may be more difficult to identify but, conditional on identification, may fare better in terms of parameter estimation. This trade-off is one reason why *calibration*, which sidesteps these difficult issues, has been so popular among applied macroeconomists. Kocherlakota's approach is another proposal in a research program that has attempted to place calibration into either an explicit classical or Bayesian framework (see Watson, 1993; Diebold, Ohanian, and Berkowitz, 1998; Schorfheide, 2000; and Fernández-Villaverde and Rubio-Ramírez, 2004).

THE PHILLIPS CURVE PRIOR IN MACROECONOMETRIC MODELING

Central bank research staff have strong priors on the class of models that are used in monetary policy analysis. The dominant class of models are those with a *Phillips curve*. Here, I define the

Figure 1

Is the NAIRU Phillips Curve Still There? Unemployment vs. Future Changes in Inflation, 1984-2006



Phillips curve as the view that during periods of slack economic capacity—such as a period of relatively high unemployment—rapid economic growth will not raise inflation very much; but during periods of low unemployment, economic growth can raise inflation considerably. The Phillips curve view has implications for monetary policy. Specifically, it implies that monetary stimulus during periods of high unemployment will not raise inflation very much and that there is scope for the Fed to moderate recessions (which are periods of slack capacity) through expansionary monetary policy. It also implies that, as capacity becomes tight, the Fed controls inflation by attempting to reduce the growth rate of the real economy through monetary contraction.

It may be reasonable for model builders and policymakers to narrowly focus on this class of models if there is strong empirical support for the Phillips curve. In contrast, if there is limited

support for this class of models, then there is scope to consider alternative theoretical channels for the determination of inflation. Here, I present U.S. time-series evidence that shows little support for the view that inflationary risks are significantly higher during periods of rapid growth and tight capacity, relative to rapid growth and slack capacity.

Atkeson and Ohanian (2001), Stock and Watson (2006), and others have recently analyzed the Phillips curve in U.S. time series. Under the Phillips curve view, low unemployment should be associated with rising inflation. This is often referred to as the NAIRU (non-accelerating inflation rate of unemployment) Phillips curve. Figure 1 shows the NAIRU Phillips curve by presenting the change in inflation and the unemployment rate for the period 1960-2006. The figure updates my earlier study with Atkeson to include data from 2000-06. The heavy gray line is an ordi-

nary least squares (OLS) regression line, which shows a modest negative relationship between the change in inflation and unemployment for 1960-83. The blue line is the OLS regression line between these variables for 1984-2006. The slope coefficient is very close to zero and is also statistically insignificantly different from zero. This latter result indicates that there has been no systematic relationship between the change in the inflation rate and unemployment since 1984. In other words, there has been no simple NAIRU Phillips curve in U.S. data for more than 20 years.

In Atkeson and Ohanian (2001), we extended our analysis of the Phillips curve by examining whether changes in unemployment, or other measures of slack capacity, help forecast future inflation relative to a naive forecasting model that simply extrapolates the current inflation rate into the future. Surprisingly, we found that inflation was not forecasted well by measures of slack capacity. In particular, the root mean-squared forecast error (RMSE) for the core consumer price index, which is a measure of inflation that excludes volatile food and energy prices and a key indicator of inflation for both financial markets and central banks, is as much as 94 percent higher compared with the forecast from the naive model that extrapolates the current inflation rate into the future. More sophisticated forecasting models did not fare much better: for example, inflation forecasts from Stock and Watson's macroeconomic activity index model, which forecasts inflation from a much larger information set than just unemployment. This model had an RMSE of between 33 and 81 percent higher than the naive forecast. These results indicate that there is no significant, predictable relationship between cyclical fluctuations in the real economy and future inflation. Paradoxically, forecasts from sophisticated models, which clearly fit much better in sample, are deficient to those from very simple models that do not fit so well in sample, such as the naive model we used in Atkeson and Ohanian (2001).

Tables 1 through 3 show the results of other tests of the Phillips curve. These tests evaluate whether economic growth generates inflation differentially when unemployment is low (tight

Table 1

Testing the Phillips Curve Correlation Between GDP Growth and Inflation, Controlling for Unemployment

Period	Deflator	Core CPI
1957:Q1–2006:Q2	–0.23	–0.20
1960s	–0.20	–0.23
1970s	–0.43	–0.40
1980s	–0.28	–0.11
1990s	–0.45	–0.44
2000:Q1–2006:Q2	–0.05	–0.22

capacity) relative to when unemployment is high (slack capacity). Table 1 shows the correlation between quarterly gross domestic product (GDP) growth and inflation, measured by the core CPI and by the GDP deflator for the period 1957-2006, and also for each decade during the past 50 years. The data are conditioned on the unemployment rate as a measure of capacity. The most striking finding is that the relationship between growth and inflation is negative, not positive as suggested by the Phillips curve. Tables 2 and 3 show correlations between GDP growth and inflation, conditioned on unemployment, for various leads and lags up to 1 year (4 quarters). The correlations are primarily negative or close to zero.

These tests indicate that there is not sufficient evidence to support the strong prior for the Phillips curve that characterizes current econometric practice in central banks. Instead, these results suggest that models with alternative inflation mechanisms should be analyzed.

THE PHILLIPS CURVE PRIOR AND MONETARY POLICY

The strong emphasis of the Phillips curve in macroeconomic research at central banks suggests that this view is also prominent among the policymakers who are the primary users of central bank economic research. The importance of the Phillips curve in policymaking appears to vary over time. There does not appear to be a substantial Phillips

Table 2

Testing the Phillips Curve Correlation Between GDP Growth and Inflation (Core CPI): Leads and Lags, Controlling for Unemployment

1957:Q1–2006:Q2	$i = 4$	$i = 3$	$i = 2$	$i = 1$
$\Delta \ln (y_{t+i}), \Delta \ln (p_t)$	–0.23	–0.45	–0.41	–0.28
$\Delta \ln (y_t), \Delta \ln (p_{t+i})$	–0.02	0.04	–0.10	–0.12

Table 3

Testing the Phillips Curve Correlation Between GDP Growth and Inflation (Deflator): Leads and Lags, Controlling for Unemployment

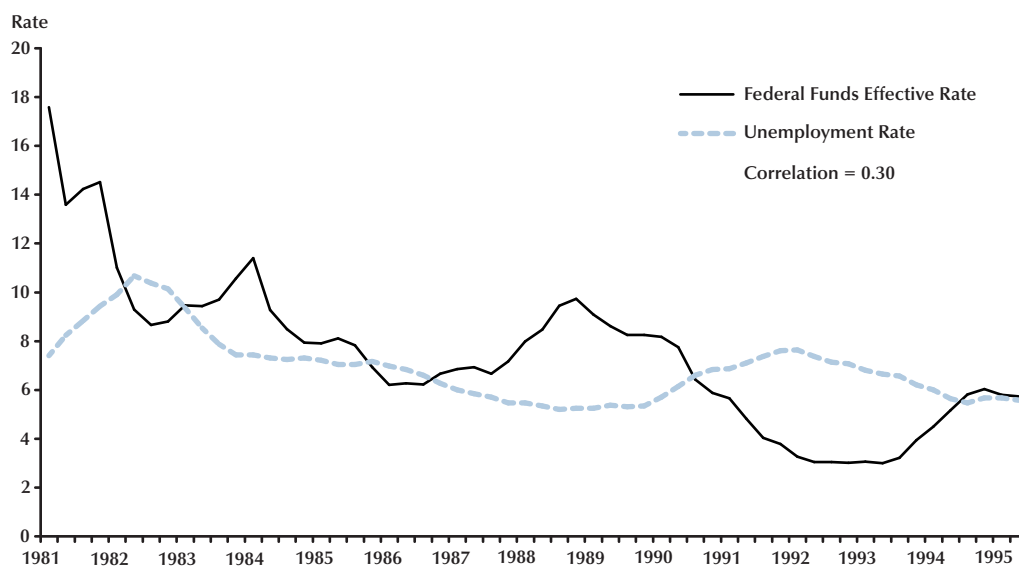
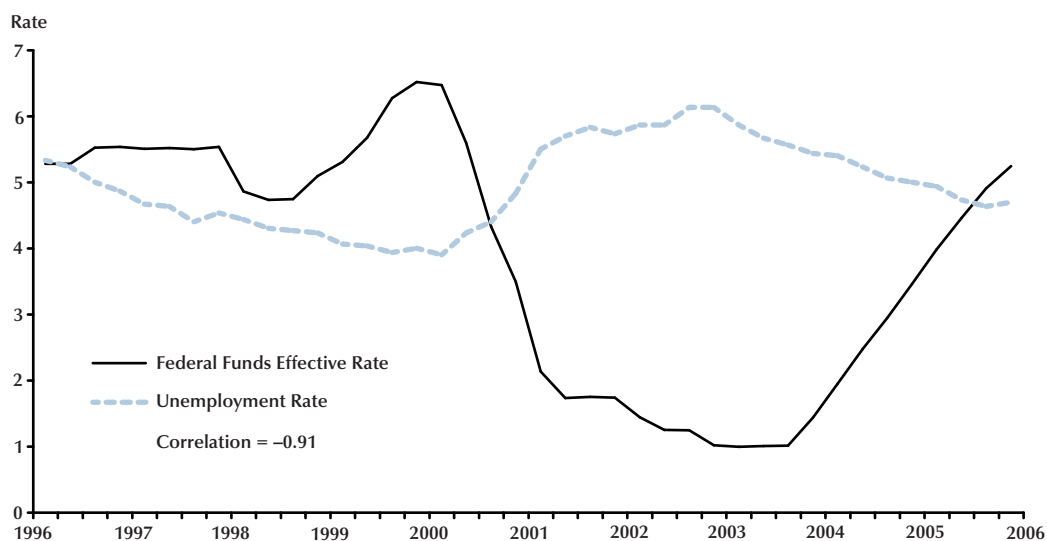
1957:Q1–2006:Q2	$i = 4$	$i = 3$	$i = 2$	$i = 1$
$\Delta \ln (y_{t+i}), \Delta \ln (p_t)$	–0.27	–0.29	–0.26	–0.22
$\Delta \ln (y_t), \Delta \ln (p_{t+i})$	0.04	–0.04	–0.09	–0.16

curve policy bias during the Volcker-Greenspan disinflation of 1982–95. This period is certainly one of the great triumphs of central banking. After engineering the largest peacetime inflation in the history of the United States, with inflation rising from about 1 percent in the early 1960s to more than 13 percent by 1980, the Fed lost credibility with financial markets. By 1980, long-term interest rates rose to 13 percent. A standard Fisher equation decomposition, which relates nominal interest rates to expected inflation over the horizon of the security, clearly indicates that financial markets were systematically expecting permanent high inflation. Beginning at this time, however, Paul Volcker initiated a low-inflation monetary policy, and inflation declined to less than 3 percent by the mid-1990s.

To analyze the potential impact of the Phillips curve on monetary policy, I examine the relationship between the federal funds rate and the unemployment rate. If there is a strong influence of the Phillips curve on policy, then we should observe a systematic inverse relationship between the funds rate and the unemployment rate. Figure 2 shows monthly data on these variables between 1981 and 1995. Note that there is little systematic relationship between the federal funds rate and the unemployment rate (the correlation is about 0.3), suggesting that policy was not par-

ticularly focused on the Phillips curve. This is not surprising, as there is little disagreement among economists or financial market participants that monetary policy during this period was unconditionally committed to reducing inflation, without much reference to the business cycle. The policy was indeed effective, as inflation fell and long-term interest rates fell.

But the nature of policy seemed to change considerably after inflation declined. Figure 3 shows the funds rate and the unemployment rate between 1996 and 2006. This figure shows a distinct and systematic inverse relationship between unemployment and the funds rate, with a correlation of –0.91. As the unemployment rate declined to 4 percent in 1999 and 2000, Fed officials worried about tight labor markets and inflationary pressures and raised the funds rate. Then, as the unemployment rate rose from 4 percent to more than 6 percent, the Fed pursued a more expansionary policy, driving the funds rate down from 6.5 percent in late 2000 to just 1 percent in late 2003. The policy record since 1995 is consistent with a strong Phillips curve prior and is reminiscent of the “fine tuning” that policymakers pursued in the 1960s and 1970s. U.S. time-series data provides little support that a fine-tuning policy based on Phillips curves will be successful.

Figure 2**Monetary Policy Without a Phillips Curve Focus****Figure 3****Unemployment Rate–Driven Monetary Policy? The Return of Phillips Curve Policymaking**

CONCLUSION

During the 1960s, policymakers believed they understood the U.S. economy so well that they could achieve virtually any desired result with the appropriate mix of fiscal and monetary stimulus or contraction. Part of this belief stemmed from the close fit macroeconomic model builders were able to achieve with large-scale macroeconomic models. This belief ended abruptly in the stagflation of the 1970s, in which the perceived stable and systematic trade-off between unemployment and inflation broke down and both unemployment and inflation rose to unprecedented postwar levels. The belief that tight-fitting models could generate accurate conditional forecasts also broke down and formed the basis of Robert Lucas's famous critique of econometric models.

Macroeconomics and economic modeling have advanced enormously since the large-scale models of the 1960s, and these advances are largely responsible for the return of macroeconomic model building to the forefront of central bank research and policymaking. But as Kocherlakota points out, identification of all parameters in these models is tenuous, particularly in models with many shock processes. Good macroeconomic practice almost by necessity requires sensitivity analysis that provides a systematic treatment of the uncertainty underlying model parameters. And when there is considerable uncertainty in conditional forecasts, policymakers should recognize this uncertainty as well.

Current policy and the current menu of models analyzed appear to be too responsive to the Phillips curve, more so than is warranted by the data. Model fit is a seductive property; it is hard for model builders to resist modifying model equations to achieve a better fit, even when the modifications do not have strong theoretical underpinnings. Fromm and Klein (1965) show case how model builders of the 1960s focused on fit and modified models to achieve low mean square error, despite the fact there were few, if any, economic foundations for these modifications.

Econometric practice today is in some ways reminiscent of 1960s practice. Shocks are being

added to various equations to achieve a close fit to the data, but without necessarily understanding deeply what the frictions or market imperfections underlying these shocks are. And current policymaking seems far too responsive to a perceived Phillips curve that is not present in the data. We know all too well the outcome of the fitting exercises of the 1970s and the reliance on the Phillips curve. Perhaps the best way to avoid the monetary policy mistakes of the past is to remember that these mistakes were partly the consequence of relying too much on an empirical relationship that does not have strong theoretical underpinnings and that is not a robust feature of U.S. data. Agnostic approaches to modeling, as suggested by Kocherlakota, can significantly aid in the process of quantifying macroeconomic uncertainty and understanding its implications for monetary policy.

REFERENCES

- Atkeson, Andrew and Ohanian, Lee E. "Are Phillips Curves Useful for Forecasting Inflation?" Federal Reserve Bank of Minneapolis *Quarterly Review*, Winter 2001, 25(1), pp. 2-11.
- Canova, Fabio and Sala, Luca. "Back to Square One: Identification Issues in DSGE Models." Working Paper 583, European Central Bank, 2006.
- Diebold, Francis X.; Ohanian, Lee E. and Berkowitz, Jeremy. "Dynamic Equilibrium Economies: A Framework for Comparing Models and Data." *Review of Economic Studies*, July 1998, 65(3), pp. 433-51.
- Fromm, Gary and Klein, Laurence R. *The Brookings Quarterly Econometric Model of the United States*. Chicago: Rand-McNally, 1965.
- Haavelmo, Trygve. "The Probability Approach in Econometrics." *Econometrica*, 1944, 12(Suppl), pp. 1-118.
- Kocherlakota, Narayana. "Model Fit And Model Selection." Federal Reserve Bank of St. Louis *Review*, July/August 2007, 89(4), pp. 349-60.

- Koopmans, Tjalling, ed. *Statistical Inference in Dynamic Economic Models*. Cowles Commission Monograph Series No. 10. New York: John Wiley, 1950.
- Liu, Ta-Chung. "Underidentification, Structural Estimation, and Forecasting." *Econometrica*, October 1960, 28(4), pp. 855-65
- Nelson, Charles R. "The Prediction Performance of the FRB-MIT-PENN Model of the U.S. Economy." *American Economic Review*, December 1972, 62(5), pp. 902-17,
- Fernández-Villaverde, Jesús and Rubio-Ramírez, Juan. "Comparing Dynamic Equilibrium Economies to Data: A Bayesian Approach." *Journal of Econometrics*, November 2004, 123(1), pp. 153-87.
- Schorfheide, Frank. "Loss Function-Based Evaluation of DSGE Models." *Journal of Applied Econometrics*, November-December 2000, 15(6), pp. 645-70.
- Sims, Christopher. "Macroeconomics and Reality." *Econometrica*, January 1980, 48(1), pp. 1-48.
- Smets, Frank and Wouters, Raf. "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association*, September 2003, 1(5), pp. 1123-75.
- Stock, James and Watson, Mark. "Has Inflation Become Harder to Forecast?" Discussion paper, Harvard University, 2006.
- Tinbergen, Jan. *An Econometric Approach to Business Cycles*. Paris: Fr. Hermann Publishing, 1937.
- Watson, Mark W. "Measures of Fit for Calibrated Models." *Journal of Political Economy*, December 1993, 101(6), pp. 1011-41.

