VOLUME 5, NUMBER 2 JUNE 24, 2010

MACROECONOMIC ADVISERS'

MACRO FOCUS

THE EMPEROR HAS NO CLOTHES

Much has been made of the failure of modern macroeconomics to predict or understand the Great Recession of 2007–2009. In this MACRO FOCUS, our resident time-series econometrician, James Morley, explains what is currently meant by "modern" macroeconomics, what is behind its failure, and what can be done to rehabilitate its reputation.*

"MODERN" MACROECONOMICS

In a recent essay, Narayana Kocherlakota, President of the Federal Reserve Bank of Minneapolis, acknowledged that modern macroeconomics failed during the recent financial crisis.[1] However, his essay misses the point of why it failed.

Like many in academia, Kocherlakota associates modern macroeconomics with a particular school of thought that takes something called the "Lucas critique" as its guiding principle. The Lucas critique refers to an argument put forth by the Nobel Prize-winning macroeconomist Robert Lucas about how the changing expectations of economic agents will confound forecasting and policy analysis based on macroeconomic data.[2] Its main implication is that an economic model with "deep structural parameters" related to preferences and technology for households and firms should provide more reliable forecasts, especially when predicting the effects of policy, than a model based more on the apparent historical correlations between macroeconomic variables. This is sometimes referred to as the "microfoundations" approach to macroeconomics because it presumes that a microeconomic structure — in particular, the metaphor of optimizing economic agents— is more robust to changes in the policy environment than macroeconomic correlations.[3]

Rather than question the relevance of the Lucas critique, Kocherlakota explains the recent failure of modern macroeconomics as due to much narrower issues. In his view, the micro-founded models failed because they lack sufficient complexity, especially in terms of their treatment of financial markets. Also, he points out, rightly, that the models are driven by "patently unrealistic shocks".

However, the rehabilitation of modern macroeconomics requires a different tack than suggested by Kocherlakota. In particular, macroeconomists need to do more than simply add complexity to their models. They should also remember that it is empirically testable whether models that put most of their weight on "deep structural parameters" produce more accurate predictions than models that put more weight on historical correlations. In doing so, it may

* Without implicating anyone, the author wishes to thank Joel Prakken, Ken Matheny, Antulio Bomfim, Alan Auerbach, Tara Sinclair, Marc Law, Jeremy Piger, Steve Fazzari, and Sharon Kozicki for helpful and critical comments.

[1] Kocherlakota, N., 2010, "Modern Macroeconomic Models as Tools for Economic Policy," *The Region*, May, [pdf].

[2] Lucas, R.E., 1976, "Economic Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1, 19–46.

[3] The optimizing agents in a micro-founded model are a metaphor because of the insurmountable problems involved in aggregating across actual households and firms. On this issue, the economist/philosopher Kevin Hoover has written many compelling critical analyses of the microfoundations approach to macroeconomics. For an example that is particularly relevant to Kocherlakota's essay, see Hoover, K., 2006, "A NeoWicksellian in a New Classical World: The Methodology of Michael Woodford's Interest and Prices," *Journal of the History of Economic Thought* 28, 143-149, [pdf].

231 S. BEMISTON
SUITE 900
ST. LOUIS, MO 63105
PHONE: 314-721-4747
FAX: 314-721-6383

1776 EYE STREET, NW
SUITE 1030
WASHIINGTON, DC 20006
PHONE: 202-521-0302
FAX: 202-449-6715

E-MAIL:
MAIL@MACROADVISERS.COM
VISIT US ON THE WEB:
WWW.MACROADVISERS.COM

be found that some macroeconomic relationships are useful even if they cannot be easily motivated as the literal outcome of a micro-founded model. This is not to deny the important role that economic theory plays in the formulation of models used for explanation, prediction, and policy analysis. However, there is no reason for models to take theory quite so literally as is typically done in modern macroeconomics. Instead, the data should be taken more seriously.

History Repeats Itself

The idea of the Lucas critique arose out of the 1970s. It was a time when large-scale macroeconometric models that relied heavily on historical correlations — especially the traditional Phillips curve tradeoff between unemployment and inflation — failed to predict or even explain "stagflation" in the form of simultaneously high rates of unemployment and inflation.

After their failure in the 1970s, the large-scale macroeconometric models were modified to include supply shocks and changes in inflation expectations. With these additions, the models could explain stagflation *ex post*. But their *ex ante* failure represented a serious blow to their reputation, especially within the ivory towers of academia.

Ironically, this historical episode should remind us somewhat of the present. Now it is "dynamic stochastic general equilibrium" (DSGE) models inspired by the Lucas critique that have failed to predict or even explain the Great Recession of 2007–2009. More precisely, the implicit "explanations" based on these models are that the recession, including the millions of net jobs lost, was primarily due to large negative shocks to both technology and willingness to work. In his essay on modern macroeconomics, Kocherlakota admits the inadequacy of these explanations:

> *… most models in macroeconomics rely on some form of large quarterly movements in the technological frontier (usually advances, but sometimes not). Some models have collective shocks to workers' willingness to work. Other models have large quarterly shocks to the depreciation rate in the capital stock (in order to generate high asset price volatilities). To my mind, these collective shocks to preferences and technology are problematic. Why should everyone want to work less in the fourth quarter of 2009?…*

So can the reputation of modern macroeconomics be rehabilitated by simply modifying DSGE models to include a few more realistic shocks? As discussed below, the problems for DSGE models run deeper than a lack of complexity, while the large-scale macroeconometric models have improved considerably since the 1970s.

The Rabbit and the Hat

A simple example helps illustrate for the uninitiated just how DSGE models work and why it should come as little surprise that they are largely inadequate for the task of explaining the Great Recession.

For this simple DSGE model, consider the following technical assumptions: i) an infinitely-lived representative agent with rational expectations and additive utility in current and discounted future log consumption and leisure; ii) a Cobb-Douglas aggregate production function with labor-augmenting technology iii) capital accumulation with a fixed depreciation rate; and iv) a stochastic process for exogenous technology shocks.

Before discussing the particular implications of this model, it is worth making two basic points about the setup. First, by construction, technology shocks are the only underlying source of fluctuations in this simple model. Thus, if we were to assume that U.S. real GDP was the literal outcome of this model, we would be assuming *a priori* that fluctuations in real GDP were ultimately due to technology. When faced with the Great Recession, this model would have no choice but to imply that technology shocks were somehow to blame.[4] Second, despite the underlying role of technology, the observed fluctuations in real GDP can be divided into those that directly reflect the behavior of the exogenous shocks and those that reflect the endogenous capital accumulation in response to these shocks.

To be more precise about these two points, it is necessary to assume a particular process for the exogenous technology shocks. In this case, let's assume technology follows a random walk with drift:

$$\text{(1)} \qquad \ln A_t = \mu + \ln A_{t-1} + u_t \,,$$

where A_t is labor-augmenting technology in the aggregate production function, μ is the rate of drift, and $u_t \sim N(0, \sigma_u^2)$ is the underlying technology shock. Why this particular form? Well, because, given the economic assumptions outlined above and, for the time being, assuming a 100% depreciation rate for capital, the model implies the following behavior for real GDP:

$$\text{(2)} \qquad \Delta \ln Y_t = (1-\alpha)\mu + \alpha \Delta \ln Y_{t-1} + (1-\alpha) u_t \,,$$

where Y_t is real GDP and α is the capital share in the aggregate production function.[5]

So, with this simple DSGE model and for typical measures of the capital share, we have the implication that output growth follows an AR(1) process with an AR coefficient of about one third. This is notable given that such a time-series model does reasonably well as a parsimonious description of quarterly real GDP dynamics for the U.S. economy. In particular, the estimated AR coefficient for different sample periods of postwar quarterly U.S. real GDP growth is fairly robust around 0.33 and such a model does well in terms of various model selection criteria.

At this point, it is tempting to think of Milton Friedman's classic essay on methodology in which he argued that we should not judge a model based on its assumptions, but instead we should focus on its predictions.[6] However, the rather absurd assumption of a 100% depreciation rate at the quarterly horizon would surely still have prompted a sharp question

[4] To the extent that other shocks, such as exogenous changes in financial conditions or oil prices, were actually to blame, they would be labeled by the model as technology shocks. In particular, the model would be misspecified and its estimates for the effects of technology shocks would be distorted.

[5] For more details of this derivation, see McCallum, B., 1989, "Real Business Cycle Models," in R.J. Barro, ed., Modern Business Cycle Theory, Harvard University Press, and Ma, J. and M. Wohar, 2009, "Real and Nominal Business Cycles: New Evidence from a Generalized Unobserved Components Model," Working Paper.

or two in a University of Chicago seminar back in the days. So, with this in mind, what happens if we consider the more general case?

Unfortunately, for more realistic depreciation rates, we cannot solve the model analytically. Instead, taking a log-linearization around steady state, we can use standard methods to solve for output growth as

(3) $$\Delta \ln Y_t = a_1 \Delta \ln Y_{t-1} + a_2 u_t - a_3 a_2 u_{t-1},$$

where the "hats" denote deviations from steady-state levels and the "*a*" coefficients are complicated nonlinear functions of the deep structural parameters describing preferences, depreciation, and the diffusion of technology. For convenience, letting "*L*" denote the lag operator such that $LY_t \equiv Y_{t-1}$, equation (3) can also be written as

(4) $$(1 - a_1 L)\Delta \ln Y_t = a_2 (1 - a_3 L) u_t.$$

The problem, as pointed out by Timothy Cogley and James Nason, is that a_1 and a_3 will be very close to each other given "realistic" values of the structural parameters.[7] Specifically, for $a_1 \approx a_3$, it is easy to see from a near cancellation of the $(1 - a_1 L)$ and $(1 - a_3 L)$ terms in equation (4) that the model now implies output growth is approximately serially uncorrelated:

(5) $$(1 - a_1 L)\Delta \ln Y_t \approx a_2 u_t$$

Thus, as the assumptions about depreciation and other parameters are made more realistic, the ability of the model to mimic the estimated AR(1) process for quarterly U.S. real GDP is severely weakened.

But fear not for this lowly single-shock DSGE model! There is still one more arrow in the quiver. What if it *just so happened* that the technology shock followed an AR(1) process instead of being serially uncorrelated? In particular, let's assume

(6) $$u_t = \phi u_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. Then, given this AR(1) process for technology shocks, we would obtain the following solution for output growth (again, using the lag operator notation for convenience):

(7) $$(1 - a_1 L)(1 - \phi L)\Delta \ln Y_t = a_2 (1 - a_3 L)\varepsilon_t.$$

As before, $a_1 \approx a_3$ implies the a near cancellation of the $(1 - a_1 L)$ and $(1 - a_3 L)$ terms and, reverting to standard notation, the process in equation (7) can be approximated as

[6] Milton Friedman, 1953, "The Methodology of Positive Economics" in Essays in Positive Economics, The University of Chicago Press.

[7] Cogley, T. and J.M. Nason, 1993, "Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model, *Economics Letters* 43, 77-81 and Cogley, T. and J.M. Nason, 1995a, "Output Dynamics in Real-Business-Cycle Models," *American Economic Review* 85, 492-511.

$$(8) \qquad (1-\phi L)\Delta \ln Y_t \approx a_2\varepsilon_t.$$

Thus, this model implies that output growth is approximately a first-order autoregressive process with an AR coefficient of ϕ, which we could simply assume to be 0.33.

However, let's take moment to reflect on this result. This simple DSGE model is able to mimic the apparent AR(1) dynamics in real GDP growth. But it does so by assuming the exogenous technology shocks also follow an AR(1) process with an AR coefficient that happens to be the same as the estimated AR coefficient for output growth. Thus, the magic trick has been revealed: a rabbit was stuffed into the hat and then a rabbit jumped out of the hat.[8]

Ever-Increasing Sophistication?

The simple DSGE model in the previous section may be familiar to readers under its alternative label of a "real business cycle" (RBC) model. Modern DSGE models typically have more complicated preferences (e.g., habit formation), production technologies, and market structures than the original RBC models. Frictions such as sticky prices even open up the possibility of using these models for monetary policy analysis, something that RBC models assume *a priori* doesn't really matter for real economic activity.

However, despite their increasing sophistication, DSGE models share one key thing in common with their RBC predecessors. After more than two decades of earnest promises to do better in the "future directions" sections of academic papers, they still have those serially-correlated shocks.[9] Thus, the models now "explain" variables like real GDP, inflation, and interest rates as the outcome of more than just serially-correlated technology shocks. They also consider serially-correlated preference shocks and serially-correlated policy shocks.

Revisiting Friedman's essay on methodology, it should be recalled that explaining economic phenomena as being due to changing preferences is basically the nightmare scenario for economists, as it implies the eventual ascendancy of sociology over economics. Meanwhile, in terms of policy, it is important to recognize the following fundamental issue with the DSGE framework. Given the assumption that the data are the literal outcome of the model, it will be the case, by construction, that policy is to blame for the severity of the welfare loss relative to a frictionless equilibrium implied by technology and preference shocks. For example, if policy in the model is monetary policy, then any bad outcome that is not directly due to a negative technology shock or a preference shock — which is harder to label as being either "good" or "bad" — must be due to suboptimal monetary policy failing to mitigate the effects of frictions, at least according to the model.

[8] It is conceivable that technology actually evolves according to the same AR(1) process as output growth. But even if this is so, there is no sense in which the behavior of technology is being explained by the endogenous economic mechanisms of the model (i.e., capital accumulation). Thus, in essence, the model provides little or no additional insight into the behavior of real GDP beyond that of a univariate AR(1) time-series model for output growth.

[9] As an early example of expressed concern about relying on serially-correlated shocks, consider the following discussion from a classic paper in the RBC literature by Robert King, Charles Plosser, and Sergio Rebelo: "But along other dimensions, the basic model seems less satisfactory. In particular, the principle serial correlation in output – one notable feature of economic fluctuations – derives mainly from the persistence of technology shocks." See King, R.G., C.I. Plosser, and S.T. Rebelo, 1988, "Production, Growth, and Business Cycles: I. The Basic Neoclassical Model," *Journal of Monetary Economics* 21, 195-232.

The example of a state-of-the-art DSGE model mentioned in Kocherlakota's essay on modern macroeconomics is a "New Keynesian" model of the European economy developed by Frank Smets and Raf Wouters.[10] On one level, their model is much more impressive than the RBC model in the preceding section. It has seven observable variables and ten different types of structural shocks, not all of which are serially correlated. The model incorporates sticky prices, sticky wages, habit formation, costs of adjustment in capital accumulation, and variable capacity utilization. Perhaps most impressively, Smets and Wouters estimate rather than calibrate their model, which had been the standard practice in the RBC literature.

Much has been made of the fact that Smets and Wouters' estimated DSGE model appears to forecast macroeconomic data almost as well as a "vector autoregressive" (VAR) model. But before getting too carried away, there are a few issues that should be considered.

First, like the RBC model in the preceding section, Smets and Wouters' model is solved by linearizing around steady state. This might sound like a minor technical detail, but it has a direct practical consequence that the model needs to be estimated using deviations from steady-state levels of variables such as real GDP. Also, the predictions of the model are for the deviations from steady state rather than the levels of the variables themselves.

Smets and Wouters "solve" the problem of measuring deviations from steady state by considering deviations of variables from their sample means or, in some cases, from estimated linear time trends. Other papers in this literature consider Hodrick-Prescott (HP) or Band-pass (BP) filtered data. From an econometric point of view, the possible presence of unit roots (a.k.a. "stochastic trends") in macroeconomic variables raises strong concerns that these procedures will lead to inconsistent parameter estimates and inaccurate out-of-sample forecasts, even in the unlikely event that the theoretical model is correctly specified.[11]

Second, even ignoring the statistical issues involved in measuring deviations from steady state, there is still the underlying economic problem that Smets and Wouters' model relies on serially-correlated shocks for productivity, the inflation target, consumer preferences, government spending, labor supply, and investment. All of these shocks are assumed to follow AR(1) processes, with posterior mean estimates of the AR coefficients ranging from 0.81 to 0.94 for the demeaned and linearly-detrended quarterly data.

The DSGE model places restrictions on the mapping of the serially-correlated shocks to the observable variables. However, it is not clear what role the shocks play in "explaining" the

[10] Frank Smets and Raf Wouters, 2003, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1, 1123–75, [pdf].

[11] Without taking too long of a digression into the literature on trend/cycle decomposition and spurious cycles, it is worth mentioning another paper by Timothy Cogley and James Nason. They consider the implications of HP filtering for RBC models, but their conclusions apply equally well for linear detrending and BP filtering for DSGE models: "When applied to integrated processes, the HP filter can generate business cycle periodicity and comovement even if none are present in the original data…We show that RBC models can exhibit business cycle dynamics in HP filtered data even if they do not generate business cycle dynamics in pre-filtered data. The combination of a unit root or near unit root in technology and the HP filter is sufficient to generate business cycle dynamics. Propagation mechanisms are unnecessary, and in many RBC models they do not play an important role." See Cogley, T. and J.M. Nason, 1995b, "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control* 19, 253-278.

observable variables relative to the endogenous economic mechanisms. Indeed, at one point in their paper, Smets and Wouters ponder as follows: "[I]t would be interesting to see which of the various frictions are crucial for capturing the persistence and covariances in the data". However, what they neglect to mention is that the answer could very well be "none". In reality, it is quite possible the persistence and covariances in the data are being captured by what amounts to a statistical factor model. This could be tested by comparing the forecasting ability of the DSGE model to a parsimonious "atheoretical" dynamic factor model. Instead, the focus is on a comparison to VAR models. But VAR models have many parameters (especially when there are large number of variables) and, as acknowledged by Smets and Wouters, may be suffering for their heavy-parameterization given the short sample period for European data.[12]

Meanwhile, Smets and Wouters readily acknowledge that some of their shocks may be serving as proxies for important, omitted macroeconomic phenomena: "Of course, these shocks could capture a whole range of shocks that are not accounted for in the stylised model such as changes in oil prices, terms-of-trade shocks, changes in taxes, etc." This simply begs the question of why not consider models, such as large-scale macroeconometric models, that directly incorporate such variables.

The Lucas Critique and Large-Scale Macroeconometric Models

Before discussing the details of large-scale macroeconometric models, it is perhaps useful to revisit Kocherlakota's essay on modern macroeconomics. He writes:

> *In terms of fiscal policy (especially short-term fiscal policy), modern macro modeling seems to have had little impact. The discussion about the fiscal stimulus in January 2009 is highly revealing along these lines. An argument certainly could be made for the stimulus plan using the logic of New Keynesian or heterogeneous agent models. However, most, if not all, of the motivation for the fiscal stimulus was based largely on the long-discarded models of the 1960s and 1970s.*

The "long-discarded models" that Kocherlakota has in mind are large-scale macroeconometric models and the natural question that arises when reading his last sentence is "discarded by whom?" Certainly, it is not by policymakers trying to predict the quantitative effects of fiscal stimulus.

The answer to "discarded by whom?" that Kocherlakota clearly has in mind is "modern" macroeconomists, by which he means academics who take the Lucas critique as the central tenet of macroeconomics. Here again Kocherlakota writes:

> *The macro models used in the 1960s and 1970s were based on large numbers of interlocking demand and supply relationships estimated using various kinds of data. In his powerful*

[12] There is a more subtle and technical issue with the model comparison. Smets and Wouters consider Bayesian estimation and use "Bayes factors" to evaluate the relative forecasting ability of models. Bayes factors are, by construction, sensitive to priors. On the one hand, macroeconomic theory provides a sensible way to formulate priors about the parameters of the model. On the other hand, if priors are being informed by the data, the exercise is one of so-called "empirical Bayes", which is highly problematic (by putting different weights on data than what is implied by Bayes rule). In the literature on estimated DSGE models, the priors are often influenced by what has been found in other studies that consider the same data. This will tend to bias results in favor of DSGE models relative to models with less informative priors.

critique, Lucas demonstrated that the demand and supply relationships estimated using data generated from one macroeconomic policy regime would necessarily change when the policy regime changed. Hence, such estimated relationships, while useful for forecasting when the macro policy regime was kept fixed, could not be of use in evaluating the impact of policy regime changes.

An immediately noticeable thing about this quote is the deliberate use of the words "policy regime" rather than just "policy". For example, consider monetary policy. This distinction would be between a change in the Federal Funds Rate (which is a change in "policy") versus a change in the coefficients for a Taylor-type rule (which is a change in "policy regime"). The choice of words is relevant because the previous quote about short-term fiscal policy would seem to be about a change in policy, not necessarily a change in policy regime. Thus, at least according to Kocherlakota's qualification about forecasting in a fixed regime, the estimated relationships considered in large-scale macroeconometric models may well be useful for predicting the effects of the stimulus plan.

Another noticeable thing about Kocherlakota's discussion of the Lucas critique is that he presents it as some sort of universal truth that estimated demand and supply relationships will be unstable and of limited use for policy analysis. This might be valid if a DSGE model were reality. But DSGE models are models, not reality. Thus, the relevance of the Lucas critique is testable and the tests have not been favorable.[13] Meanwhile, according to a meta-critique of the Lucas critique by Christopher Sims, the lack of practical relevance should come as no surprise.[14] He writes:

The only coherent interpretation of the Lucas critique is that it states that if one uses a model which incorrectly describes the reaction of expectations formation to policy choice, it will produce incorrect evaluations of policy. The implication is not that econometric evaluation of policy using models fitted to history is impossible, but that it requires correct specification of the reaction of the economy to policy… There may be some policy issues where the simple rational expectations policy analysis paradigm – treating policy as given by a rule with deterministic parameters, which are to be changed once and for all, with no one knowing beforehand that the change may occur and no one doubting afterward that the change is permanent – is a useful approximate simplifying assumption. To the extent that the rational expectations literature has led us to suppose that all "real" policy change must fit into this internally inconsistent mold, it is has led us onto sterile ground.

[13] See, for example, Favero, C. and D. Hendry, 1992, "Testing the Lucas Critique: A Review," *Econometric Reviews* 11, 265-306. Meanwhile, the implied change in demand and supply relationships given a change in policy regime may not even occur within DSGE models when allowing for indeterminate equilibria. See Farmer, R., 2003, "Why Does Data Reject the Lucas Critique?" *Annales d'économie et de statistique* special issue on "The Econometrics of Policy Evaluation", 67-68, [pdf]. As for the economic relevance of the Lucas critique under the assumption that a DSGE model is reality, there are different views. Glenn Rudebusch argues that the Lucas critique does not invalidate the use of reduced-form models to forecast aggregate output and inflation, while Thomas Lubik and Paolo Surico argue that it is more relevant when considering the effects of policy regimes on macroeconomic volatility. See Rudebusch, G.D., 2005, "Assessing the Lucas Critique in Monetary Policy Models," *Journal of Money, Credit, and Banking* 37, 245-272, [pdf] and Lubik, T.A. and P. Surico, 2010, "The Lucas Critique and the Stability of Empirical Models," *Journal of Applied Econometrics* 25, 177-199, [pdf].

[14] Sims, C., 1987, "A Rational Expectations Framework for Short Run Policy Analysis," in New Approaches to Monetary Economics, W. Barnett and K. Singleton, eds., Cambridge University Press, 293-310, [pdf].

Thus, contrary to the precepts of "modern" macroeconomics, the Lucas critique in no way proves that DSGE models will predict the effects of policy better than large-scale macroeconometric models based more on historical correlations.

However, to avoid getting lost in a long, fruitless debate over the semantics of the Lucas critique, it is perhaps more constructive to simply review the features of modern macro models that Kocherlakota argues have been inspired by it. In his words, modern macro models have the five following properties:

- *They specify budget constraints for households, technologies for firms, and resource constraints for the overall economy.*
- *They specify household preferences and firm objectives.*
- *They assume forward-looking behavior for firms and households.*
- *They include the shocks that firms and households face.*
- *They are models of the entire macroeconomy.*

The thing that is notable about this list is that these features are far from the sole domain of DSGE models. They are also present in contemporary versions of the large-scale macroeconometric models of the U.S. economy such as those developed by Macroeconomic Advisers, the Federal Reserve Board (FRB/US), and the Bank of Canada (MUSE). In this sense, it might well be reasonable to acknowledge the impact of some interpretations of the Lucas critique on macroeconomics. However, it is quite a leap to go from this acknowledgment to discarding all but DSGE models, as Kocherlakota would seem to have us do.

So what about contemporary large-scale macroeconometric models? They have a number of advantages over DSGE models. First, there are many more variables in the models. This allows for more useful details that are typically missing from DSGE models, from "small" things like the consideration of different types of consumption (e.g., durables vs. nondurables and services) and different forms of fiscal policy (i.e., more than just lump-sum transfers), to larger things like foreign trade.[15] Second, the models consider levels data rather than deviations from steady state. This is helpful for statistical and economic identification, as well as for forecasting. Third, the models are grounded in macroeconomic theory, but they are not intended to be a literal description of reality.

An example might help illustrate the nature of large-scale macroeconometric models. The model developed by Macroeconomic Advisers is based in part on the life-cycle hypothesis in which households are forward-looking and smooth their consumption across their lifetime income profiles. This theoretical setting implies a consumption function in which the marginal propensities to consume can be thought of as complicated functions of "deep

[15] There, of course, exist DSGE models that incorporate some of these additional details found in large-scale macroeconometric models. But for reasons of computational tractability, they never come close to incorporating all of the details at the same time. As for the notion that DSGE models can, in principle, be made as complicated as necessary to capture reality, there is a problem that the microfoundations are likely to be so complicated that the models would never be operational for forecasting or policy analysis. For example, the large disconnect between the complicated micro-founded models on the existence of money and the practice of monetary policy comes to mind. Indeed, by focusing on a micro-founded "medium-of-exchange" role for money, the models may be reducing rather than increasing their links to reality. See Goodhart, C.A.E., 2009, "The Continuing Muddles of Monetary Theory: A Steadfast Refusal to Face Facts," *Economica* 76, 821-830, [pdf].

structural parameters".[16] A key point is that the estimates for the marginal propensities to consume are remarkably stable over the postwar period, implying that the deep structural parameters for this model are also fairly stable.

The Macroeconomic Advisers model is estimated using the directly observable data rather than deviations from steady state. One possibly surprising benefit of this is that estimation of the model, which takes unit roots and cointegration into account, makes use of permanent variation to help identify parameter values. Also, exogeneity restrictions which might be debatable for short-run fluctuations are arguably more reasonable in terms of long-run variation. Thus, long-run variation helps with both statistical and economic identification.

A more obvious benefit of using the levels data is that the Macroeconomic Advisers model provides direct forecasts of the data rather than deviations from steady state. Meanwhile, the model allows for relatively flexible short-run dynamics. Thus, the macroeconomic theory pinning down long-run relationships and some of the adjustment dynamics provides a guide to explaining the data, but theory is not taken quite so literally as it is in estimated DSGE models.[17]

As a key concession to reality, the Macroeconomic Advisers model allows for residuals. Consequently, when faced with the Great Recession, the model did not explain 100% of the movements in real GDP. Although the consumption function in the model implied a decline in economic activity in response to the dramatic declines in household wealth associated with the collapse of the housing market and then financial markets, the model — along with pretty much every other pre-existing quantitative model — did not quite predict the severity of the recession. But this is surely better than pretending that the model explains 100% of real economic activity by relying on serially-correlated technology and preference shocks.

Indeed, the Macroeconomic Advisers model can capture the effects of wealth on consumption more accurately *because* it does not assume that the model must explain 100% of real economic activity. The residuals act as a kind of "safety pressure relief valve" to address the fact that models are not reality. The notion that estimated DSGE models are too literal corresponds to the idea that they have no such safety valve.[18] A consequence is that, as discussed in more detail below, estimates of the supposed deep structural parameters for DSGE models can actually be fairly sensitive over time, including to changes in policy regimes.

---

[16] Prakken, J., 2002, "Deep Parameters and the Consumption Function," Technical Note, *Macroeconomic Advisers*. In addition to depending on the interest and tax rates, which are accounted for directly in the estimated consumption function, the marginal propensities to consume also depend on the rate of time preference, both the inter- and intra-temporal substitution elasticities, the intensity factors in the utility function, the expected rate of growth in the consumer's real wage, the age of the consumer, life expectancy, and retirement age.

[17] Somewhat ironically, the literalism of estimated DSGE models for short-run dynamics does not generally extend to long-run relationships. In particular, by estimating linear time trends or filtering the data prior to estimation of the log-linearized system, the implied long-run relationships between the levels data can be contrary to any cointegration relationships implied by theory.

[18] Some estimated DSGE models allow for "measurement error". However, Smets and Wouters celebrate the fact that their model does not rely on such shocks to "explain" the data (while, of course, measurement error does not explain actual economic activity). Meanwhile, simply interpreting the technology and other shocks broadly does not serve the same purpose as including residuals. Again, as discussed in footnote 4, this is because a model forces the same endogenous reaction to all shocks given the same label. Thus, estimates of the endogenous reaction to the true more-narrowly-defined shocks will be distorted.

So, returning to Kocherlakota's requirements for a "modern macro model", the treatment of households in the Macroeconomic Advisers model i) includes intertemporal budget constraints, ii) specifies preferences for households that appear to be stable, iii) assumes households are forward looking, iv) includes realistic shocks that households face such as different types of fiscal interventions, and v) is part of a model of the entire economy. Evidently, even though the model differs from DSGE models in terms of certain assumptions and general implementation, it has a claim to being "modern" by Kocherlakota's criteria.[19]

Can't We All Just Get Along?

In some sense, the various approaches to macroeconomics are not really as different as they are sometimes made out to be. Although Finn Kydland and Edward Prescott, the Nobel Prize-winning gurus of the RBC camp, once wrote dismissively of a "system-of-equations" approach,[20] the reality is that VARs, large-scale macroeconometric models, and DSGE models all imply systems of equations. Policy forecasts for these different approaches are all based on some assumptions from macroeconomic theory and some consideration of how economic agents perceive a given change in policy — i.e., was it anticipated or unanticipated and will it be permanent or transitory? The main differences across approaches are in terms of how estimation is carried out and how the theoretical assumptions are imposed. The VAR places the least (but not zero) weight on theory, while the DSGE models place the most, even to the extent of imposing strong restrictions on some parameters across equations. It is ultimately an empirical question as to whether the imposition of these cross-equation restrictions really helps with predicting the effects of policy and forecasting more generally.[21]

An important issue, related to the Lucas critique, is the stability of estimates over time. For any model, parameter estimates will sometimes change, some more than others. For example, time-series models of real GDP growth typically have fairly stable estimates for parameters related to dynamics, but variance estimates have changed greatly since the onset of Great Moderation in the mid-1980s. However, the possibility of time-varying parameters hardly invalidates the use of a given model for forecasting or policy analysis. It merely begs the question of how adept estimates for that model are at tracking parameters that can generate accurate predictions in real time. In macroeconomics, parameter stability is a relative concept, not an absolute.

---

[19] The Macroeconomic Advisers model also implicitly assumes that firms are forward-looking in their investment decisions. Of course, this distinction between implicit and explicit assumptions, which is related to how literally the theory is taken, is the most essential difference from the DSGE framework. Meanwhile, the Phillips curve in the model makes use of survey measures of inflation expectations. This is forward-looking, but does not require identification via rational expectations in the sense of economic agents — unlike lowly econometricians — knowing the exact structure of the economy, including all parameter values, and observing all variables and shocks in real time. This assumption about expectations is also an important difference from DSGE models.

[20] Kydland, F.E. and E.C. Prescott, 1990, "The Econometrics of the General Equilibrium Approach to Business Cycles," *Federal Reserve Bank of Minneapolis Staff Report*, no. 130, [pdf].

[21] There is an old argument dating back (at least) to the Cowles Commission in the 1940s and 1950s that if different models fit the data equally well, their relative usefulness for policy evaluation can only be determined on *a priori* theoretical grounds. However, even if models are observationally equivalent within a given sample period, the very fact that they can produce different policy predictions implies that they will have different out-of-sample forecasting performances (at least assuming the forecasters can observe the policy interventions). Thus, an empirical evaluation of the usefulness of a model for predicting the effects of policy is possible by tracking its real-time out-of-sample forecasting performance over periods in which policy has changed.

As an example, one of the most prominent controversies throughout the history of macroeconomics has been over the stability of the Phillips curve. This was certainly the example from Robert Lucas's original critique that stuck in the collective consciousness of academic macroeconomists. One reason is that estimates of Phillips curve parameters have indeed changed over time. However, the notable thing is that this is also true about the estimates of the supposed "deep structural parameters" of DSGE models that generate a Phillips curve. For example, Luca Benati finds that estimates of a key structural parameter that determines the New Keynesian Phillips curve is not at all robust across different monetary policy regimes.[22] Meanwhile, instabilities in Phillips curve parameters do not appear to translate into instabilities in other macroeconomic correlations. It might be argued that this is due to low statistical power. But, contrary to this view, it is notable, for example, just how robust point estimates for consumption functions are to different sample periods. Thus, macroeconomists should pursue stable relationships as much as possible, but they should not assume *a priori* that a DSGE model is where they will find them. In the words of Robert Lucas and Thomas Sargent, "[T]he question of whether a particular model is structural is an empirical, not theoretical, one."[23]

Doing Better in the Future

It is a safe bet that future versions of DSGE models will incorporate more complicated financial sectors and allow for different types of fiscal policies. And guess what? The new-and-improved DSGE models will turn out to imply (*ex post*) that the Great Recession was actually due to serially-correlated financial intermediation shocks and suboptimal fiscal policy.

Alas, these conclusions will be driven much more by the DSGE framework than by the data. In general, the implications of DSGE models for policy are more assumed than estimated. For example, the typical assumption is that households are "Ricardian" in the sense that they are infinitely-lived and anticipate that the government's intertemporal budget constraint will always hold. Thus, households will perceive a current increase in government spending as being offset by a future increase in taxes. This clearly places strong restrictions on estimates of fiscal multipliers — i.e., by construction, everything is a "balanced-budget" multiplier. In some recent DSGE models, a large fraction of households are assumed to be "hand-to-mouth" in the sense that they consume all of their disposable income immediately. Although this assumption represents somewhat of a departure from the original intentions of the microfoundations agenda, it appears to be necessary to replicate less restrictive estimates of fiscal multipliers from VAR models.[24]

[22] Benati, L., 2008, "Investigating Inflation Persistence across Monetary Regimes," *Quarterly Journal of Economics* 123, 1005-1060, [pdf]. The parameter is the "indexation" parameter for Calvo-style pricing that corresponds to the fraction of firms that cannot re-optimize prices in a given time period. Also see Glenn Rudebusch's analysis of the Lucas critique mentioned in footnote 13 in which he discusses some other models with supposed deep structural parameters that have turned out to be unstable.

[23] Lucas, R.E. and T. Sargent, 1981, "After Keynesian Macroeconomics," in Rational Expectations and Econometric Practice, University of Minnesota Press, 295-319. This quote is, of course, somewhat disingenuous. Lucas and Sargent emphasize that success in terms of fitting the data or general short-term forecasting is no guarantee of success at forecasting conditional on a change in the policy environment. They clearly see the failures of large-scale macroeconometric models in the 1970s as empirical validation of this point and strongly believe micro-founded models will provide stable parameters. But their point about the empirical nature of whether a model is structural has often been forgotten in modern macroeconomics.

[24] See Galí, J., J.D. Lopez-Salido, and J. Valles, 2007, "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association* 5, 227-270, [pdf].

However, rather than just sticking to the same basic framework with relatively small changes in assumptions to help fit the data, modern macroeconomists should also focus on a more fundamental issue: a "shock" should be a surprise — i.e., a given shock process should really be serially uncorrelated within a fully-specified theory. At the very least, the analysis of DSGE models should be geared much more towards convincing a skeptic that results are being driven by endogenous economic mechanisms that are consistent with data, not by assumed exogenous processes or, more subtly, by an absence of residuals that could help cope with model misspecification of the real world.

In general, promoters of DSGE models need to convince non-believers that estimates are robust across policy regimes in the sense of producing better forecasts than other models in changing policy environments. Although it is certainly true that estimated models can "over-fit" in sample, out-of-sample comparisons address this problem. Indeed, even if models happen to fit equally well in sample, they will, if based on different economic assumptions, produce different out-of-sample predictions for the effects of policy. The bottom line then is that DSGE models should be subject to the same (market-based) forms of evaluation that large-scale macroeconometric and other forecasting models have been subject to — i.e., they need to forecast well in real time.

Meanwhile, it should be acknowledged that, despite including oil prices and credit conditions, large-scale macroeconometric models didn't exactly predict the severity of the Great Recession of 2007–2009. There were still sequences of negative forecast errors for various measures of real economic activity. More consideration of nonlinear time-series dynamics could potentially help on this front.[25] However, in terms of really predicting the crisis, the award obviously goes to theories of endogenous financial crises inspired by the ideas of Hyman Minsky. Formal evaluation of these more narrative approaches is hard and there may be an element of the "stopped-clock syndrome" at play. But it would be foolish to dismiss such theories out of hand. In particular, a ludicrous notion sometimes expressed in the ivory towers of academia is that, for Minsky to be taken seriously, his ideas need to be put into a DSGE model.[26] Instead, the converse is true. For DSGE models to be taken more seriously outside of academia, they need to explain and predict as well as Minsky. And serially-correlated preference and technology shocks aren't going to do it!

To be critical of the Lucas critique is not to say it is completely irrelevant. An important goal of macroeconomic models is to have stable parameters given changes in the policy environment. But how we get there is not necessarily through a particular class of microfounded models. More broadly, if macroeconomists want to regain the trust of the public at large, they need to resist the notion that "macroeconomics" is defined as a method, rather

[25] A young macroeconomist who shall remain nameless once responded to evidence of nonlinearity in the form of business cycle asymmetry by saying (without any hint of embarrassment) that a DSGE model could easily explain such a phenomenon — all that was necessary was to assume the appropriate nonlinear shock process!

[26] Consider the following claim by V.V. Chari, Partrick Kehoe, and Ellen McGratten: "[M]ost macroeconomists now analyze policy using some sort of dynamic stochastic general equilibrium (DSGE) model. This type of model can be so generally defined that it incorporates all types of frictions, such as various ways of learning, incomplete markets, imperfections in markets, and spatial frictions. The model's only practical restriction is that it specify an agreed-upon language by which we can communicate, a restriction hard to argue with. An aphorism among macroeconomists today is that if you have a coherent story to propose, you can do it in a suitably elaborate DSGE model." See Chari, V.V., P.J. Kehoe, and E.R. McGratten, 2008, "New Keynesian Models: Not Yet Useful for Policy Analysis," *Federal Reserve Bank of Minneapolis Staff Report*, no. 409, [pdf].

than as a subject matter. Specifically, macroeconomists need to be more pluralistic. They should draw from different types of analysis, be it time-series models, large-scale macro-econometric models, DSGE models, and more narrative approaches. Ultimately, we will know that the reputation of macroeconomics has been rehabilitated when "modern macro-economics" is no longer used as a label for a particular school of thought, but instead refers to a body of knowledge of substantive and useful insights into how the macroeconomy actually works and what will happen to it in the future.