The Past, Present, and Future of Macroeconomic Forecasting

Francis X. Diebold

The reports of the death of large-scale macroeconomic forecasting models are not exaggerated. But many observers interpret the failure of the early models as indicative of a bleak future for macroeconomic forecasting more generally. Such is not the case. Although the large-scale macroeconomic forecasting models didn’t live up to their original promise, they nevertheless left a useful legacy of lasting contributions from which macroeconomic forecasting will continue to benefit: they spurred the development of powerful identification and estimation theory, computational and simulation techniques, comprehensive machine-readable macroeconomic databases, and much else. Moreover, past failures do not necessarily imply a bleak future: we learn from our mistakes. Just as macroeconomics has benefitted from rethinking since the 1970s, so too will macroeconomic forecasting.

Understanding the future of macroeconomic forecasting requires understanding the interplay between measurement and theory, and the corresponding evolution of the nonstructural and structural approaches to forecasting. Nonstructural macroeconomic forecasting methods attempt to exploit the reduced-form correlations in observed macroeconomic time series, with little reliance on economic theory. Structural models, in contrast, view and interpret economic data through the lens of a particular economic theory.

Structural econometric forecasting, because it is based on explicit theory, rises and falls with theory, typically with a lag. Structural Keynesian macroeconomic forecasting, based on postulated systems of decision rules, enjoyed a golden age in the 1950s and 1960s following the advances in Keynesian theory in the 1930s and 1940s,

and the two then declined together in the 1970s and 1980s. The evolution of non-structural economic forecasting, in contrast, is less bound to fashions in economic theory; its origins long predate structural Keynesian macroeconomic forecasting and progress continues at a rapid pace.

One is naturally led to a number of important questions. What of the impressive advances in economic theory of the 1980s and 1990s? Should we not expect them to be followed by a new wave of structural macroeconomic forecasting, or has nonstructural forecasting permanently replaced structural forecasting? Is it necessary to choose between the structural and nonstructural approaches, or might the two be complements rather than substitutes? If a new structural forecasting is likely to emerge, in what ways will it resemble its ancestors? Our answers will take us on a whirlwind tour of the past, present and future of both structural and nonstructural forecasting. We'll begin by tracing the rise and fall of the structural Keynesian system-of-equations paradigm, and then we'll step back to assess the long-running and ongoing progress in the nonstructural tradition. Finally, we'll assess the rise of modern dynamic stochastic general equilibrium macroeconomic theory, its relationship to nonstructural methods, and its implications for a new structural macroeconomic forecasting.

The Rise and Fall of Keynesian Macroeconomic Theory and Structural Forecasting

Some important forecasting situations involve conditional forecasts; that is, forecasts of one or more variables conditional upon maintained assumptions regarding, for example, the behavior of policymakers. Conditional forecasts require structural models. Structural econometrics, and hence structural macroeconomic forecasting, makes use of macroeconomic theory, which implies that developments in structural forecasting naturally lag behind developments in theory. The first major wave of twentieth century macroeconomic theory was the Keynesian theory of the 1930s and 1940s, and it was followed by major advances in structural macroeconomic forecasting.

When Keynes's *General Theory* was published in 1936, theory was distinctly ahead of measurement. Measurement soon caught up, however, in the form of the systems of equations associated with Klein's (1946) *Keynesian Revolution* and Klein and Goldberger's (1955) *Econometric Model of the United States: 1929–1952*. Indeed, the period following the publication of the *General Theory* was one of unprecedented and furious intellectual activity directed toward the construction, estimation and analysis of Keynesian structural econometric models. The statistics side of the structural econometrics research was fueled by the advances of Fisher, Neyman, Pearson, and many others earlier in the century. The economics side, of course, was driven by Keynes's landmark contribution, which spoke eloquently to the severe economic problems of the day and seemed to offer a workable solution.

The intellectual marriage of statistics and economic theory was reflected in the
growth of the Econometric Society and its journal, *Econometrica*, and beautifully distilled in the work of the Cowles Commission for Research in Economics at the University of Chicago in the 1940s and early 1950s. The intellectual focus and depth of talent assembled there were unprecedented in the history of economics: Cowles researchers included T.W. Anderson, K. Arrow, G. Debreu, T. Haavelmo, L. Hurwicz, L.R. Klein, T. Koopmans, H. Markowitz, J. Marshak, F. Modigliani, H. Simon, A. Wald, and many others. A central part (although by no means the only part) of the Cowles research program was identification and estimation of systems of stochastic difference equations designed to approximate the postulated decision rules of Keynesian macroeconomic theory.

Just as the blending of mathematical statistics and economics associated with the Cowles Commission was historically unprecedented, so too was the optimism about solving pressing macroeconomic problems. Early on, the macroeconomic system-of-equations research program appeared impressively successful, and structural econometric forecasting blossomed in the late 1950s and 1960s, the heyday of the large-scale Keynesian macroeconomic forecasting models. There was strong consensus regarding the general paradigm, even if there was disagreement on details such as the relative slopes of IS and LM curves, and the models were routinely used for forecasting and policy analysis in both academia and government.

But cracks in the foundation, which began with intellectual dissatisfaction with the underpinnings of Keynesian macroeconomic systems of equations, started to appear in the late 1960s and early 1970s. First, economists became dissatisfied with the lack of foundations for the disequilibrium nature of the Keynesian model. A new and still ongoing research program began which sought microfoundations for Keynesian macroeconomic theory, particularly for the central tenets of sticky wages and prices. Many key early contributions appear in the classic Phelps et al. (1970) volume, and more recent contributions are collected in Mankiw and Romer (1991).

Second, just as macroeconomists became increasingly disenchanted with the ad hoc treatment of sticky prices in traditional models, they became similarly disenchanted with ad hoc treatment of expectations. Building on early work by Muth (1960, 1961), who introduced the idea of rational expectations and showed that schemes such as adaptive expectations were rational only in unlikely circumstances, the "rational expectations revolution" quickly took hold; Sargent and Wallace (1975) is an important and starkly simple early paper.

Third, and most generally, economists became dissatisfied not only with certain parts of the Keynesian macro-econometric program, such as the assumptions about price behavior and expectations formation, but rather with the overall modeling approach embodied in the program. The approach was dubbed the "system-of-equations" approach by Prescott (1986), in reference to the fact that it concentrated on the estimation of parameters of equation systems representing ad hoc postulated decision rules ("consumption functions," "investment functions," and so on) as opposed to more fundamental parameters of tastes and technology. Newly

---

1 For a concise history of the Chicago days of the Cowles Commission, see Hildreth (1986, ch. 1).
emerging macroeconomic work in the late 1960s and early 1970s, in contrast, was firmly grounded in tastes and technology; Lucas and Prescott (1971) and Lucas (1972) remain classic examples. Work in the tastes-and-technology tradition accelerated rapidly following Lucas’s (1976) formal critique of the system-of-equations approach, based on the insight that analysis based on decision rules is a fundamentally defective paradigm for producing conditional forecasts, because the parameters of decision rules will generally change when policies change.

Finally, if the cracks in the foundation of Keynesian structural forecasting began as intellectual dissatisfaction, they were widened by the economic facts of the 1970s; in particular, the simultaneous presence of high inflation and unemployment, which naturally led economists to question the alleged inflation/unemployment tradeoff embedded in the Keynesian systems of equations. In addition, a series of studies published in the early 1970s revealed that simple statistical extrapolations, making no assumptions at all about economic structure, often forecasted macroeconomic activity just as well as large-scale Keynesian macroeconomic models; Nelson (1972) remains a classic. Keynesian macroeconomics soon declined, and Keynesian structural econometric forecasting followed suit.

**Nonstructural Forecasting**

By the late 1970s, it was clear that Keynesian structural macroeconomic forecasting, at least as traditionally implemented, was receding. One response was to augment the traditional system-of-equations econometrics in attempts to remedy its defects. Important work along those lines was undertaken by R. Fair and J. Taylor (for example, Fair, 1984, 1994; Taylor, 1993), who developed methods for incorporating rational-expectations into econometric models, as well as methods for rigorous assessment of model fit and forecasting performance. Models in the Fair-Taylor spirit are now in use at a number of leading policy organizations, including the Federal Reserve Board and the International Monetary Fund, as described, for example, in Brayton et al. (1997). They are an important step forward, even if the theory on which they are built remains largely in the system-of-equations tradition.

Another response, involving a more radical change of direction, was to explore alternative, nonstructural forecasting methods. Many forecasting chores involve unconditional, rather than conditional, forecasts—that is, interest often centers on the likely future path of the economy when policy remains unchanged, so that the Lucas critique is not relevant—and unconditional forecasting does not require a structural model. That insight, together with the emerging discontent with Keynesian macroeconomic theory and the lack of a well-developed alternative, produced tremendous interest in nonstructural econometric forecasting in the 1970s. The title of an important paper by Sargent and Sims (1977), “Business Cycle Modeling Without Pretending to Have too Much a Priori Theory,” nicely summarizes the spirit of the times.

The impressive intellectual development of nonstructural forecasting spans many decades; it predates the Keynesian episode and continues to the present.
Macroeconomists and econometricians didn’t pay much attention at first, in spite of the fact that key early contributions were made by economists; they were too busy with Keynesian theory and Keynesian structural econometrics. Nevertheless, rapid development took place in the hands of some of the most talented mathematicians, statisticians and engineers of the 20th century.

Let us begin our account in the 1920s, a period of fertile intellectual development in nonstructural modeling and forecasting. Many ideas were hatched and nurtured, and the groundwork was laid for the impressive technical advances of the ensuing decades. In particular, Slutsky (1927) and Yule (1927) argued that simple linear difference equations, driven by purely random stochastic shocks, provide a convenient and powerful framework for modeling and forecasting a variety of economic and financial time series. Such stochastic difference equations are called autoregressive processes, or autoregressions. They amount to regression models in which the current value of a variable is expressed as a weighted average of its own past values, plus a random shock. Autoregressive processes are closely related to moving average processes, also studied by Slutsky and Yule, in which the current value of a variable is expressed as a weighted average of current and lagged random shocks alone. In fact, under reasonable conditions, one can convert an autoregressive process to a moving average process, and conversely. Either way, the key insight is that system dynamics can convert random inputs into serially correlated outputs, a phenomenon often called the Slutsky-Yule effect. Frisch (1933) put the Slutsky-Yule insight to work in formulating the idea of “impulse” and “propagation” mechanisms in economic dynamics.

In the 1930s, the mathematician-turned-economist H. Wold made a stunning contribution, paving the way for later work by the mathematicians N. Wiener and A. Kolmogorov, and the engineer R. Kalman. Wold showed that, given sufficient stability of the underlying probabilistic mechanism generating a time series, its stochastic part can be represented as a model of the form studied by Slutsky and Yule. Thus, the Slutsky-Yule models are not only convenient and powerful, they are absolutely central—they’re the only game in town. Wiener and Kolmogorov worked out the mathematical formulae for optimal forecasts from models of the type studied by Slutsky, Yule, and Wold. Kalman extended the theory in the late 1950s and early 1960s by relaxing some of the conditions that Wiener and Kolmogorov required; his forecasting formula, known as the Kalman filter, is designed to work with a powerful model representation known as state-space form and has a convenient recursive form amenable to real-time forecasting. The Wold-Wiener-Kolmogorov-Kalman theory, which effectively represents the pinnacle of the Slutsky-Yule research program, is beautifully exposited in Whittle (1963, second edition 1983). Appropriately enough, a leading economist, T. Sargent, wrote the second edition’s introduction, which catalogs the tremendous impact of the prediction theory on modern dynamic economics.

See Harvey (1989) for extensive discussion of state space representations and the Kalman filter in relation to forecasting.
In part, the nonstructural econometric forecasting explosion of the 1970s was driven by econometricians absorbing the powerful earlier advances made by the likes of Wold, Wiener, Kolmogorov and Kalman. But there was a major additional push: in 1970, just as discontent with Keynesian structural econometric forecasting was beginning to emerge, Box and Jenkins (1970; third edition, Box, Jenkins and Reinsel, 1994) published a landmark book on nonstructural time series analysis and forecasting.

Many of the Box-Jenkins insights started literatures that grew explosively. For example, prior to Box and Jenkins, trend was typically modeled as a simple linear deterministic function of time, whereas Box and Jenkins allowed trend to be driven by the cumulative effects of random shocks, resulting in “stochastic trend.” Stock and Watson (1988a) provide an insightful discussion of stochastic trend and its wide-ranging implications. Shocks to series with stochastic trend have permanent effects, an idea amplified in the empirical macroeconomics literature associated with Nelson and Plosser (1982) and Campbell and Mankiw (1987), among others. The direct implication for forecasting is that long-run forecasts fail to revert to any fixed trend; effectively, the underlying trend location is redefined each period, as emphasized, for example, in Diebold and Senhadji (1996).

The most important contribution of Box and Jenkins, however, is their sweeping vision, articulation, and illustration of an operational framework for applied nonstructural forecasting, consisting of iterative cycles of model formulation, estimation, diagnostic checking, and forecasting. Autoregressive moving average (ARMA) models are the centerpiece of the Box-Jenkins framework. ARMA models are combinations of the autoregressive and moving average models of Slutsky and Yule, and they have the potential to approximate dynamics more parsimoniously than purely autoregressive or moving average models.

An ongoing flood of work followed Box and Jenkins. Macroeconomics, in particular, is crucially concerned with cross-variable relationships, whereas the basic approach of Box and Jenkins uses only the past of a given economic variable to forecast its future. In other words, much of macroeconomics is concerned with multivariate relationships, whereas the basic Box-Jenkins models are univariate. Thus, many extensions of the Box-Jenkins program involve multivariate modeling and forecasting, and vector autoregressive models have emerged as the central multivariate model. Vector autoregressions were forcefully advocated in econometrics by Sims (1980) as a less restrictive alternative to traditional econometric systems-of-equations models, in which variables were arbitrarily labeled “endoge-

---

5 Processes with stochastic trend are also called integrated processes, or unit-root processes. The pioneering work of Dickey and Fuller (for example, Fuller, 1976) on unit root testing grew from a desire, motivated by Box and Jenkins, to determine whether various series displayed stochastic trend. The similarly pioneering work of Granger and Joyeux (1980) on “long memory,” or “fractionally-integrated,” processes grew from attempts to generalize the idea of integration on which Box and Jenkins relied so heavily; see Diebold and Rudebusch (1989) for a macroeconomic application of long memory models and Baillie (1996) for an insightful recent survey.
nous’ or ‘exogenous.’ In vector autoregressions, in contrast, all variables are endogenous.

The mechanics of vector autoregressions are straightforward. Recall that we approximate dynamics with a univariate autoregression by regressing a variable on its own past values. In a vector autoregression, by way of logical extension, we regress each of a set of variables on past values of itself and past values of every other variable in the system. Cross-variable linkages are automatically incorporated because we include lags of all variables in each equation, and because we allow for correlations among the disturbances of the various equations. It turns out that one-equation-at-a-time least squares estimation of vector autoregressions is statistically efficient, in spite of the potential correlation of disturbances. Moreover, it is relatively simple and numerically stable, in contrast to the tedious numerical optimization required for estimation of multivariate ARMA models.

Many multivariate extensions of the Box-Jenkins paradigm are conveniently implemented in the vector autoregressive framework. Here we introduce a few to help convey a feel for the breadth of modern time-series econometrics and forecasting. The discussion is necessarily brief; for a more detailed introduction to modern time series forecasting, see Diebold (1998).

Granger (1969) and Sims (1972) made important early multivariate contributions, providing tools for exploring causal patterns in multivariate systems. The Granger-Sims causality notion is predictive; we say that \( x \) Granger-Sims causes \( y \) if the history of \( x \) is useful for forecasting \( y \), over and above the history of \( y \). We commonly use Granger-Sims causality tests to help identify and understand the patterns of cross-linkages and feedback in vector autoregressions.

The dynamic factor model of Sargent and Sims (1977) and Geweke (1977) is another important early multivariate contribution. The essential idea of dynamic factor models is that some economic shocks are common across sectors and others are idiosyncratic, so that large sets of variables may depend heavily on just a few common underlying sources of variation, a common feature of economic models and evidently also of economic data. The common shocks, or “factors,” produce comovements and facilitate parsimonious modeling and forecasting of large numbers of variables. Dynamic factor models have proved particularly useful with the emergence of macroeconomic panel datasets, including cross-country, cross-region, and cross-state data. Important recent contributions include Stock and Watson (1989), Quah and Sargent (1993), Forni and Reichlin (1997), and Stock and Watson (1997).

Granger (1981) and Engle and Granger (1987) develop the related idea of cointegration. We say that two series are cointegrated if each contains a stochastic trend, yet there exists a linear combination of the two trends that does not. Thus, for example, each of two series \( x \) and \( y \) may contain stochastic trend, but the spread

---

4 Ben McCallum notes in private communication that in an important sense, the situation was even worse: the endogenous-exogenous labeling was arguably not arbitrary, but rather systematic, with policy variables labeled as “exogenous” on the grounds that they could have been managed exogenously by policymakers if they had been unorthodox enough to do so.
x-y may not. It is apparent that in such situations stochastic trends are shared, which makes the series move together. This is the essence of the Stock-Watson (1988b) “common trends” representation of cointegrated systems and is precisely the same idea as with the intimately-related dynamic factor model: comovements may be due to dependence on common factors. Cointegration is also intimately connected to the idea of error-correction, pioneered by Sargan (1964) and long a cornerstone of “LSE econometrics,” in which the current deviation of a system from equilibrium conveys information regarding its likely future course and is therefore useful for forecasting. Indeed, there is a formal equivalence between cointegration and error correction, as established by Engle and Granger (1987).

All of the discussion thus far has been based on linear models. Nonlinear forecasting methods have also received increasing attention in recent years, as the Slutsky-Yule theory of linear modeling and forecasting has matured, and that trend will likely continue. Models of volatility dynamics, which permit volatility forecasting, are an important example; the literature began with Engle’s (1982) seminal contribution, and recent surveys include Bollerslev, Chou and Kroner (1992) and Bollerslev, Engle and Nelson (1994). We will, however, avoid discussion of nonlinear methods here for the most part, because although they are clearly of value in areas such as finance, they are less useful in macroeconomics. There are two reasons. First, many of the nonlinear methods require large amounts of high-quality data for successful application, whereas in macroeconomics we typically have short samples of data contaminated by substantial measurement error. Second, many of the nonlinearities relevant in fields such as finance simply don’t appear to be important in macroeconomics, perhaps because macroeconomic data are highly aggregated over both space and time. Early on, for example, models of time-varying volatility were fit to macroeconomic data, such as aggregate inflation, but those ventures were quickly abandoned as it became clear that volatility dynamics were much more important in high-frequency financial data.

One strand of nonlinear literature, however, is potentially relevant for macroeconomic forecasting—the idea that business cycle expansions and contractions might be usefully viewed as different regimes, which focuses attention on tracking the cycle, charting the timing of turning points, and constructing business cycle chronologies and associated indexes of leading, lagging and coincident indicators (Diebold and Rudebusch, 1996, 1998). Burns and Mitchell (1946) is a classic distillation of early work in the nonlinear tradition, much of which was done in the first four decades of the twentieth century, and which was extended in breadth and depth by G. Moore, V. Zarnowitz, and their colleagues at the National Bureau of Economic Research in the ensuing decades.

Regime-switching models are a modern embodiment of certain aspects of the Burns-Mitchell nonlinear forecasting tradition. The idea of regime switching is implemented through threshold models, in which an indicator variable determines...
the current regime (say, expansion or contraction). In the observed indicator models of Tong (1990) and Granger and Teräsvirta (1993), the indicator variable is some aspect of the history of an observable variable. For example, the current regime may be determined by the sign of last period’s growth rate. In contrast, Hamilton (1989) argues that models with unobserved regime indicators may be more appropriate in many business, economic and financial contexts. In Hamilton’s widely-applied model, sometimes called a ‘‘Markov-switching’’ or ‘‘hidden-Markov’’ model, the regime is governed by an unobserved indicator.

The future of nonstructural economic forecasting will be more of the same—steady progress—fueled by cheap, fast computing, massive storage, and increased sophistication of numerical and simulation techniques. Such techniques are rapidly allowing us to estimate complicated models not amenable to treatment with standard methods, and to dispense with the unrealistic assumptions often invoked in attempts to quantify forecast uncertainty. Efron and Tibshirani (1993) and Gourieroux and Monfort (1996), for example, provide good examples of recent developments. The future of nonstructural macroeconomic forecasting will likely also involve combining aspects of the linear and nonlinear traditions, as for example, with vector autoregressive models that allow for factor structure and regime switching (Diebold and Rudebusch, 1996; Kim and Nelson, 1998a, b).

A New Wave of Macroeconomic Theory—and Structural Forecasting

Nonstructural models are unrestricted reduced-form models. As such they are useful for producing unconditional forecasts in a variety of environments ranging from firm-level business forecasting to economy-wide macroeconomic forecasting. Again, however, in macroeconomics we often want to analyze scenarios that differ from the conditions presently prevailing, such as the effects of a change in a policy rule or a tax rate. Such conditional forecasts require structural models.

As we have seen, an early wave of structural econometrics followed the development of Keynesian theory. But the Keynesian theory was largely based on postulated decision rules, rather than the economic primitives of taste and technology; the system-of-equations approach to structural econometric forecasting inherited that defect and hence, wasn’t really structural. Ultimately the system-of-equations approach to both theory and forecasting declined in the 1970s. Progress toward a new and truly structural macroeconomic forecasting had to await a new wave of powerful theory developed in the 1970s and 1980s. The new theory has its roots in the dissatisfaction, percolating in the late 1960s and early 1970s, with the system-of-equations approach. In many respects, the essence of the new approach is methodological and reflects a view of how macroeconomics should be done. Lucas (1972), in particular, paved the way for a new macroeconomics based on dynamic stochastic model economies with fully-articulated preferences, technologies, and rules of the game. Hence the descriptively accurate name: dynamic stochastic gen-
eral equilibrium (DSGE) modeling. The key innovation is that DSGE models are built on a foundation of fully-specified stochastic dynamic optimization, as opposed to reduced-form decision rules, and are therefore not subject to the Lucas critique. But ultimately the "new" theory is neither new nor radical; rather, it is very much in the best tradition of neoclassical economics.

The new research program has sought from the outset to make clear that dynamic stochastic general equilibrium models can address practical, empirical questions. Early on, for example, Kydland and Prescott (1982) used DSGE models to argue that a neoclassical model driven purely by real technology shocks could explain a large fraction of U.S. business cycle fluctuations. Hence the early name "real business cycle" models. Later work, however, broadened the approach to allow for rich demographic structures, imperfect competition and sticky prices (and hence real effects of monetary shocks), and much else; the papers collected in Cooley (1995) offer a good overview. Ultimately, again, the essence of the new approach is not about whether the shocks that drive the cycle are real or monetary, whether prices are flexible or sticky, or whether competition is perfect or imperfect, but rather about the way macroeconomic questions should be approached.

The earliest and still rapidly developing strand of the dynamic stochastic general equilibrium literature makes use of simple "linear-quadratic" models, in which agents with quadratic preferences make optimizing decisions in environments with linear production technologies. Linear-quadratic models are surprisingly more flexible than a superficial assessment might indicate; they nest a variety of popular and useful preference and technology structures. Linear-quadratic models are also convenient, because a large literature provides powerful methods for solving, analyzing and forecasting with them. Moreover, it turns out that optimizing behavior within linear-quadratic economic models implies decision rules, such as those that govern consumption or investment behavior, that are stochastic linear functions of other variables. In particular, the decision rules conform to the great workhorse nonstructural model, the vector autoregression, subject to restrictions arising from theory. The result is a marvelous union of modern macroeconomic theory and nonstructural time-series econometrics, paving the way for a new structural econometrics.

Maximum likelihood methods are central to linear-quadratic DSGE modeling and trace to the important early work of Hansen and Sargent (1980); the modern approach is to construct and maximize the likelihood function using a state space representation in conjunction with the Kalman filter. Initially, maximum likelihood estimation was challenging in all but the simplest cases, but recent improvements in numerical algorithms and computing power have begun to make estimation and forecasting with linear-quadratic DSGE models workable for routine analysis and forecasting. Hansen and Sargent (1998) provide a powerful overview, synthesis, and extension of linear-quadratic DSGE modeling.

Kydland and Prescott (1982) started a distinct, but intimately related and equally important, strand of the DSGE literature. Two key features differentiate their product. First, Kydland and Prescott do not require that preferences be quadratic and technology be linear; instead, they use non-linear-quadratic models that are (arguably) more natural. Non-linear-quadratic models are challenging to solve,
and the Kydland-Prescott program spurred a great deal of important research on numerical and computational aspects of model solutions. One interesting outcome of that research is that, although non-linear-quadratic models don’t have tidy vector-autoregressive systems of decision rules, they nevertheless often have decision rules that can be accurately approximated by vector autoregressions.

Second, Kydland and Prescott acknowledge from the outset that their models, like all models, are false, and they recognize that traditional econometric estimation procedures such as Gaussian maximum likelihood may loose some of their appeal in such situations. Partly for that reason, and partly because of the sheer difficulty, non-linear-quadratic DSGE modelers often eschew formal estimation in favor of less structured “calibration” methods, as described in this journal in Kydland and Prescott (1996). Calibration means different things to different people, but the central idea is learning about the properties of a complicated DSGE model and attempting to assess its agreement with the data, based on simulations of the model economy. The parameters underlying the simulated model economy are typically set informally, sometimes by statistical considerations such as generating realistic amounts of volatility in observed variables, sometimes by economic considerations such as producing “reasonable” steady state behavior, and sometimes by appealing to previous empirical studies.

Calibration is the natural response of economic theory to the computer age; hence the commonly-used synonym “quantitative economic theory.” Calibration, however, fails to provide a complete and probabilistic assessment of agreement between model and data and therefore, fails to deliver the goods necessary for forecasting with DSGE models. Econometric discontent based on recognition of that fact has been simmering for some time and is expressed forcefully by Sims (1996) in the Winter 1996 symposium in this journal on calibration and econometrics. The growing list of such symposia includes a special issue of Journal of Applied Econometrics (see the introduction by Pagan, 1994) and an Economic Journal “Controversy” section (see the introduction by Quah, 1995).

---

7 See, for example, Rust (1996) and Judd (1998), who describe and contribute to the impressive advances being made for solving non-linear-quadratic stochastic dynamic programming problems.

8 The reasoning is straightforward. Loosely speaking, under correct specification, Gaussian maximum likelihood estimates converge to the true parameter values as the sample size grows; hence the estimated model converges to the true model, which is the best model to use for any purpose. Under misspecification, however, the parameters can’t converge to the “true” values, because an incorrect model has been fitted. Instead, the parameters converge to values that make the fitted model the best approximation to the data, where the measure of goodness of approximation is induced by the estimation procedure. The key insight is that, under misspecification, the best approximations for one purpose may differ from the best approximation for another purpose. The measure of goodness of approximation associated with Gaussian maximum likelihood estimation is 1-step-ahead mean squared forecast error. Thus, if the model is to be used for 1-step-ahead forecasting, and if mean squared error is the relevant loss function, Gaussian maximum likelihood estimation is a logical choice. If, on the other hand, the model is to be used for another purpose, such as 4-step-ahead forecasting, Gaussian maximum likelihood estimation is less appealing.

9 Important exceptions exist, however, such as McGrattan, Rogerson, and Wright (1997), who estimate non-linear-quadratic DSGE models by maximum likelihood methods.

10 See also Hansen and Heckman (1996), in the same symposium, the lead paper in which is Kydland and Prescott (1996).
If dynamic stochastic general equilibrium models are to be used for forecasting, formal econometric analysis is desirable for at least two reasons. First, forecasting is intimately concerned with the quantification of the uncertainties that produce forecast errors. Accurate assessment of such uncertainties is key, for example, for producing credible forecast confidence intervals. Calibration methods, unlike probabilistic econometric methods, are ill-suited to the task.

Second, simply using a priori “reasonable” parameter values, although useful as a preliminary exercise to gauge agreement between model and data, is not likely to produce accurate forecasts. For example, it might be commonly agreed that technology shocks are likely to be serially correlated, and for purposes of a preliminary calibration exercise we might adopt a simple first-order autoregressive scheme and set the serial correlation coefficient to an arbitrary but “reasonable” value, such as .95. But the first-order autoregressive process might be an oversimplification of reality, and even if adequate, the serial correlation coefficient that maximizes forecast accuracy might be, say, .73, not .95. Although such details might make little difference to a qualitative analysis of the model’s properties, they are likely to make a major difference for forecast accuracy. In short, **accurate forecasting demands quantitative precision.**

The upshot is that for forecasting we need to take seriously the “fit” of DSGE models and search for best-fitting parameters. Moreover, we need estimation methods that are tractable, yet capable of delivering probabilistic inference, and we need to take misspecification seriously. Calibration and maximum likelihood estimation meet some, but not all, of those goals. Calibration is tractable and takes misspecification seriously, but it is not probabilistic. Maximum likelihood is probabilistic, but it is often challenging to implement and may not take misspecification seriously enough.\(^{11}\)

The choice set, however, now includes a number of procedures other than calibration and maximum likelihood; in particular, new estimation procedures are being developed that attempt to find a middle ground. There are a variety of ways to proceed. Sims and his co-authors, including Leeper and Sims (1994), Leeper, Sims and Zha (1996), and Sims and Zha (1996), use a strategy based on examining the entire likelihood function, rather than just its maximum. Christiano and Eichenbaum (1992) match selected moments of real macroeconomic data and data simulated from a DSGE model. In similar fashion, Canova, Finn and Pagan (1994) match vector autoregressions, Rotemberg and Woodford (1997) match impulse-response functions, and Diebold, Ohanian and Berkowitz (1997) match spectra. Finally, Rotemberg and Woodford (1996) and Diebold and Kilian (1997) develop

\(^{11}\) Nevertheless, if calibration and Gaussian maximum likelihood estimation were the only strategies available for parameterizing a serious DSGE forecasting model, the choice would probably not be difficult: maximum likelihood estimation appears preferable, because: a) it enables probabilistic inference; b) recent improvements in computing and algorithms are making implementation less tedious, especially in the linear-quadratic case; and c) although the measure of goodness of approximation associated with Gaussian maximum likelihood estimation is 1-step-ahead mean squared forecast error, which may not be appropriate in all situations such as when interest centers on longer-horizon forecasts, short-horizon forecasts often are of interest.
procedures that enable us to assess agreement between model and data predictability at various horizons of interest.

If structural modeling and forecasting have come a long way, they still have a long way to go; in closing this section, it is tempting to comment on a few aspects of their likely future development. DSGE theory will continue to improve and will begin to take certain aspects of reality, such as heterogeneity, more seriously. In particular, recent work (Geweke, 1985; Kirman, 1992; Altissimo, 1997) highlights the fact that aggregator functions may not be structural with respect to policy interventions, which suggests that current-vintage representative-agent DSGE models may not fully address the Lucas critique.¹²

The stochastic dynamics of driving variables, such as technology shocks, will be similarly enriched to reflect recent developments in nonstructural modeling, such as the possibility of regime switching, and to allow for multiple sources of uncertainty, including measurement error. The resulting models will have approximate representations as vector autoregressions with factor structure, possibly involving cointegration, as in King, Plosser, Stock and Watson (1991), and possibly with regime switching, as in Diebold and Rudebusch (1996) and Kim and Nelson (1998a, 1998b). Formal econometric procedures will be used to diagnose possible model inadequacies, as in Chow and Kwan (1997).

One might expect the scale of DSGE forecasting models to grow over time. That is likely to happen, and current models that determine, for example, three or four variables in equilibrium, are likely to evolve into richer models that determine, say, eight or ten variables in equilibrium.¹³ But the expansion in scale is likely to stop there, for two reasons. First, the demise of the large-scale models heightened professional awareness of the fact that bigger models are not necessary better, an idea memorably enshrined in Zellner’s (1992) KISS principle: Keep It Sophisticately Simple. Second, in contrast to models in the system-of-equations tradition, which are typically estimated equation-by-equation and then assembled in modular fashion, the nature of DSGE models requires that their parameters be jointly estimated, which limits the complexity of the models that can be entertained.

Last and not least, shrinkage will likely emerge as a key component of estimation techniques for DSGE forecasting models. Shrinkage refers to the idea of coaxing, or “shrinking,” parameter estimates in certain directions. Shrinkage can be implemented using Bayesian methods to coax parameter estimates away from the likelihood maximum and toward the prior mean. It seems obvious that shrinkage in a “correct” direction will likely improve forecast performance. Less obvious, but equally true, is the insight that even shrinkage in “incorrect” directions can improve forecast performance, by drastically reducing forecast error variance at the potentially low price of a small increase in bias.

Shrinkage has a long history of productive use in nonstructural modeling and forecasting. For example, it has long been known that vector autoregres-

---

¹² See also Oliner, Rudebusch, and Sichel (1996), who document the instability of empirical Euler equations for investment.

¹³ The work of Sims and his coauthors has already reached that point.
sions estimated using Bayesian shrinkage techniques produce forecasts drastically superior to those from unrestricted vector autoregressions. The “Minnesota prior,” a simple vector random walk, remains widely used[14] Shrinkage has an equally bright future in the new structural modeling and forecasting. Shrinkage is a potentially tailor-made device for estimating potentially misspecified DSGE forecasting models, because, as we have seen, DSGE theory essentially amounts to restrictions on vector autoregressions. At one extreme, we can ignore the theory and forecast with an estimated unrestricted vector autoregression: no shrinkage, loosely corresponding to a Bayesian analysis with a diffuse prior. At the other extreme, we can directly impose the theory and forecast with a restricted vector autoregression: complete shrinkage, loosely corresponding to a Bayesian analysis with a “spiked” prior. Intermediate cases, corresponding to forecasting with vector autoregressions estimated with various informative, but not spiked, priors are potentially more interesting. First, we may use statistically-oriented priors, such as the familiar Minnesota prior, which shrinks toward a vector random walk. Second, we may use statistically-oriented, but theory-inspired, priors, such as one corresponding to factor structure. Third, we may use DSGE theory-based priors, as in Ingram and Whiteman (1994), to coax the estimates in directions implied by an explicit economic theory.

Concluding Remarks

In a recent New York Times article entitled “The Model Was Too Rough: Why Economic Forecasting Became a Sideshow,” economics writer Peter Passell noted: “Americans held unrealistic expectations for forecasting in the 1960’s—as they did for so many other things in that optimistic age, from space exploration to big government . . .” Our expectations for forecasting were quite appropriately revised downward in the 1970s and 1980s, and the ensuing era of humility has been good for all. The new humility, however, is not symptomatic of failure, just as the bravado of the 1960s was not symptomatic of success.

As the 1990s draw to a close, we find ourselves at a critical and newly optimistic juncture, with the futures of structural and nonstructural forecasting very much intertwined. The ongoing development of nonstructural forecasting, together with the recent developments in dynamic stochastic general equilibrium theory and associated structural estimation methods, bode well for the future of macroeconomic forecasting. Only time will tell whether linear-quadratic or non-linear-quadratic approximations to the macroeconomy are the best approach for practical macroeconomic forecasting, but regardless, the seeds have been

[14] For an extensive discussion, see Doan, Litterman and Sims (1984). The Bayesian vector autoregressive tradition continues to progress, as for example with the work of Sims and Zha (1997), who develop methods applicable to large systems.
sown for a new structural econometrics and structural econometric forecasting, a modern and thorough implementation of the Cowles vision. The new structural econometrics is emerging more slowly than did the earlier wave following Keynes, because the baby was almost thrown out with the 1970s bathwater; the flawed econometrics that Lucas criticized was taken in some circles as an indictment of all econometrics. It has taken some time to get on with macroeconomic work, but progress is evident.

The hallmark of macroeconomic forecasting over the next 20 years will be a marriage of the best of the nonstructural and structural approaches, facilitated by advances in numerical and simulation techniques that will help macroeconomists to solve, estimate, simulate, and yes, forecast with rich models. Moreover, related developments will occur in a variety of fields well beyond macroeconomics. It’s already happening and in some cases, progress has been underway for years, as evidenced by example from the recent literatures in industrial organization (Ericson and Pakes, 1995), labor economics (Eckstein and Wolpin, 1989; Stock and Wise, 1990; Rust, 1994), public economics (Rios-Rull, 1995), agricultural economics (Rosen, Murphy and Scheinkman, 1994), health economics (Gilleskie, 1997), development economics (Rosenzweig and Wolpin, 1993), environmental economics (Rothwell and Rust, 1995), and international economics (Backus, Kehoe and Kydland, 1994).

Helpful input was provided by the panel members and audience at the American Economic Association’s 1997 New Orleans roundtable, “Monetary and Fiscal Policy: The Role for Structural Macroeconomic Models,” as well as participants at seminars at the Federal Reserve Bank of Atlanta, the International Monetary Fund, and the 1997 International Symposium on Forecasting. I am also grateful to Brad De Long, Robert Fildes, Lutz Kilian, Alan Krueger, Ben McCallum, Timothy Taylor, Ken Wolpin, Mike Woodford, and seminar participants at the Federal Reserve Bank of San Francisco, especially Tim Cogley, Ken Kasa, Glenn Rudebusch and Tom Sargent. I, however, bear full responsibility for all remaining flaws. The National Science Foundation, the Sloan Foundation, and the University of Pennsylvania Research Foundation provided research support. This paper is available on the World Wide Web at (http://www.ssc.upenn.edu/~diebold/).

References


Granger, C. W. J. and R. Joyeux, “An Intro-


Yule, G. U., “On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer’s Sunspot Numbers,” Philosophical Transactions, 1927, 226A.


This article has been cited by:

1. Alok Pratap Singh, Jayendra Kumar, Iti Chauhan, Prasoon Kumar Saxena, Praveen Kumar Gaur, Satish Kumar Sharma, Dinesh Kumar Gupta, Nalini Kanta Sahoo, Pranjal Kumar Singh. Role of climate change in disasters occurrences: Forecasting and management options 149-180. [Crossref]
7. Nicholas G. Reich, Matthew Cornell, Evan L. Ray, Katie House, Khoa Le. 2021. The Zoltar forecast archive, a tool to standardize and store interdisciplinary prediction research. Scientific Data 8:1. . [Crossref]
10. Vangala Sarveswararao, Vadlamani Ravi. Chaos, Machine Learning and Deep Learning based Hybrid to forecast Consumer Price Index Inflation in India 2551-2557. [Crossref]
13. Apurv Jain. Macro forecasting using alternative data 273–327. [Crossref]
14. David Higuita-Alzate, Marisol Valencia-Cárdenas, Juan Carlos Correa-Morales. 2018. Método de combinación de pronósticos usando modelos Bayesiano y una metaheurística, caso de estudio. DYNA 85:207, 337-345. [Crossref]
15. Lasha Kavtaradze, Manouchehr Mokhtari. 2018. FACTOR MODELS AND TIME-VARYING PARAMETER FRAMEWORK FOR FORECASTING EXCHANGE RATES AND INFLATION: A SURVEY. Journal of Economic Surveys 32:2, 302-334. [Crossref]
17. John D. Levendis. ARMA(p,q) Processes 11–46. [Crossref]


22. Oleksandr Bandura. 2016. General economic cycles model – cumulative inefficiency model. *Ekonomichna teoriya* 2016:1, 86-100. [Crossref]


24. James Morley. Macroeconomics, Nonlinear Time Series in 1–30. [Crossref]


29. S. Mahdi Barakchian. 2012. Do Long-Run Theory Restrictions Help in Forecasting?. *Journal of Forecasting* 31:5, 401-422. [Crossref]


33. Gonzalo Fernández-de-Córdoba, José L. Torres. 2011. Forecasting the Spanish economy with an augmented VAR–DSGE model. *SERIES* 2:3, 379-399. [Crossref]


38. Dan S. Rickman. 2010. MODERN MACROECONOMICS AND REGIONAL ECONOMIC MODELING. *Journal of Regional Science* 50:1, 23-41. [Crossref]
41. James Morley. Macroeconomics, Non-linear Time Series in 5325-5348. [Crossref]
42. Lyle Kane. Linking TVET to Economic Opportunities in Post-Conflict Liberia 787-797. [Crossref]
43. James Morley. Macroeconomics, Non-linear Time Series in 525-548. [Crossref]
44. Charles Renfro. Econometric Diagnostic Tests 97-121. [Crossref]
45. Charles Renfro. Cointegration and Alternative Specifications 177-230. [Crossref]
46. Raphael Kahan. 2009. Forecasting GDP with the Leading Indicators: A VAR Approach. *SSRN Electronic Journal*. [Crossref]
47. Viviana Fernandez. 2008. Traditional versus novel forecasting techniques: how much do we gain?. *Journal of Forecasting* 27:7, 637-648. [Crossref]
54. LANCE BACHMEIER, SITTISAK LEELAHANON, QI LI. 2007. MONEY GROWTH AND INFLATION IN THE UNITED STATES. *Macroeconomic Dynamics* 11:1, 113-127. [Crossref]
55. Javier Andrés, Fernando Restoy. 2007. Macroeconomic Modelling in EMU: How Relevant is the Change in Regime?. *SSRN Electronic Journal*. [Crossref]


64. Monetary Transmission in Diverse Economies 25, . [Crossref]


68. Peter Anker. 2001. ECB monetary policy and the DM–dollar exchange rate: evidence from a Bayesian VAR. *Applied Economics* 33:12, 1553-1562. [Crossref]

69. Francis X. Diebold. 2001. Econometrics: Retrospect and prospect. *Journal of Econometrics* 100:1, 73-75. [Crossref]


72. Simonetta Longhi, Peter Nijkamp. Economic Integration and Labour Markets: Ways Forward 3-23. [Crossref]

73. Xi Chen, Ye Pang, Guihuan Zheng. Macroeconomic Forecasting Using Genetic Programming Based Vector Error Correction Model 759-773. [Crossref]