Reflections on Economic Forecasting

Daniel Bachman

(this version: 2/23/2025)

Abstract: The failure of economic forecasters to foresee key economic changes is well documented. I examine the inability to predict the 2007-09 financial market crash and recession in detail, and discuss some lessons to be learned from the experience. Based on this and my own experience as an applied forecaster, I suggest a taxonomy of forecast users, and a taxonomy of approaches that economists currently use to forecast the economy. One of those approaches—-structural models—remains the mainstay of applied forecasting. I discuss why applied forecasters continue to rely on this method, and suggest an agenda for economic forecasting research that would be more likely to supply useful tools to applied forecasters than the current focus on econometrics.

The reputation of economic forecasters would appear to be well deserved. That's based on the well-documented fact that economists don't forecast business cycles very well. (See, for example, Zidong An, Joao Tovar Jalles, and Prakash Longani (2018)). It's a bit unfair, since there are a lot of other things that economists forecast. In fact, I believe that our inability to forecast business cycles says more about business cycles than it says about economists. But we are supposed to help people understand and anticipate problems in the economy. And our track record is mixed—at best.

Consider a few of the failures to forecast recessions in the past 50 years.

1. The 1973-5 oil price shock and recession left economists scrambling to give a more prominent role to energy prices and the supply side in determining business cycles.
2. The 1980 and 1982 interest rate shocks created (in 1982) the deepest recession in the United States since the Great Depression. We now view these as having been "engineered" to reduce inflationary expectations, but one doubts that, had policymakers known the price beforehand (an unemployment rate of 10.8% in November 1982), that they would have employed such harsh therapy. Why didn' t we tell them?
3. In the runup to the 2007-09 recession, few economists caught on to proliferation of exotic financial instruments that increased leverage leading the very damaging near shutdown of the financial markets. The "financial plumbing" of the economy wasn't interesting or important—until it proved to be key to understand current events.

4. The 2020 recession was caused by a completely unpredictable (to economists) event—a pandemic. Of course, epidemiologists had been warning of the possibility of a pandemic for years. But it's outside the scope of our usual discourse, and thus economists did not present that risk to forecast users. We might want to say that it's outside our scope, but forecast users don't necessarily agree.

To be clear: these spectacular failures did not lead to improved theory or econometrics (at least as measured by ability to predict business cycle downturns). Fifty years of research on expectations consistency, underlying behavior of economic agents, regime change, and econometrics left economists just as bereft in 2007 and 2020 as they were in 1973.

The 2008 financial crisis provides an interesting case study in the failure of economics to warn policymakers, businesses, and households that trouble lay ahead. It's worth looking at the incident in more detail.

I. The failure of economic forecasting in the early 2000s.

After the fact, we understood what happened. A framework in which mortgage originators did not take on risk, but passed it on to asset purchasers, created incentives that led to too much risk in financial markets. Overly risky investments and leverage eventually brought down major financial intermediaries, threatening day to day economic transactions. The basic outline of a financial crisis and crash is (or should have been) all too familiar from even a cursory study of financial history.

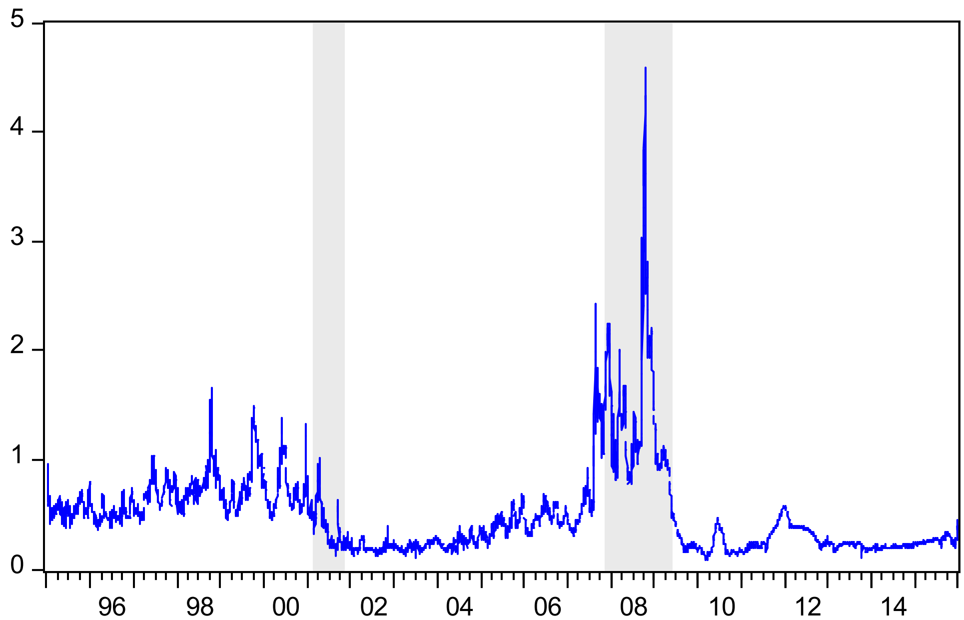
Some economists did signal trouble. They were ignored by business decision makers, policymakers, and individuals caught up the possibility of large short-term gains. It was easy to pretend that the high volume of housing lending and complex financial structure behind it were really risky given the fairly robust debate about whether there was really a bubble. Kristopher S. Gerardi, Christopher L. Foote, and Paul S. Willen [2010] provide a fine description of the contentious and technical discussion that preceded the bursting of the bubble. As they point out, any decision-maker eavesdropping on this debate would have not heard sufficient professional agreement to have driven significant decisions (either by policymakers to burst the bubble or by investors to bet against it).

The events of late 2008 posed a huge challenge to forecasters, especially those using econometrics. It's not surprising that any sort of econometric-based forecast would have been wrong. Econometric forecasts, after all, are based on historical relationships. And historical relationships simply fell apart.

The TED spread, the "unknown unknown"

Figure 1 shows the historical spread between LIBOR interest rates and 3-month U.S. Treasury rates. The figure shows the dangers of using history to predict the future. We can certainly imagine an econometrician looking at this spread and saying in late 2006 that this spread could not go above 200 basis points--if anybody had been interested in this spread in late 2006. Not only had this spread never reached the heights experienced during the crisis, there was no evident relationship between the spread widening and the U.S. business cycle.

Figure 1: Ted Spread
Percentage Points



Source: Federal Reserve Bank of St. Louis

This gives us one important clue to part of the reason we failed. Estimation cannot capture what is not in the data. Using the past to predict the future only works when the future resembles the past, not when it is different. Although every econometrician learns this, the implications never really get into our heads. "This time is different" is the real truth: and that limits what econometrics can tell us. *The sample is always small*.

William Krasker's (1980) paper described the problem of "small samples" in macroeconomic forecasting, but the complete implications of this problem have never really been accepted by the profession. When we don't see all states of nature in the historical data, as was certainly the case before the 2008 financial collapse, econometrics will be misleading. Consider the recessions I noted at the beginning of the paper. In all cases, they involve a state of nature that hadn't occurred before.

The Lucas critique writ large

Really good economic modeling should be able to get around this problem, you might think. Yes, the spread relationship pictured above might well have been unprecedented. But if we understood and modeled the underlying (presumably fixed) determinants of behavior, we should have seen it coming. All we have to do is map the unchanging determinants of behavior to the existing economic data. A small task.

Here, the reader will recognize Lucas' famous critique [1976] of structural model-based simulations rearing its head. Without knowing the "deep underlying parameters," Lucas claimed, we don't know when observed relationships (like the interest rate spread) have changed.

Let's think about this in the context of the events of September 2008. To correctly model and forecast the spread in Figure 1, we would have to know the profit functions of institutions that trade these assets. To be more precise, we would really need to know the utility functions and compensation schemes of the people in charge of trading for those institutions. This—as we have painfully learned—might well not line up with the firm profit functions at all. And, of course, we have to assume that people exhibit unchanging, rational behavior.

But, wait, there is more we would have to know! A fully expressed model would have to include not only deep underlying economic behavior of investors and consumers, but the deep underlying determinants of political and organizational behavior as well. The model would have to explain *why* bank employees had incentives that put their own employers at risk. It would have to be able to explain *why* U.S. government officials chose to force a rescue of Bear Stearns, but not of Lehman Brothers . Without modeling institutional and political behavior—an area that macroeconomics and economic forecasting has almost entirely neglected—accurate economic predictions will not be practical. This goes quite a bit beyond simply writing a representative utility function and profit function, and calibrating the resulting model until it looks like the past. And it goes beyond looking mechanically for "regime change." The specifics of the regime change—the institutional specifics—matter.

And yet, people still want—and need—economic forecasts. I will propose some explanations of why that is the case, and some taxonomies of users and types of forecast to help situate better what we economic forecasters are trying to do. This is absolutely necessary if our research is to have any real meaning for us and for our "customers." And it might even help lead research in more useful directions.

II. Why do we forecast?

It’s a simple question with multiple answers. And it’s one we don’t ask ourselves enough, because the reason for the forecast is the context for everything else about it. It is impossible to judge whether a forecast is “useful” or “correct” or whether a forecasting method is accurate or unbiased without knowing why we care.

I propose the following taxonomy of forecasting:

Type 1: Academic or pure forecasting.

These are forecasts that are designed mainly for the audience of researchers in economics and related fields. Certainly, the forecasts may be picked up and used by others, and the forecaster might be quite pleased if that happens. But the primary purpose of the forecast is to advance the discussion within economics. There are two reasons for these forecasts:

Forecasting is interesting. There is a pure challenge in seeing if we can successful discern patterns in our world, a challenge that has driven the development of all science. There’s nothing wrong with knowledge for it’s own sake, although economists normally focus on usefulness of their work.

Out of sample forecasting is the ultimate test of a theory or model. The whole point of the model, after all, is to explain the world; if it gives us a bad prediction, it hasn’t explained the world well. Thus, forecasting is an important part of theory and model validation.

Type 2: Applied forecasting

Applied forecasters serve customers or users who must make decisions. There are a wide variety of potential users, something that many of us forget. Broadly speaking, forecast users are divided into two groups: policymakers, who have to shape government policy in one manner or another, and individual economic actors (businesses and households) who must make decisions about the future.

In the spirit of the businesslike nature of applied forecasting, I am going to create the taxonomy of users with an approach from marketing: the “persona”. A “persona is an example a specific customer a marketer will encounter. In our case, these “personas” are non-exhaustive list of the types of people who need our help.

Policy customers are those who want to understand economic conditions so they can take decisions that might affect those very conditions. Examples of such personas might include:

1. A legislator (and their staff) who need to understand the implications of different legislative proposals.
2. A central bank governor charged with determining the short-term interest rate and availability of liquidity.
3. A task force charged with developing energy policy.
4. A budget office determining the assumptions government agencies should use in developing their budgets.

Nonpolicy economic actors need to make decisions, but aren't concerned that their own decisions might have an impact on the economy. They may be making decisions in th public sector, or in the private sector. Examples of such personas might include.

1. A foreign intelligence officer developing estimates of the long-term military capability of an adversary.
2. A Treasury official charged with planning the future mix of issues of government debt.
3. Local school officials budgeting for future school capacity.
4. Corporate planners weighing different long-term investments.
5. A plant manager concerned with managing inventory levels.
6. An auto company's production team planning the number and mix of cars to be delivered to dealers over the next month.
7. A purchasing manager budgeting for specific raw material needs.
8. The CEO of a home construction company deciding whether to borrow to build a development of houses for sale.
9. A bank's comptroller engaged in short term money management.
10. The investment committee of a mutual fund.
11. A marketing executive setting goals for a sales team.
12. A college student deciding on a career.

Notice the larger number of examples of decision makers who are not policymakers. Researchers in economic forecasting tend to focus on policymakers. But the uses of economic forecasts go far beyond determining optimal monetary and fiscal policy.

III. What are we forecasting?

Economic forecasting research tends to focus on forecasting two variables: output (usually measured by GDP) and inflation (usually measured by either the CPI, the personal income expenditure index, or the GDP deflator). A perusal of the personas I've listed suggests that these are not actually variables of interest for most customers of applied forecasts. In many cases, industry or sector forecasts are the focus of the customer.

That doesn't mean that the traditional focus is wrong. An industry or sector forecast relies on an underlying forecast of the macroeconomy. But it's still only half the story for most economic actors. Indeed, it's usually a mistake to try to forecast a sector or specific commodity price without having an explicit assumption about the trajectory of the overall economy. Those who do this (yes, they exist) often wind up with inconsistent or illogical forecasts. (An interesting technical example involves a naïve forecaster who assumes that an average of past values is a reasonable forecast when the variable of interest has a unit root, and is not mean reverting.) So it's not wrong for economists to spend considerable effort on how to forecast output and inflation.

This also has implications for forecast evaluation. From the customer's point of view, there are two potential sources of forecast error. First, there is the forecast error associated with getting the overall economy wrong. And, second, there is the forecast error associated with getting the relationship between the overall economy and the part that matters to them wrong. Economic research rarely focuses on the second question, although it's of great importance in applied forecasting.

Lending a further wrinkle to the "use case" and perhaps to how we evaluate forecasts is the fact that most decisions aren't "scalar," while our forecasts are. What I mean by this is that decisions tend to be qualitative: whether to become a welder or a hospital technician, whether to build a school or rely on portable classrooms, whether to sign a long-term supply contract or not. When a user translates the forecast into a yes or no decision, scalar measures of success or failure may not be relevant. A business in a cyclical industry mainly cares whether the economist can predict a recession, not whether GDP will grow 1.5% or 2.0%. Loss functions matter, and while we sometimes talk about these, standard forecasting evaluation turns (all too often) back to reducing mean squared error---which is at best a simplification of the problem.

IV. How do we forecast?

The question of how to forecast is, of course, at the heart of a long literature in economics. But this literature is surprisingly biased.

Forecasts do not have to be expressed as rational numbers. In fact, in many (most?) cases outside economics, the need is to forecast events, conditions, or states of the world. While economics can do this, we tend to think of the problem as one of solving a set of simultaneous equations. Thus we get Robert Lucas and Thomas Sargent [1978] describing economic forecasting this way:

> For purposes of ex ante forecasting, or the unconditional prediction of the vector y_{t+1}, y_{t+2} given observation of y_s and x_s, s < t, the estimated reduced form, together with [a vector autoregression of x_t, the exogenous variables] is sufficient. This is simply an exercise in a sophisticated kind of extrapolation, requiring no understanding of the structural parameters, that is, the economics of the model.

Here, forecasting is reduced to a problem of engineering: with good enough methodology, we'll get a best-in-show forecast. Hence the emphasis on formal estimation of exogenous variables (x(t)), something no practical forecaster bothers do to formally. (One wonders exactly how one runs a vector autoregression on the political process that drives key policy variables.)

This is a very limited type of forecasting, and useful only in those specific cases where the y(t) and x(t) vectors can be expressed in continuous rational measures. Many—most—social science and even economic questions do not fit this case. To take a few examples; the engineering approach would be useless in predicting whether Iraq would invade Kuwait in 1991, or what type of exit treaty the British government would successfully negotiate with the EU, or whether a mentally disturbed individual poses a threat to the community. Forecasting continuous time rational variables is a specialized type of forecasting. It is by no means the only way to approach the problem.

Even in economics, there is a long history of acknowledging (in applied forecasting) that numerical methods are insufficient for the purpose. Judgement often enters the forecast through the somewhat disrespected form of "addfactors," an issue I will discuss below.

But it's true that macroeconomic forecasting is fortunate to have the problem posed (mostly) in a way that allows for extensive use of quantitative techniques (econometrics). And so, in macroeconomics, "forecasting" is almost synonymous with projections from statistical models. Within that general category, there are three approaches.

Approach 1: The search for underlying parameters

The first approach, which most economists would view as the "gold standard," is to attempt to create a true, complete picture of the behavior of economic actors, and link it to economic measures of interest. This is what Robertt Lucas had in mind when he claimed that analysts ultimately needed to know the unchanging (in the face of policy shifts) —deep structural parameters‖ that determine economic behavior. In a practical sense, it's the flavor of the DSGE approach.

Of course ,we don't have direct measures of deep structural parameters or the determinants of economic behavior (such as utility). Economists have made great efforts to uncover some of these parameters, but their efforts depend on a mix of strong assumptions about human behavior and advanced econometric techniques with their own quirks and arbitrary statistical assumptions. I would suggest that, given how poorly economists did in predicting the 2008 financial crisis, we haven't gotten very far. We might want to ask why.

Contrast what we know about, say, risk aversion among workers of different education levels (the sort of thing that might be considered a deep structural parameter) to what we know about wages and employment by industry and occupation. Or compare what we know about production and profits levels in the oil and gas industries compared to what we know about the marginal elasticity of substitution between oil and natural gas. We are quite certain about the outcomes (e.g., wages by occupation and industry or the profitability of different industries) but have little direct observation to guide us about economic concepts such as risk aversion or the marginal elasticity of substitution in the energy industry. And forests worth of paper for econometric journals haven't gotten us very far along the road of teasing the deep structural parameters out of the outcomes data.

Further, to work at the macro level, these models must ignore many problems of aggregation. Aggregation has always been a theoretical and practical problem for macroeconomics. The aggregation assumptions required in a seven-equation DSGE model are hardly less breathtaking and artificial in this regard than the assumptions required in a traditional 800 equation macro model.

The DSGE approach also ignores the deep structural parameters behind the *institutional and political framework* of the economy. It's necessary to do this, or the models would become intractable very quickly. This is not a trivial matter. It might even be possible to claim that most of our forecast error is generated in the political sphere—that is, we haven't modeled the political economy very well. Why not put more effort there? Of course, we are then back to the problem that we have to leave econometrics behind. We might even need to talk to political scientists.

Finally, I think we all have to acknowledge that the behavioral assumptions behind these models are almost certainly wrong, even if they are a useful simplification for economists.

Most psychologists would no doubt laugh at our rational expectations ideas. Given what we see in financial markets, the general public would have reason to laugh, too. Though, come to think of it, the general public might not find the resulting analytic failure so funny. DSGE models can't take into account fear and greed—motivations which had a major role in the more than one financial crisis.

Approach 2: Atheoretic statistical models

The second approach is to refuse to consider theory at all. Instead, the forecaster simply characterizes the statistical relationship between the variables over some historical period. That, of course, is the essence of the Vector Autoregression (VAR) approach. This has a lot of appeal. We simply let the data do the talking, and accept its limits. We directly use the data at hand, and don't have to worry about mapping from behavioral assumptions to official economic measures.

This approach avoids the problems involved in the search for deep structural parameters that may or may not exist. But it still suffers from the other shortcomings of the theoretical approach. Many of the drivers, and even outcomes, that we are trying to predict may not be amenable to measurement by continuous function variables. It's unlikely that all states of nature are included, and virtually impossible to include all possible variables of interest. So, for example, VAR models estimated before 2008 did not include measures of the health of the financial plumbing, because nobody thought about it. VAR models estimated before 2020 didn't include any public health measures. What measures would we need to include now to anticipate the next crisis?

And the small sample problem remains. Despite what appear to be long estimation periods, samples used in VAR estimation are unlikely to include all states of nature.

Designers of VAR systems are faced with an unpleasant tradeoff. Long runs of historical data provide better estimation properties, and reduce the small sample problem (but don't by any means eliminate it). But this is at the expense of possible misleading parameter estimates because parameter changes are more likely over long periods of time. There has been some useful advancement in identifying structural changes, but doing so in real time is, by it's very nature, almost impossible.

Users of VAR systems face a further problem of interpretation. Experienced economic forecasters know that decision makers don't find black box methods very convincing. Decision-makers are typically looking for stories or narrative that explains how the behavior of economic actors is translated into observed data. VARs do not provide such a story. As Dan Hamilton [2011] explains, this limits the use of VARs in an applied framework.

Structural Models: back to Larry Klein

So that leaves the third approach. It's not as intellectually consistent as the other two. But it is more flexible, and, I think, better reflects the limits and use of our knowledge. And here I refer to the traditional structural macroeconometric models, as originally developed by Larry Klein (see, for example, Klein and Goldberger [1952]. To estimate these, we use economic theory to postulate relationships in existing data (the data which matters to decision makers). Many of the relationships and estimates are admittedly ad hoc, both in the way they are specified and the period over which they are estimated. It is a bit unfair to claim that anything goes in these models. It is fair to say that the theory corresponds to that taught in an intermediate undergraduate class in macroeconomics. I'm not sure this is a bad thing, since the intermediate course probably represents the closest we can come to a consensus within the profession.

A structural modeler starts with some a priori assumptions about how the model is expected to behave. For example, research literature might suggest that the short-term multiplier for government spending should be between 1.2 and 1.8, and that an acceleration of the money stock should, after several years, result in an acceleration of inflation. Of course, while the modeler prefers to aim towards consensus, such consensus may not be forthcoming. But the object of the modeling exercise is to reflect the general understanding of economists as much as possible.

Structural models are normally estimated on an equation by equation basis. Simultaneous estimation isn't very practical in models with several hundred stochastic equations. In addition, modelers engage in explicit parameter search (obviously biasing any relevant test statistics) as the characteristics of the model matter more than the test statistics for an individual equation Sometimes this requires simply "calibrating" rather than estimating parameters. Models are typically judged based on responses to a set of standard shocks.

But the model isn't the end of the work. The model is a tool. The forecasts are the result of a system in which a team of analysts *interacts* with the model, and is willing to overwrite the model when there is evidence that it is not picking up behavioral or institutional changes.

The model itself—like the deep structural models or the atheoretic models—still doesn't make explicit assumptions about the policymaking process or anything about the economic system as measured by data, the *system* can make allowances for unexpected policy and environmental shifts. It can also make allowance of fear and greed. Among other advantages, the model/system allows forecasters to react more quickly to unusual events than a pure VAR or DSGE model forecast. The forecasting team can alter or adjust the model's output when it becomes clear that events have broken out of historical norms.

The key factor that ties together model and judgment, and therefore creates a flexible forecasting system, is the use of addfactors. Addfactors are simply assumptions about the likely forecast errors. Using addfactors allows the forecaster to assume that an equation is wrong, or likely to give biased results, under current conditions.[1]

Now, addfactors are kind of a secret of the forecasting world. Here, for example, is a description of the use of addfactors in 1974 (E. Phillip Howrey, Lawrence R. Klein, and Michael D. McCarthy, 378).

> The contribution of econometric models to applied economics is increased substantially when it is used together with a priori information on error values to help reduce residual variance. And the intelligent user of econometric models would contend that there is a substantial amount of information that can be used to improve predictive performance. Many of the structural shifts that take place in the economy can be foreseen a quarter or more in advance; tax law changes and important strikes are examples…Even in the case of those shifts that were not foreseen, the model forecaster would surely have learned by experience: having observed that some of his structural equations had begun to exhibit large errors, he could be expected to take corrective measures….Sturctural change does erode forecasting performance, but to the extent that structural shifts are correctly anticipated or rapidly detected, the actual forecast error will not be as large as the MSE [Mean Squared Error] obtained by simple extrapolation of the model.

This, by the way, came in an article warning against running horse races between time series and unadjusted structural models that appeared almost 50 years ago! The article also spends quite a lot of time emphasizing that estimating structural models requires what the authors describe as "TLC" or tender loving care.

James Morley (2010) more recently described it this way: the residuals[2] act as a kind of safety pressure relief valve to address the fact that models are not reality. (See also Robert Fildes and Herman Stekler (1999, pp. 38-41).)

---

[1] There is another, more technical use of addfactors worth mentioning. The very last historical observation for a given series is associated with a specific error. If the error is quite large, a naive forecast for the next period from the same equation may show a substantial change as the error suddenly becomes zero. Forecasters therefore often ―smooth‖ the error to zero, typically by a simple rule (the error declines 10% each quarter, for example). The forecasting literature of the 1960s and 1970s tended to emphasize this use of addfactors, although the ability of addfactors to allow subjective alterations to the forecast is probably of more importance in understanding how the traditional econometric modelbased forecasting system works.

[2] Morley calls calls addfactors "residuals," following the usual practice at Macro Advisers, the publisher of this paper. I prefer the "addfactors" because it distinguishes the forecaster's assumptions from the actual residuals created by the estimation of the equation.

Addfactors are—to use a very bad word—subjective. This makes them apparently "unscientific." But there is far too little serious consideration given to the question of why this method of forecasting has survived the market test for so long, while other methods have not.

Addfactors allow us to have our cake and eat it, too. We can have the discipline of a model tool, but we can override the model when we know that historical elasticities no longer apply. That's the key to the long survival of structural model forecasting.

A further note about structural models, especially the large ones. Most of our non-policymaking personas aren't really interested in GDP and CPI. They need forecasts of specific sectors or prices of specific commodities. Of course, there has to be an underlying macroeconomic forecast underneath the sector forecast, but there is also some sector knowledge. Structural models provide a natural tool for this, but using input/out matrix to translate forecasts of demand components to forecasts of industry output. Neither DSGE models with their very tiny data sets, or VAR models can provide a useful tool to create industry and sector forecasts that are consistent with a macroeconomic point of view. That almost certainly helps to keep large scale structural models in business outside the policymaking community.

V. The evidence for judgment

The importance of judgment is not new, although it seems not to have had much impact on the research dynamics of the economics profession. See, for example, Fildes and Stekler, Michael R. Donihue (1993) and Roy Batchelor and Pami Dua (1991). It's worth quoting Batchelor and Dua's conclusion on this matter: "the irrational forecasters give a significantly higher weight to use of an Econometric Model, and a significantly lower weight to the exercise of Judgment, than do the rational forecasters." (p. 703).

While this conclusion may not be universally accepted, there seems to be little willingness to even acknowledge the possibility. For example, Graham Elliott and Allan Timmermann's (2008) survey article in the *Journal of Economic Literature* attempts to be a comprehensive survey of forecasting—and doesn't even mention the possibility of judgmental adjustments to forecasts.

Economists might want to be aware that outside macroeconomic forecasting, adding judgment to statistical forecasts is taken quite seriously. Paul Goodwin (2002) provides one survey and discussion of a very extensive literature.

This doesn't mean that the model isn't important. Although forecasters who use structural models exercise a lot of judgment, they are still working within a model-based framework. The model plays several roles in the forecast process:

1. Anybody forecasting the economy has an implicit model in mind. Having an explicit macro model forces the forecasters to better understand their model and to explain—to themselves as well as to users—the justifications for forecast adjustments.
2. The model ensures consistency, particularly throughout the large, detailed picture of the economy our users demand. Forecasters don't need to make judgments about everything. They can let the model determine how a relatively small number of judgments play out in a complete macroeconomic framework. This is efficient, since the meta-work of modeling then informs every forecast without requiring revisiting every relationship in the model.
3. Models help to ensure that the stories that matter to decision-makers are consistent with data and historical relationships among different economic sectors and actors. As I noted earlier in the discussion of VARs, decision makers want stories, not just numbers. The model ensures that the stories we tell are consistent with current economic knowledge and the statistical relationships between different economic measures.

VI. Conclusion: Some thoughts on the future direction of forecasting research

Although I started my career as an academic, I have spent most of it as an applied economic forecaster. And, in truth, my graduate education was a surprisingly poor preparation for that role. In fact, as I have trained other young and mid-career economists in forecasting, I've found that the first thing they need to do is "unlearn" some of the dogma that pervades graduate education and macroeconomic research in forecasting and econometrics. By no means does that mean that my graduate education, or that of my students, lacked value, or that we haven learned nothing since the 1960s. But all applied forecasters have to learn how to reverse the clock, and use methods that we are taught are methodologically wrong.

In the 1960s, there was a revolution in economic forecasting. The earliest Klein type structural models were a huge improvement over the indicator approaches, charts, and sheer guesswork that characterized the scene previously, as described in Walter A. Benjimin (2014). The truth is that we've advanced little since then. If I were to write a research agenda that would advance applied forecasting, and provide better tools for all economic decision makers, it would include the following.

1. *Less focus on econometrics*. There have been some useful advances in terms of techniques over the past 60 years. Atheoretic methods, the concepts surrounding cointegration and better methods of determining regime change are probably the most important of these advances. But, in the scheme of things, they are simply not that important, and don't measure up to the revolution caused by postwar econometrics and the use of early computers.

   In truth, the marginal value of econometrics for improving forecasting is probably quite low. The key problems that applied forecasters face are not those of improving the accuracy or obtaining better confidence limits, but of interpreting the wider context of the economy much better. An example of research that is particularly helpful might be the Economic Uncertainty Index developed by Scott R. Baker, Nicholas Blook, and Steven J. Davis (2016). Creating and improving such data sources is one of the keys to better forecasting.

2. *Cooperation with other social scientists*. Political scientists, sociologists, and psychologists all have the potential to improve our forecasts. Political scientists, of course, understand the nature of policymaking better than we do. Sociologists and anthropologists and help us see mass behavior differently—and if there's one thing we need, it's to understand how bubbles work. And, dare I say, psychologists might just be better at leading us to understanding the deep underlying parameters of human behavior than econometricians. Or, at least, provide high marginal value given how little we pay attention.
3. *Careful study how applied forecasters do their job*. The academic profession tends to dismiss applied forecasters because of the many theoretical and conceptual problems with the current approach. There's no question that the Structural Model approach has methodological problems. But it's what people do: and, as social scientists, we should pay attention to what people do rather than what we say they *should* do. Or, to appeal to our economic souls, notice that, of all the forecasting approaches developed since World War 2, only the Structural Model approach has

resulted in a profitable business model. Sixty years on, that is still true. It might be a clue to academic economists that there is something worth examining in the approach.

4. *Creating tools for rigorously applying and evaluating judgement in forecasting*. This is not impossible. In fact, it is being done! Philip Tetlock (2015) claims to have found ways to develop particularly accurate judgmental approaches. He seems to be unware of (or uninterested in) traditional econometric methods where applicable, while at the same time his name is absent from economic research on forecasting. We need to engage his and related research. At the very least, horse races between his approach and econometrics would seem to be necessary, even if they might just prove embarrassing.
5. *Widening our forecast space.* Too much of the current research focuses on issues involving monetary policy. That's probably because monetary policy institutions have the best reach into academia, and also have money for their own staffs. But the variety of personas that need good economic forecasts goes far beyond monetary authorities. Research into sector and longer term forecasting also has a marginal value, and presents interesting problems that are worth our time.

Economic forecasting is a fascinating project in itself, and when done correctly can help people make better decisions. I've been privileged to have spent much of my career doing that, and I like to think that I helped people in business understand economics and the economy better than they did. But I know that our profession has a lot of room for improvement. I hope the reader takes my suggestions in that spirit.

Bibliography

An, Zidong, Joao Tovar Jalles, and Prakash Longani, (2018) "How well do economists forecast recessions?", IMF Working Paper 18/39, International Monetary Fund: Washington, DC.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, (2016) Measuring Economic Uncertainty. Manuscript.

Batchelor, Roy, and Pami Dua, (1991), "Blue Chip Rationality Tests," *Journal of Money, Credit, and Banking*, vol. 23(4), November.

Donihue, Michael, (1993), "Evaluating the Role Judgment Plays in Forecast Accuracy," Journal of Forecasting, vol. 12.

Elliott, Graham, and Allan Timmerman, (2008), "Economic Forecasting," Journal of Economic Literature, vol. 46(1).

Fildes, Robert, and Herman Stekler, (1999), *The State of Macroeconomic Forecasting*, Lancaster University Management School Working Paper 1999/002.

Friedman, Walter A. (2014), *Fortune Tellers*, Princeton: Princeton University Press.

Gerardi, Kristopher S., Chistopher L. Foote, and Paul S. Willen, (2010) *Reasonable People Did Disagree: Optimism and Pessimism About the U.S. Housing Market Before the Crash*, Federal Reserve Bank of Boston Public Policy Discussion Paper No. 10-5.

Goodwin, Paul, (2002), "Integrating Management Judgment and Statistical Methods to Improve Short-Term Forecasts," *Omega*, vol. 20.

Hamilton, Dan (2011). *Forecasting with Structural Models and VARs: Relative Advantages and the Client Connection*. Manuscript.

Howrey, E. Philip, Lawrence R. Klein, and Michael McCarthy, (1974), Notes on testing the predictive performance of econometric models," International Economic Review, vol. 15(2), June.

Klein, Lawrence R. and Arthur S. Goldberger (1955), *An Econometric Model of the United States, 1929-1952*, North-Holland Publishing Company.

Krasker, William S. (1980), "The ` peso problem' in testing the efficiency of forward exchange markets,‖ *Journal of Monetary Economics*, Elsevier, vol. 6(2), pages 269-276, April.

Lucas, Robert (1976), "Econometric Policy Evaluation: A Critique", in Brunner, K.; Meltzer, A., *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, 1, New York: American Elsevier, pp. 19–46.

Lucas, Robert E. and Thomas Sargent (1978), "After the Phillips Curve" in Rational *Expectations and Econmetric Practice*, edited by Robert E. Lucas and Thomas Sargent, Minneapolis, University of Minnesota Press.

Morley, James, (2010), "The Emperor has no clothes," Macroeconomic Advisers Macro Focus, vol. 5(2), June 24.

Tetlock, Philip E. (2015), *Superforecasting: The art and science of prediction*. New York: Crown Publishing.