Forecasting Pitfalls

David F. Hendry*
Economics Department, Oxford University

October 1, 2003

Abstract

The paper highlights the pitfalls of forecasting by considering a trip by car—but with inaccurate maps in a world where roads are unexpectedly shut or opened. Behind the scenes lies a new theoretical framework for economic forecasting which acknowledges that economic models are simplified representations of an economy prone to sudden shifts. That approach explains the prevalence of forecast failure; accounts for the results of forecasting competitions; and explains the surprisingly good performance of 'consensus' forecasts. The paper offers a non-technical discussion of that theory, and draws some lighthearted conclusions on how to avoid some of the pitfalls awaiting unwary forecasters.

1 Introduction

> "Trying to predict the future is a mug's game. But ... we need to have some sort of idea of what the future's actually going to be like because we are going to have to live there, probably next week." in Douglas Adams, *Predicting the Future, The Independent on Sunday*, November 1999 (reprinted in *The Salmon of Doubt*, MacMillan, 2002).

One of my more amusing taxi rides was early one morning in Oslo, *en route* to lecture at Norges Bank on Economic Forecasting. I was asked the purpose of my journey by the driver, who instantly expressed infinite scepticism as to the ability of anyone to forecast, hence my lectures must be pointless. I gently enquired why he was out that morning, and on his expected reply 'to make a living', I asked if he had not therefore forecast that there would be customers seeking his service? The ensuing discussion revealed that he was in fact pretty good at forecasting when and where to seek trade, but had never thought of it that way. He even conceded that perhaps there was a point to forecasting. In fact, we all forecast many things many times every day, albeit probably unconsciously. Forecast errors can be expensive: missed appointments, 'accidents', and inadvisable asset purchases for individuals, possibly bankruptcy by failing to forecast a major change in the demand for its output for companies. Thus, the study of forecasting potentially offers handsome dividends—but is riddled with pitfalls.

Historically, the theory of economic forecasting has been based on two key assumptions—that the forecasting model is a good representation of the economy, and that the structure of the economy remains relatively constant. In reality, forecasting models are crude approximations, while the economy is subject to major intermittent unanticipated shifts. Thus, in practice, inaccurate forecasts have been all too common.

*I am grateful to the UK Economic and Social Research Ccouncil for financial support under grant L138251009, and to Guillaume Chevillon, Michael Clements, Michael Massmann, Grayham Mizon and participants in the Oxford Economics Faculty Workshop for helpful comments on an earlier draft. Prepared for a public lecture, Brisbane. The paper draws on earlier work with Neil R. Ericsson, and was published in *Bulletin of EU and US Inflation and Macroeconomic Analysis*, 2003, 100^{th} Issue, 65–82.

This paper highlights the pitfalls of forecasting by considering a trip by car—but alas, the maps are inaccurate and the driver lives in a world where roads are unexpectedly shut or opened, created or destroyed. We will consider the factors that influence how close the driver's forecasts of her journey time are to the outcomes. Throughout, I assume forecasts are instrumental to a decision, not an end in themselves: we rarely care if we do not have the most accurate, forecast possible, provided it is good enough for the purpose at hand. This assumption determines the form and focus of any forecast, as well as how to evaluate it.

Section 2 provides some necessary background, before section 3 spells out the problems confronting our poor driver on her journey. Then section 4 illustrates that economic data do take the form likely to disrupt 'the best laid plans of mice and men', whichever of the methods discussed in section 5 one chooses to implement. The key section is 6, which provides advice to forecasters on avoiding pitfalls. The title *Forecasting Pitfalls* is deliberately ambiguous, as success in doing so may be the key to a viable future for economic forecasting, so section 7 addresses that dimension. Section 8 concludes.

2 Background

> **Economics** *n*. "An arcane language, used by its own cognoscenti for reviewing past events in the production and distribution of wealth. There are some who would define economics as a science rather than a language; but, in the absence of any evidence that future events can be predicted by economists on the basis of fixed laws, this approach can hardly be supported by the objective lexicographer." Peter Bowler, *The Superior Person's Book of Words*, p. 37.

Behind the scenes of the paper lies a new theoretical framework of economic forecasting which acknowledges that economic models are greatly simplified representations of an economy which is dynamic and prone to sudden shifts. This newer approach has achieved a number of successes: it explains the prevalence of forecast failure; accounts for the results observed in forecasting competitions; shows why some apparently *ad hoc* approaches do well in practice; and reveals the basis for the surprisingly good performance of 'consensus' forecasts. Moreover, it suggests that some of the 'folklore' of forecasting is incorrect, including the belief that 'simple methods do best', and more surprisingly, that forecast failure is due to 'poor methods', 'inaccurate data', 'incorrect estimation', or 'data-based model selection'. My aim is to explain the underlying theory, but without the technical trappings that made it possible (see Hendry and Ericsson, 2001).

A forecast is any statement about the future, so forecasting is potentially a vast subject. There are two basic methods of forecasting. In the first, we have a crystal ball that can 'see' into the future; in the second, we extrapolate from present information. Unfortunately, as the *Washington Post* headlined in relation to the probability of a recession in the USA last summer, "Never a crystal ball when you need one".[1] Indeed, demonstrably functional examples of the first method appear unavailable to humanity, so we are forced to focus on the second, dramatically inferior, method. Here, we restrict ourselves to *systematic* forecasting rules. Even so, there exist dozens of methods of extrapolating, as well as numerous choices of what to forecast: section 5 discusses these.

In the last decade, interest in economic forecasting has increased markedly. New theories of forecasting and new methods for their evaluation have been developed, and much more empirical evidence has been acquired. Drawing on these recent developments, this paper explains some of the central issues in economic forecasting.

[1] Robert J. Samuelson, 16 June, 2001, p.A23.

As is often remarked, the problem with forecasting is that the future is uncertain. Forecast uncertainty is intrinsic to the task, but arises from two sources: one that we know is present and for which we understand the probabilities involved; and one due to factors that we do not even know exist.

> "Because of the things we don't know [that] we don't know, the future is largely unpredictable." Maxine Singer, *Thoughts of a Nonmillenarian*, p. 39.[2]

In tossing a pair of dice, the two sources might correspond to the following:

- the probability that a certain pair of numbers will appear face up on any given throw, and:
- the uncertainty arising from not knowing that the dice are loaded.

Clements and Hendry (1999) summarize the latter type of problem in relation to the above quote by Singer (1997).

Once the unpredictable has occurred, we can usually account for its effects. As a result, we can explain the past quite well. Indeed, most schoolchildren seem to learn history as if its outcomes were inevitable, rather than being a single and highly improbable sequence of events from a complicated process in which contingency has played a large role. New unpredictable events will always intrude in the future, making the future appear much more uncertain than the past. As the old joke goes, when the historian on an Oxford College Investment Committee was asked about whether it should change its strategy given recent trends in asset markets, he replied 'but the last 1000 years have been an exceptional period'. There is every likelihood that the future will see more large, unanticipated shocks – indeed, the recent collapse of the telecoms industry is a reminder that new uncertainties occur.

Statistics seeks to render 'regular' on average events that are individually unpredictable: that rendering underlies the theory of economic forecasting. For example, the age at which any individual person will die is uncertain, whereas the average age at death in a large population is highly predictable, and forms the basis of the life-insurance industry. To achieve their objectives, statisticians create models of the processes in question, check how well their models characterize the evidence, and solve the models for their average outcomes. Economic forecasting uses a similar principle: investigators develop models of the economy that seek to average over likely future 'shocks' and so deliver a useful statement about the average future. This procedure works well for 'measurable uncertainty'—that is, for the regularly occurring events that are individually unpredictable, but nevertheless average out. Singer's quote suggests that unmeasurable (or at least unmeasured) uncertainty is also important in explaining the actual uncertainty about the future.

To illustrate, imagine in 1945 predicting the number of coal miners in the UK over the period 1975–1985. Because the Oil crisis, OPEC, and the discovery of North Sea Oil and Natural Gas were all not envisaged, almost any forecast would have been woefully inaccurate. Still, the United Kingdom might have decided to protect the coal industry, leaving some forecasts quite accurate. Or, history might have taken an entirely different course, with nuclear fusion being implemented at an incredibly cheap price. Or there might have been no recoverable oil in the UK sector of the North Sea...... It is hard to imagine how anyone could conceive of the myriad possibilities that discoveries might bring. Singer alludes to this second aspect of uncertainty, which, by its very nature, is almost impossible to model, and is the source of the main pitfalls. In economics, events equivalent to earthquakes in geology seem to occur all too often, rendering both short-run and longer-term forecasts seriously inaccurate: Barrell (2001)

[2] This contrasts with the earlier view:
"The one who does not know, and does not know that he does not know;
shall remain in blackest ignorance for ever and ever."
Khajeh Abdullah Ansari, *Robai*, 1550 (approx). Translated by Prof. Hashem Pesaran, Cambridge University.

discusses six episodes of sudden shifts in the 1990s alone. UK forecasts one-year ahead made in August 1992 would have required major revisions by the end of September, on its abrupt departure from the ERM. In the next section, an analogy will help to develop this background to economic forecasting.

3 The pitfalls

> "Forecasters have sometimes described their task as similar to driving in a thick fog using only the rear-view mirror, but I think that is an understatement. To make the metaphor more exact, add misted windows, an unreliable clutch, a blindfold, and handcuffs, not to mention the unsignposted cliff a hundred meters down the road." Diane Coyle, in *Making Sense of Published Economic Forecasts.*

This section motivates the need for forecasting and clarifies several aspects of forecasting, including the uncertainty inherent in forecasting, the effects of shifts in underlying economic behavior, and the costs of making forecast errors. To highlight the problems faced in economic forecasting, I draw on a commonplace activity—travelling by car.

Planning a car-trip typically involves deciding on a destination, then consulting a (physical or cognitive) map to see how to get there. Maps seek to schematically represent connections between locations, but otherwise can seriously mislead: roads shown in red on a map are not red in reality, nor is the width of a road to scale. Nevertheless, maps that accurately portray *connections* are invaluable when planning a trip. The economic equivalent of a road map is an econometric model, which seeks to embody our best knowledge of the linkages in an economy. Evaluating a map's accuracy involves checking whether or not the roads do link up as marked on the map. Evaluating an econometric model is similar in principle, but not so easy in practice. Maps can be on the wrong scale, with too much—or too little—detail: for example, tourist maps sometimes offer a 3D impression of buildings. That maps are designed for particular purposes is reflected in the fact that models are of different types and are valuable for different purposes: if the aim is forecasting, models have to be targeted to that end.

Many factors will influence journey time, including the distance to be driven, the type of car, the road quality, the driver's skill (and personality), the traffic density, the time of day for travelling, and the weather. Given these, plus weather and traffic reports, an initial estimate of the trip's duration can be made. In many instances, that estimate will be sufficiently accurate to ensure arrival at the destination in good time. Many small factors will cause variation around this estimate: bad luck in being stopped at a sequence of traffic lights, or heavier traffic than usual, and so on. The variability around the average journey time is measured by the variance of the forecast error or, more usefully, its square root, the forecast-error standard deviation. This particular measure—the forecast error standard deviation—can be expressed as a percentage of the journey time. A large value, such as 50%, denotes an unreliable route, where a journey may well take between one-and-a-half times and half as long as expected. A small value, say 5%, is what most motorists would like, to ensure arrival within roughly ± 5 minutes on a journey of about an hour.

Similarly, with economic forecasts, a large standard deviation for a forecast entails an unreliable forecast. To illustrate, consider forecasting Gross Domestic Product (GDP): many factors influence its value, such as consumers' spending, company investment, government decisions, and exports. Over the last 200 years, per capita GDP in many developed countries has grown at about 2% per annum, with standard deviations around that mean growth of about 3% per annum. Absent a model, and the forecast interval would have to be 2% $\pm 6\%$ to ensure almost always covering the outturn. A forecast interval six times larger than the average growth rate may seem poor, but substantially smaller values may be difficult to achieve in practice. Let us now see why, drawing on our analogy to car travel.

Motorists are well aware that unexpected events can upset carefully laid plans. A pothole that punctures a tyre, a crash in the traffic ahead, or unexpectedly bad weather all can create an extended delay, as can more extreme events, such as a bridge collapsing or an earthquake. These events occur intermittently, and they can be viewed as rare realizations from a set of adverse factors. Their effect is to shift the mean of the journey time from its norm to a much larger value; and that shift is important to our explanation of forecast failure in economics. If an alternative route is not available, or the car is stuck in the jam before being alerted of the possible delay ahead, the forecast error will be large. Such effects are called location shifts, as they shift the location (mean) of the variable (here, the journey time) from one value (the norm) to another (after the adverse event), and the new value persists. Economists have sought to develop methods for handling location shifts when forecasting, and economic models can be viewed as an attempt to forecast the effects of a specific, but important, class of location shifts, namely changes in economic policy regimes.

The car analogy also highlights how minor surprises can have sudden and large effects: serious traffic delays can occur, even without an extreme event as a cause. For example, when a moderate volume of traffic passes through a roadwork that reduces the number of lanes, no problems may arise. If the traffic volume increases only somewhat, horrendous congestion may result. Such effects are classified as 'non-linearities' and 'regime switches'. Up to a point, increasing the traffic density has only a modest effect on travel time; but beyond that point, increasing the traffic density leads to serious delays, with gridlock as an extreme outcome. Similar situations arise in economics, causing difficulties for economic forecasters. Within a certain range, an exchange rate might respond roughly proportionally to a balance-of-payments deficit; but the exchange rate might suddenly nose-dive as concerns over bank solvency develop. As another example, booms and busts in the economy may derive from non-linearities that generate regime switches: the end of the dotcom investment boom is a recent example.

We now extend the travel analogy to consider additional problems in economic forecasting. Imagine that the road map is not actually correct: perhaps it is out-of-date, and some roads shown on the map no longer exist. Serious forecast errors can result from following a road that unexpectedly terminates in a field, as once happened to Ken Wallis, Jean-Francois Richard and myself in Austria alongside the Danube: at least it did not end in the river! Modern economies are sufficiently complicated that no one could hope for a 'correct' roadmap thereof: actual econometric models incorrectly omit important linkages and include ones that (unknowingly) are irrelevant. This class of problems is called model mis-specification, and it adds to the forecasters' difficulties. In particular, model mis-specification complicates calculating the likely magnitudes of forecast errors. Sometimes, no problems ensue, as when non-existent roads are not part of the route. Other times, large errors occur. However, if the mistakes themselves were 'regular', then forecasts of journey times could be adjusted: for example, if 1 in 10 roads were falsely shown, inducing an average increase in journey times of 10% (say), then unbiased forecasts could be obtained by increasing all forecasts by 10%. Thus, by itself, mis-specification is not that problematic.

Returning again to the analogy of driving by car, now consider a world that changed so fast that, by the time any map was available, some roads had vanished and new ones had appeared but were unrecorded. Route planning would become exceptionally hazardous, and large forecast errors would abound. Estimating the forecast errors' variance would itself be hazardous, as the forecast errors would depend on which roads had vanished and which replacements could be used as substitutes. No simple corrections for mis-specification could be calculated. The economic forecaster confronts a similar environment. For this and other reasons, some forecasting agencies maintain several models: the Bank of England's approach to modelling and forecasting includes using a suite of models rather than a single model. Very different forecasts from their models warn that some of the models must be at odds with reality: later outcomes may help isolate the source of the problem.

Good guides to the future are sparse when the future is not like the past. Economists refer to this situation as non-stationarity: the distributions of events change over time. A stationary series is one where those features are constant – and hence it is ahistorical in the sense that actual dates do not matter. Non-stationarity is clearly a characteristic of economies: technology, legislation, politics, and society all change over time, markedly affecting living standards, the variability of unemployment, the level of inflation, and so on, as the next section illustrates. Modern econometrics devotes considerable effort to developing models of non-stationarity, and those models fall into two distinct classes. The first includes models of regular and persistent changes (called 'stochastic trends' because there are fluctuations around the average growth). The second class includes models of 'structural breaks', which are large sudden, and usually unanticipated, changes. The models with location shifts, described above, are an important member of this class. Clements and Hendry (1998, 1999) provide a detailed analysis of economic forecasting when both types of change are present, along with the problems of mis-specification and data measurement error already mentioned. Structural breaks appear to explain why it is so hard to reduce forecast-error standard deviations: the outcome is sometimes very far from the forecast. The 1929 crash and ensuing Great Depression are a classic example of when large forecast errors occur.

While the causes often seem obvious in retrospect, many forecast errors result from unanticipated events, both large and small. Thus, Diane Coyle (2001) correctly chastised Nigel Lawson when Chancellor of the Exchequer for blaming his mistakes on Her Majesty's Treasury for mis-forecasting over the late 1980s. Even today, the economics profession is not unanimous on which factors induced those forecast errors, although many believe that financial deregulation was responsible, as it led to an unexpectedly sharp reduction in credit rationing. Unanticipated events, like the sudden rise of oil prices in 2000, frequently occur; and they thereby disrupt forecast accuracy. However, not all methods of forecasting are equally disrupted: section 5 discusses how forecasting is done, and notes which methods might be less prone to pitfalls. First we need to see the source of the problem, manifest in almost any economic time series.

4 Economic data

> "The scientists have their machines, while the economists are still waiting for their data..... Econometrics is an attempt to compensate for the glaring weakness of the data base available to us by the widest possible use of more and more sophisticated statistical techniques", Wassily Leontief (1971), in 'Theoretical assumptions and nonobserved facts', *American Economic Review*.

To forecast the future values of any economic time series, one needs to know how such series behave in general. There are ten general features of many economic time series:

(i) very smooth and 'exploding' in their original levels.

Figure 1 illustrates (i), in panels of four, denoted a, b, d, c, respectively clockwise from the top left, commencing with the smoothest through to the least smooth. These four graphs show money, prices, constant-price (real) GDP and employment in the UK, over 1870-1990. As can be seen, the first has grown almost 600-fold over the century and a quarter, the second 55-fold, and the third just over 10 fold. Since total employment has barely doubled, per capita output has risen 5-fold: 2.5% pa has massive long-run effects – just imagine living on 20% of your present income. Equally, pity the poor

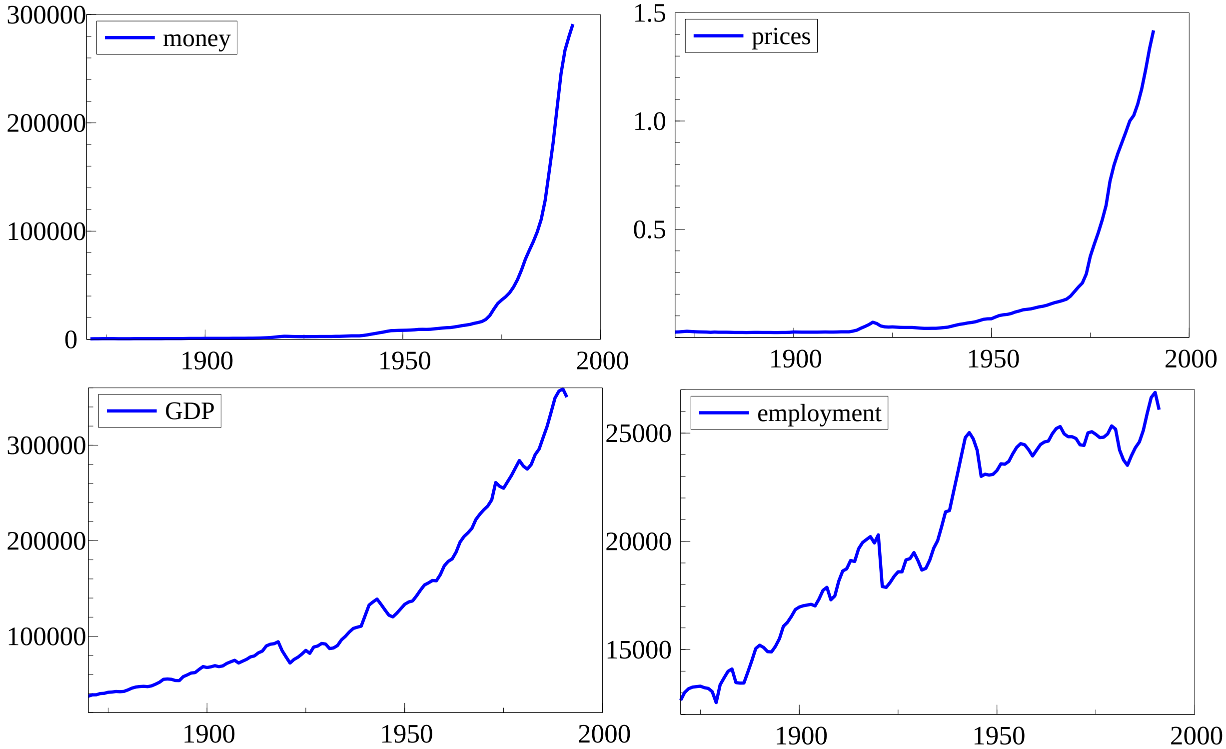

Figure 1 Levels of money, prices, real GDP and employment.

forecaster: in the century from 1870 to 1970, the price level rose just over 7-fold, then rose another 7-fold over the next twenty years!

(ii) Relatively smooth when graphed in logs;
(iii) usually trending linearly in log levels.

Figure 2 shows (ii) and (iii), using the logs of the same four variables. The trends are more nearly linear, albeit perhaps increasing over time for the first three. 'Business cycle' deviations around the trend are clearly visible in the last two panels (and show up in the 5-year averages noted below).

(iv) Somewhat erratic when graphed in (log) changes;
(v) but such differences are less trending.

Figure 3 illustrates (iv) and (v) for the annual changes of the same four variables.

(vi) Smoother for multi-period differencing;
(vii) but rather erratic in second differences.

Figure 3 also illustrates (vi); whereas figure 4 illustrates (vii).

(viii) Prone to jumps in location (and trend);
(ix) so means and variances of data not constant over time;

Figure 5 illustrates both (viii) and (ix) for UK unemployment, as did figure 3 for the annual changes in the other four variables.

Figure 6 shows the consequence of attempts to forecast the start of the interwar period, and the early years of the Thatcher experiment. Panels a and c show the ex ante forecasts, whereas b and d add in the outturns that eventuated – clearly, both sets of forecasts failed badly. This is included to emphasize that forecast failure is easy to generate, not to suggest that anyone might in fact have foreseen the large rises in unemployment witnessed in both periods.

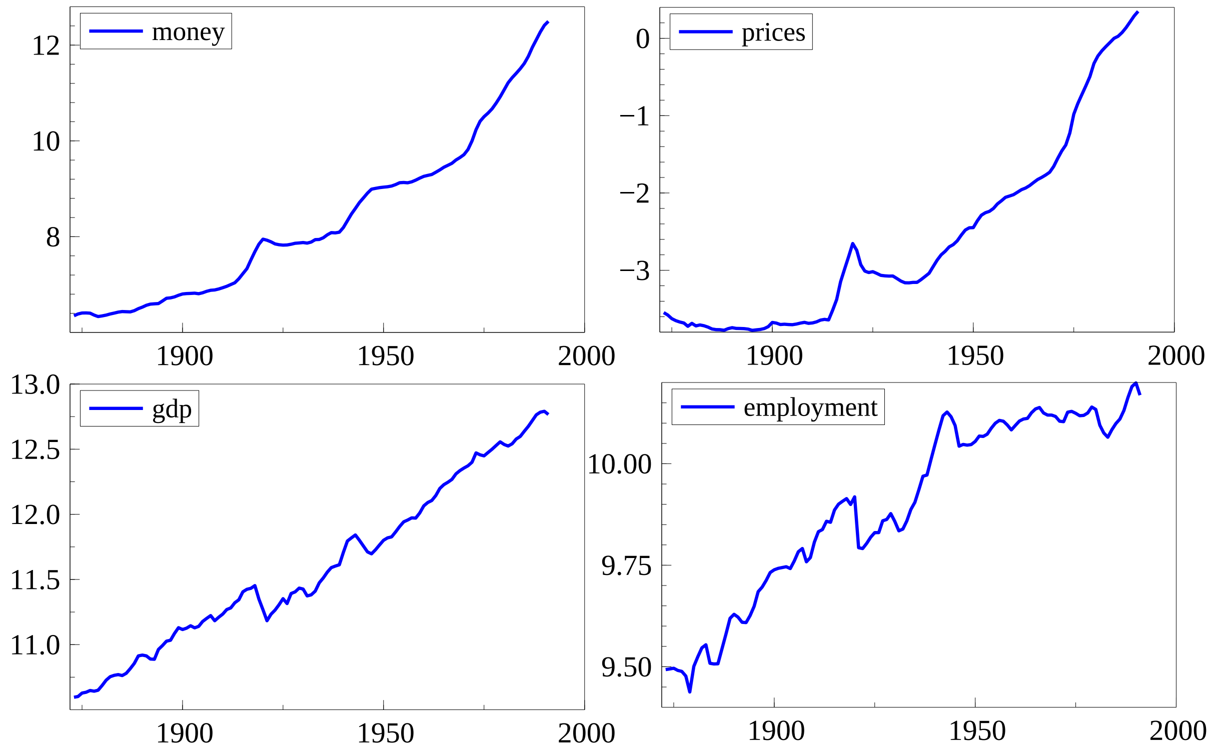

Figure 2 Logs of money, prices, real GDP and employment.

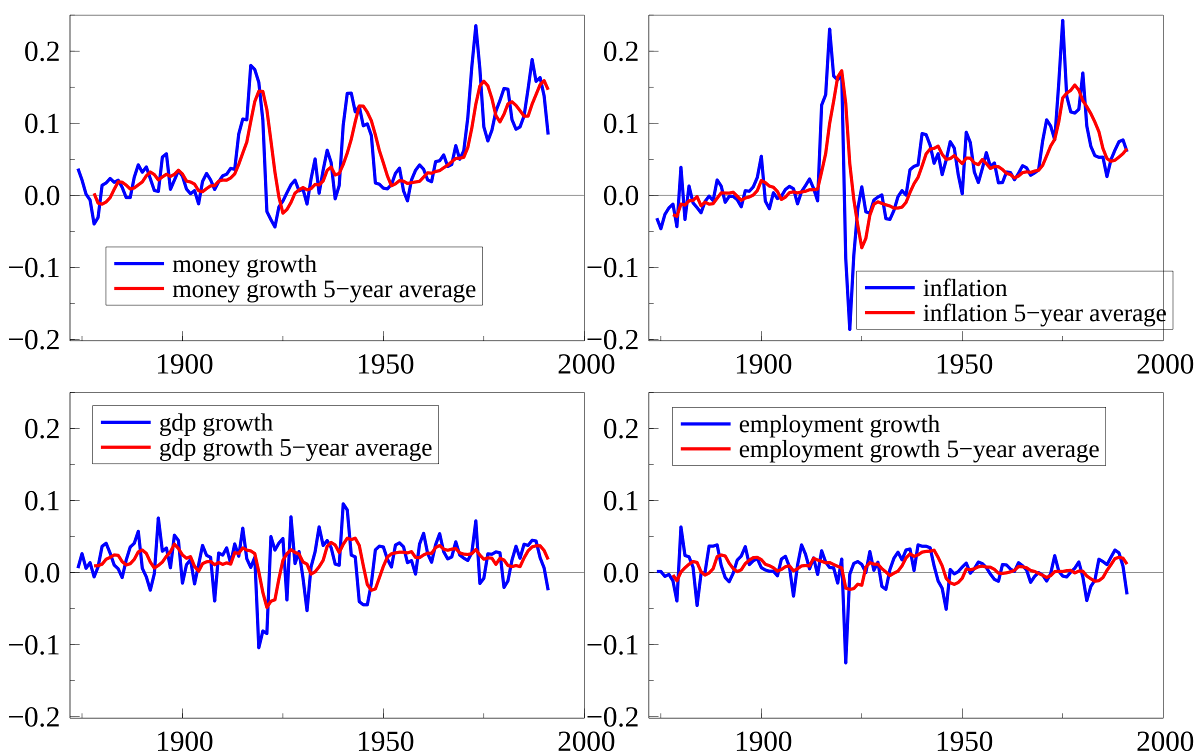

Figure 3 Annual changes in money, prices, GDP and employment.

(x) high sample correlations with other series.

One must allow for all ten properties when both model building and forecasting, but the most important for the latter are the location shifts. The changes in the locations, trends and variabilities of all these series over time are all examples of non-stationarity. We have now seen the problem for forecasting: the

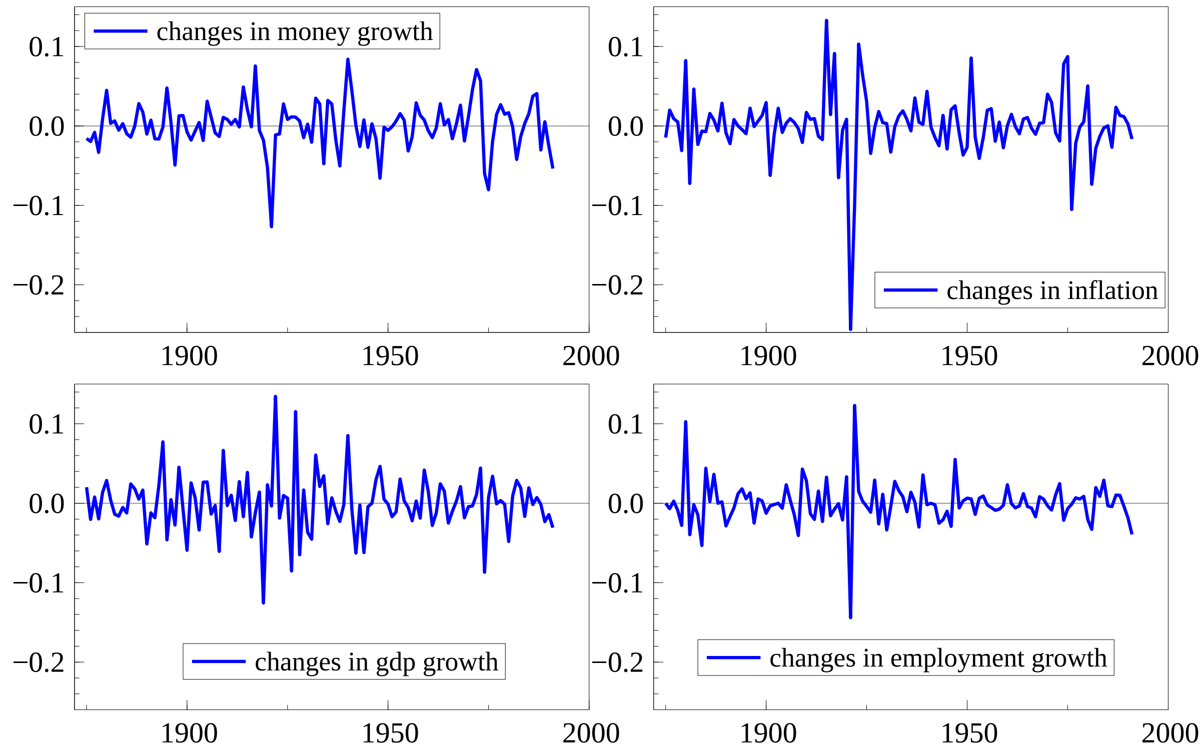

Figure 4 Changes in growth rates of money, prices, GDP and employment.

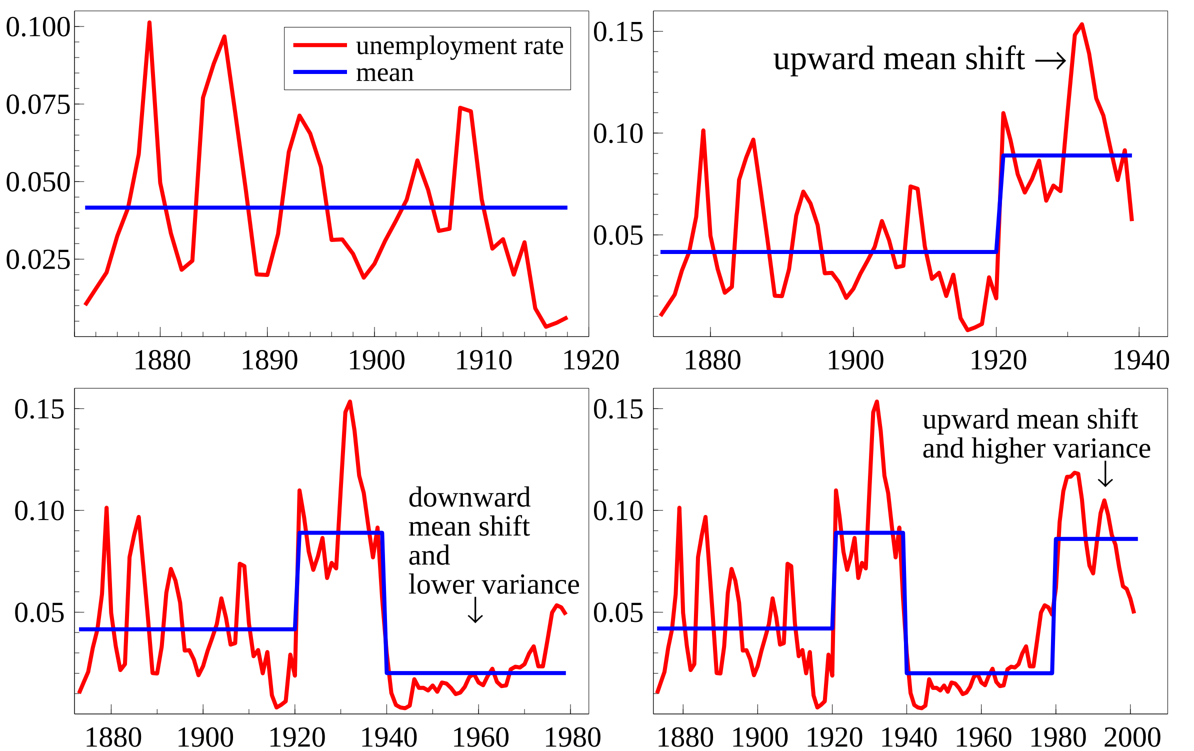

Figure 5 Shifts in UK unemployment.

future is not like the past at the phenomenological level. Success will entail modelling the evolution, and the breaks....

Figure 6 Forecasting two episodes of UK unemployment.

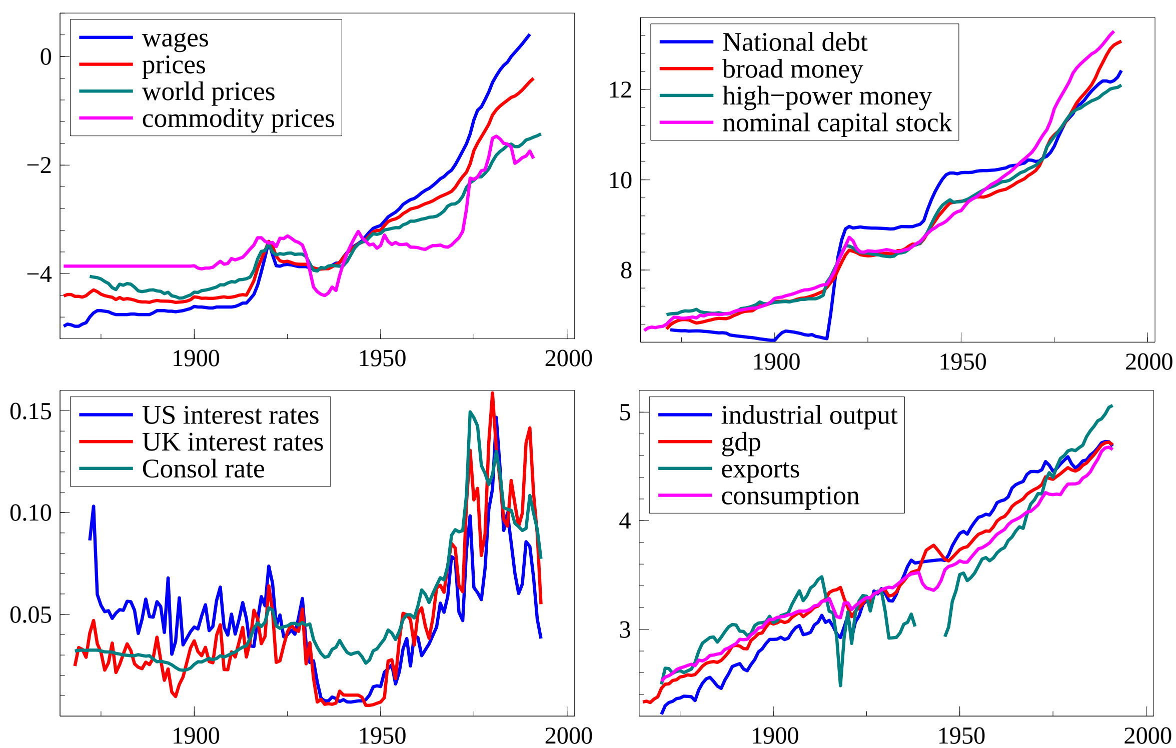

Figure 7 Price, nominal asset, interest rate and physical output measures.

5 Forecasting methods

> "Here is Edward Bear, coming downstairs now, bump, bump, bump, on the back of his head, behind Christopher Robin. It is, as far as he knows, the only way of coming downstairs, but sometime he feels there really is another way, if only he could stop bumping for a moment and think of it." From *Winnie-the-Pooh* by A.A. Milne (1926), Methuen and Co. Ltd.

That is how forecasters feel most of the time.[3] So let us step back and look at some possible methods.

Since forecasts are a human construction, many methods are available for making them. However, the success of any approach requires that:

(a) there are regularities to be captured;
(b) those regularities are informative about the future;
(c) the proposed method captures such regularities; and
(d) excludes non-regularities.

The first and second are properties of the economy that may, or may not, be present; the third and fourth are properties of the forecasting method. The first seems clearly true, however, since today looks a lot like yesterday even if some things have changed, and unanticipated events occur. The second is less obvious, but we must assume some persistence. Ensuring the third without violating the fourth can be difficult in practice, and helps determine which methods will prove successful in practice.

Widely used methods of forecasting include:

- guessing,
- 'rules of thumb' or 'informal models',
- naive extrapolation,
- leading indicators,
- surveys,
- time-series models, and
- econometric systems.

In order to produce good forecasts, the first only requires luck; the second an intuitive 'feel' for the variables to be forecast (flying by the seat of your pants – rather than your brain?)[4]; the third merely needs that present trends continue (see Irvine Fisher below: do you have any friends who bought their house at the peak of the last boom?); the fourth that what led yesterday will still lead today (what football team do you support that achieves that?); the fifth that the individuals asked form serious plans and truthfully reveal them (noting mis-predictions of election outcomes); sixth that the evolution of the process is relatively regular (see the graphs above); and the seventh that....
well, rather a lot of assumptions are satisfied.
Nevertheless, formal econometric systems of national economies consolidate existing empirical and theoretical knowledge of how economies function, provide a framework for a progressive research strategy, and help explain their own failures as well as provide forecasts. Economic forecasting based on econometric and time-series models are our primary methods.

In the earlier analogy, the forecast of the journey duration by car was based on the distance (measured on the map) and average speed (dependent on road quality and traffic conditions), mediated by any likely special factors. Such an approach is inherently 'causal' or structural: the effects that are likely to

---

[3] I first read this quote as a lead to Nerlove and Wallis (1966).

[4] This is clearly unfair on the immense skills of, say, Alan Greenspan in absorbing a vast array of disparate information to 'foresee' local economic trends.

influence the forecast are quantified in the context of a model. Its advantage is an accurate forecast when all the factors are known: its drawback is that if the map is inaccurate, so may be the forecasts. Econometric models in effect are of this type. However, a poor model need not lead to inaccurate forecasts. For example, the Ptolemaic system is hardly a realistic model of planetary motions, but nonetheless it predicted lunar eclipses reasonably well (see e.g., Harré, 1985). A map likewise may always have been incorrect, so forecasts for trip times are wrong, but no worse than usual. Alternatively, a map may have become incorrect because a bridge was closed: that sort of event leads to 'forecast failure', where a previously well-performing approach suddenly does much worse.

For repeated journeys, an alternative forecasting method is available, which does not rely on a map at all: forecast the journey's time by how long it took on the previous occasion. If the road network does not change greatly, and if no special factors disturbed the previous journey, then such a pure extrapolation could be quite accurate, *and it is independent of how good or bad the map is*. If the relation between the map and reality is poor, or if their relationship continually alters, then extrapolating from the previous outcome may be as accurate as one can achieve. In essence, that is what time-series models do: the simplest case, a random walk, precisely forecasts the previous value (more general models equilibrate towards a mean).

These situations parallel those faced by economists when forecasting: models are far from accurate representations of the economy, and their accuracy can change abruptly. A model that closely represents the economy is a good basis for forecasting only if the future economy is 'close' to the current one. The culprit of forecast failure is not rapid evolution *per se*, as commonly occurs in technology, but abrupt changes, as may derive from legislative and political developments. In such a setting, a model structured around past economic behavior may not be the best forecasting device available. Rather, a 'naive predictor' that simply extrapolates from previous outcomes may be more successful. If the economy stays on track, such extrapolative forecasts can be reasonably accurate. If the economy crashes, any extrapolative forecast will miss the crash itself, as typically will other types of forecasts. However, after the crash has occurred, the extrapolative forecast tends to come back on track, whereas many other types of forecasting models continue to mis-forecast systematically. Regretably, some extrapolative forecasts have little to commend them: a stockbroker who continually chants the mantra that 'the bull market will go on', can become famous before she gets it badly wrong, then return rapidly to guru status by repeatedly intoning 'the bear market will go on'.....

Nevertheless, to be competitive, forecasts from economic models must adjust rapidly to major changes, even when those changes were themselves unpredictable. A serious criticism of historical forecast errors is how *systematic* they have been—not so much that the forecast errors were sometimes large. For the United Kingdom, long sequences of under-prediction and over-prediction have been recorded: Chancellor Lawson may have been right to complain about those. Still, while avoiding such errors would be an improvement over having them, simply adapting does not actually provide knowledge of the future, it merely eliminates tracking error.

Reverting to our analogy, modern technology has developed instruments for informing drivers about congestion on the road ahead, so drivers could anticipate looming problems by selecting a diversion before becoming stuck in a traffic jam, whatever its source. In effect, the driver temporarily alters the map, pencilling in the blockage ahead. The extent to which this will improve forecasts of journey time will depend on how accurately drivers' can estimate the effects of the congestion spilling over on to alternative routes etc. Such flexibility enhances the probability of correctly forecasting the overall journey time. Of course, really major shocks (such as an electrical storm) might disrupt even that provision of information. The analogue in economics is the ability to predict the changes in the economic system that will result from the structural changes elsewhere.

This leaves the forecaster with two possibilities: avoid the pitfalls, or forecast them. We consider these in turn.

6 Advice to forecasters on avoiding pitfalls

> "Oddly, the industry that is the primary engine of this incredible pace of change—the computer industry—turns out to be rather bad at predicting the future itself. There are two things in particular that it failed to see: one was the coming of the Internet, ...; the other was the end of the century." Douglas Adams (2002).

Adams goes on to quote (from *The Experts Speak – And Get It Wrong* by Christopher Cerf and Victor Navasky) a range of famous authoritative predictions that 'turned out to be wonderfully wrong, almost immediately', including Irving Fisher, October 17, 1929: 'stocks have reached what looks like a permanently high plateau'; and in 1897, Lord Kelvin (he who estimated that the sun could not be more than 100 million years old – without a confidence interval of ± 4 billion) stated that 'radio has no future'; as well as IBM's chairman (Thomas J. Watson) in 1943 that the world will never need more than five computers: and the President of Digital Equipment Corporation (DEC) in 1977: 'There is no reason for any individual to have a computer in their home'.

In this section we will describe how to avoid the six main pitfalls from the viewpoint of the forecaster (rather than of the user): *caveat emptor*.

The quotes from Adams reveal the first pitfall, and hence how to avoid it:

never be precise.

It is essential when forecasting to be imprecise about the magnitude, and preferably the timing, of what is forecast. The great seers of the past were well aware of the need for imprecision, indeed, preferably ambiguity, so any outcome would be consistent with their forecast or prophecy: viz., the oracle at Delphi (now, incidentally, embodied in a method of forecasting – you got it at once – the Delphic method). [5]

The second pitfall evident in the quotes can be avoided by:

always make more than one statement.

If Kelvin had also remarked at some other time and place, 'radio has a great future', he would not be castigated now. Tell different audiences different messages, and some will be able to confirm that you were right: is that what the internal e-mails at stockbrokers were really for – a two-way bet?

To avoid the third pitfall:

never make unconditional statements.

Instead, try 'radio has no future unless ...' (the price of sets falls by 95%, the quality of broadcasts triples, and leisure time doubles to suggest a few potential qualifiers....). Or as La Palice (1470-1525) expressed the matter, 'Un quart d'heure avant sa mort, il etait encore vivant.'

Adams ends his article with the remark: "Predicting the future is a mug's game, but any game is improved when you can actually keep the score". But that is actually my fourth pitfall – to avoid it:

never allow post-event evaluation:

—unless you can claim complete success. As FDR was told when re-visiting a town during his whistle-stop re-election campaign on wishing to show that he had achieved in office what he had promised: 'Mr. President, you must deny you have ever been here before'. In fact, Adams's genius is to ask "will the e-commerce bubble burst?" – in November 1999 remember – leaving us admiring his perspicacity, not

[5] Sen (1986) concludes that 'The imprecisions of economic predictions have to go hand in hand with a precise recognition of these imprecisions'.

a bland "the e-commerce bubble will burst in January 2000", which could, of course, have been badly wrong.

Alternatively, you may have to rely on Hendry's first forecasting theorem:

any forecast can be made to look good or bad by an appropriate choice of metric.

You may already know this method of presentation – the Russian and American Ambassadors ran in a race: the American came second and the Russian came second last – but does not mention that only those two took part.... Figure 8 illustrates this theorem.[6]

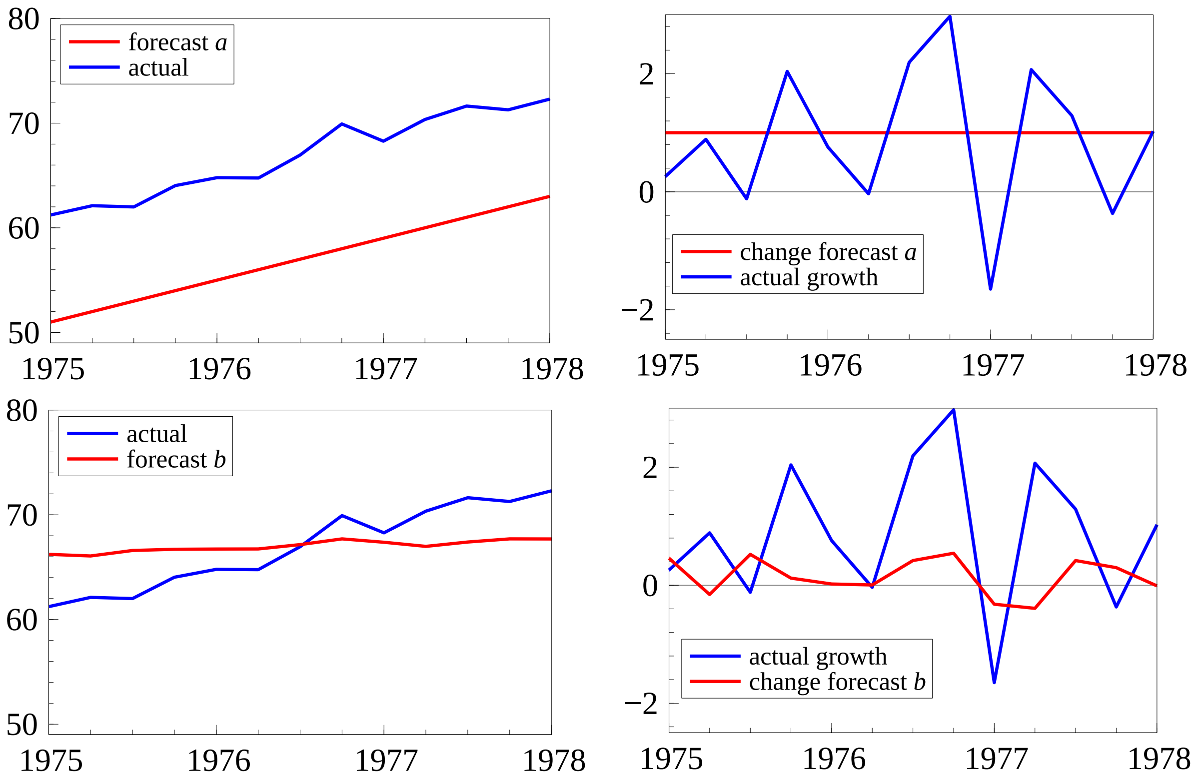

Figure 8 Who wins: a or b?.

In terms of accuracy, for forecasting levels, *b* wins; but for growth, *a* wins:

> "..the Dodo suddenly called out 'the race is over!' and they all crowded round it panting, and asking 'But who has won?' This question the Dodo could not answer without a great deal of thought,... At last the Dodo said '*Everyone* has won, and *all* must have prizes.' "
> Lewis Carroll, *Alice's Adventures in Wonderland*, 1865 (p.33).

Next:

don't stand out from the crowd.

Diane Coyle was the presenter for many years of the Golden Guru award for each year's best forecast of the UK misery index (measuring inflation, unemployment, and economic growth). In Coyle (2001), she notes that forecasters have herded together on these three key measures, with the Golden Guru usually being won by only a narrow margin, and with very few repeat winners.[7]

[6] The theorem assumes—realistically—that forecast errors are not everywhere zero.

[7] Coyle proceeds to apologize for perpetrating in the Golden Guru contest the notion that the difference between 2.2% and 2.5% measured GDP growth in a given year is at all meaningful. However, Coyle also stresses that journalists could improve their presentations of forecasts by avoiding spurious precision and by explaining the uncertainties involved.

Finally:

never predict large shifts.

This is similar to the preceding advice, but emphasizes the need to avoid forecasting at all – just say where we are. The 'bus-stop game' in Hendry (2001a) illustrates why that strategy works. Such advice may seem at odds with the main problem confronting economic forecasting, namely large shifts, but assumes we are not be able to forecast the big changes anyway. Figure 9 illustrates the apparent effectiveness of forecasts that never have any 'future' content. I have marked what the forecast for 1947–48 might have been, given a 'history repeats itself' notion based on the experience in 1920–21: 40% government deficit-GNP ratio removed in 3 years; troops demobilized; European trade collapses; pre-war parity sought—and output crashes in the ensuing post-war slump. No amount of missing the target by forecasting the previous value would ever overwhelm one huge error like X. This suggests that the valuation criterion itself is inappropriate, a pitfall discussed earlier.

Nevertheless, perhaps it is worth trying to anticipate events....

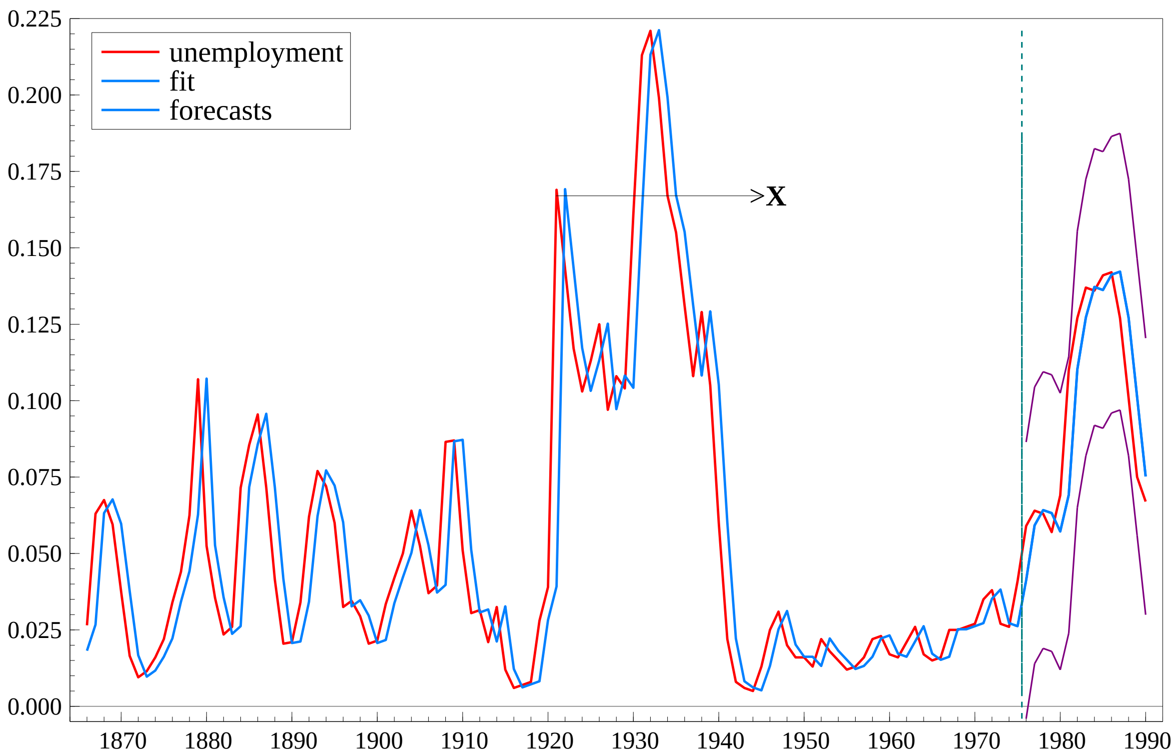

Figure 9 'Forecasts' of the unemployment rate.

7 Forecasting the pitfalls

> "Toad-stone: popular name for the mineral *bufonite*, supposed to be a natural concretion found in the head of the common toad – and predicted to offer protection from poison". From Robert Hunter, *Encyclopaedic Dictionary*, 1894, reported in *Forgotten English*, by Jeffrey Kacirk, 2001.[8]
>
> "Sweet are the uses of adversity, which like the toad, ugly and venomous toad wears yet a precious jewel in his head". (in Shakespeare's *As You Like It*.)

[8]Toad fairs were common in the early 1800s in Dorset: the toadman made predictions based on the patterns a handful of dried toad bones made when they fell.

A regime shift would reveal significant predictive failure here, namely wear the stone and drink poison—shades of Voltaire: "An incantation can kill a flock of sheep, if accompanied by a sufficiently large dose of arsenic". In economics, policy-regime shifts are among the most important members of this class, although 'economic crises' follow a related close second. Forecasts made before a break and in ignorance of its impending occurrence are bound to suffer its full effects. Consequently, attempts to forecast future 'rare events', especially those which entail location shifts, could pay handsome dividends. Thus, we turn to the second sense of our title: how can we forecast the pitfalls that might confront us?

Sen (1986) suggests two main difficulties confronting successful economic 'prediction': choice and interaction. The former entails that different economic agents can choose differently in near identical situations, and the latter that interactions between millions of decisions can lead to unexpected outcomes.

Several possible approaches are available for accounting for rare events in modelling; rather fewer for forecasting rare events directly, thereby avoiding the resulting large forecast errors. Note that modelling rare events is particularly onerous because there are so few observations on the rare events themselves. Models of past shocks fall into three groups, where the events are modelled:

(a) as one-off aberations (usually removed by indicator variables: see e.g., Hendry, 2001b);
(b) by adding observable factors to account for the shifts (as with financial liberalization: see e.g., Muellbauer, 1994); and
(c) as part of more regular 'regime shifts' (as in Markov switching: see e.g., Hamilton, 1989, and Krolzig, 1997).

A natural step in approach (a) is to document the rare events, to see how often they occurred and how big they were historically. Then one could seek to modify the forecasts accordingly. Two forecasts might be produced, one accounting for, and the other neglecting, some large potential shock, such as another oil crisis. Each person can then select the forecast that he or she deems most likely. Or, the two forecasts can be averaged, weighted by how likely the shock (or its absence) is believed to be. To illustrate that by our analogy, one might expect a 4-hour journey without delays and a 6-hour journey (equally likely) with delays, and so schedule 5 hours total in order to allow for both possible outcomes. Note that actual outcomes always differ from the weighted forecast in this case—the actual journey time is either 4 hours or 6 hours and never the 5 hours planned. Whether or not such discrepancies between forecasts and outcomes matter depends on the costs of forecast errors.

Route (b) is the most satisfactory in intellectual terms, as insights are gleaned into how economies function—but only after the event. To be of value in forecasting, all new shifts would need to be forecast, as well as prior explanations provided.

As an example of (c), Osborn, Sensier and Simpson (2001) treat recessions as sufficiently rare that adding 'leading indicators' to a regime-shift model might help in their prediction, and claim some success in practice.

Unfortunately, many rare events are not part of a sequence like business cycles on which a small sample of observations is available: examples include the 1984 Banking Act and the 1986 Building Societies Act in the UK, both of which had massive unforseen consequences which later could be modelled as in (b). Even so, 'rare events' should be partly predictable since they have causes, and some of those causes may be discernible in advance. Environmental rare events such as hurricanes, earthquakes and volcanic eruptions usually issue 'advance signs' that are harbingers of impending problems. Recent advances in (say) earth sciences for forecasting volcanic eruptions seem to have considered both leading indicators (e.g., the temperature of the vented steam, where rises indicate increased activity), and 'pressure analysis', which uses low-frequency sounds as an indicator of internal pressure that will lead to an explosion (the lower the sounds, as in an organ-pipe, the greater the pressure). If economic counterparts have corresponding attributes, then a search for ‘early-warning signals’ is merited. One route may be by monitoring high-frequency data, which should reflect location shifts much sooner in real time, although there is the corresponding drawback that such data tend to be noisier. Nevertheless, ‘early-warning’ signals merit serious consideration, so high-frequency readings on the state of the economy must play a role in this area.

7.1 Forecasting crises

> “...whereas the weather forecast has no effect on the weather, the economic forecast may often affect the economy...” Sir John Mason, in discussion of Hendry (1986).

An extreme pitfall is a crisis – an event which occurs remarkably frequently, given the examples of crises in the 1990s that Barrell (2001) discusses. Say we try to ‘predict’ their impending occurrence, then we face the above problem, first discussed by Morgenstern (1928), and partially resolved by Marget (1929):
Monday: garage mechanic ‘Your brakes are about to fail’; driver ‘Better fix them’;
Wednesday, back at garage: driver – ‘Why did you change my brake pads: yesterday’s emergency stop was fine?’.

We all see the silliness of this interchange—until it comes to predicting crises. If we succeeded, and altered policy such that no crisis eventuated, would that be a forecast failure? We could, of course, emphasize the conditional nature of the crisis forecast, implicit in the mechanic’s comment (‘unless you fix the brakes, you will have an accident’), and most Central Banks PR certainly tries to (‘we averted the recession by lowering interest rates’). Nevertheless, the combination of failure to forecast crises that later occur (East Asia in 1997), and forecasting ones that do not happen conspire to overwhelm the few successful positive predictions (no points are awarded for the mantra ‘no crisis ahead’ even when it is correct). Most such attempts in fact focus on ‘pressure analysis’, so can claim a good pedigree, but as yet have not delivered much in the way of ‘best in show’ accolades.

8 Conclusion

> “When weather forecasters go awry, they get a new super-computer; when economists mis-forecast, we get our budgets cut”. DFH (originally quoted following the unpredicted 1987 storm).

There are many pitfalls in all forms of forecasting, most well exemplified by economic forecasting. Many are unavoidable, because they are due to events outside the purview of almost anybody: September 11, 2001 is a tragic example. However, a recent general theory of economic forecasting provides a framework for analyzing methods in the context of a world where such unanticipated location shocks occur, and models are inexact representations of the economy. That theory helps explain many historical episodes of forecast failure, and suggests genuine ways of avoiding systematic forecasting errors, none of which relies on the lighthearted advice to forecasters provided above.

The car journey analogy highlighted many of the pertinent issues, and suggested why some methods might work even when maps were unreliable guides to an ever-changing route network. If econometric models are to succeed as forecasting devices, they must be adaptive, and could try to forecast some of the pitfalls that lie ahead.

References

Adams, D. N. (2002). *The Salmon of Doubt*. London: MacMillan.

Barrell, R. (2001). Forecasting the world economy. in Hendry, and Ericsson (2001), pp. 149–169.

Clements, M. P., and Hendry, D. F. (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.

Clements, M. P., and Hendry, D. F. (1999). *Forecasting Non-stationary Economic Time Series*. Cambridge, Mass.: MIT Press.

Coyle, D. (2001). Making sense of published economic forecasts. in Hendry, and Ericsson (2001), pp. 54–67.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**, 357–384.

Harré, R. (1985). *The Philosophies of Science*. Oxford: Oxford University Press.

Hendry, D. F. (1986). The role of prediction in evaluating econometric models. In *Proceedings of the Royal Society*, Vol. A407, pp. 25–33.

Hendry, D. F. (2001a). How economists forecast. in Hendry, and Ericsson (2001), pp. 15–41.

Hendry, D. F. (2001b). Modelling UK inflation, 1875–1991. *Journal of Applied Econometrics*, **16**, 255–275.

Hendry, D. F., and Ericsson, N. R. (eds.)(2001). *Understanding Economic Forecasts*. Cambridge, Mass.: MIT Press.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference and Application to Business Cycle Analysis*. Berlin: Springer–Verlag. Lecture Notes in Economics and Mathematical Systems, 454.

Leontief, W. (1971). Theoretical assumptions and nonobserved facts. *American Economic Review*, **61**, 1–7.

Marget, A. W. (1929). Morgenstern on the methodology of economic forecasting. *Journal of Political Economy*, **37**, 312–339.

Milne, A. A. (1926). *Winnie-the-Pooh*. London: Methuen and Co. Ltd.

Morgenstern, O. (1928). *Wirtschaftsprognose: eine Untersuchung ihrer Voraussetzungen und Möglichkeiten*. Vienna: Julius Springer.

Muellbauer, J. N. J. (1994). The assessment: Consumer expenditure. *Oxford Review of Economic Policy*, **10**, 1–41.

Nerlove, M., and Wallis, K. F. (1966). Use of the Durbin-Watson statistic in inappropriate situations. *Econometrica*, **34**, 235–238.

Osborn, D. R., Sensier, M., and Simpson, P. W. (2001). Forecasting and the UK business cycle. in Hendry, and Ericsson (2001), pp. 104–123.

Sen, A. K. (1986). Prediction and economic theory. In *Proceedings of the Royal Society*, Vol. A407, pp. 3–23.

Singer, M. (1997). Thoughts of a nonmillenarian. *Bulletin of the American Academy of Arts and Sciences*, *51*(2), 36–51.