CHAPTER ONE

An Overview of Economic Forecasting

Michael P. Clements and David F. Hendry

1.1. Introduction

The chapters of this *Companion* address a wide range of issues and approaches to economic forecasting. We overview its material in terms of 11 key questions that might fruitfully be asked of any practical venture, but here specifically concern forecasting.[1]

1. What is a forecast?
2. What can be forecast?
3. How confident can we be in forecasts?
4. How is forecasting done generally?
5. How is forecasting done by economists?
6. How can one measure the success or failure of forecasts?
7. How does one analyze the properties of forecasting methods?
8. What special data features matter most?
9. What are the main problems?
10. Do these problems have potential solutions?
11. What is the future of economic forecasting?

Some of these questions can be dealt with quickly and are not specifically addressed by any of the contributors, while others provide the subject matter for several chapters. Moreover, several chapters relate to a number of the questions. To set the scene, the first part of this overview provides preliminary remarks in response to each question. The second part of the overview then summarizes each chapter in relation to these 11 questions. It is hoped that organizing the *Companion* in this way will be helpful to the nonspecialist reader. The questions we ask, and the answers that are provided here, require no foreknowledge of the subject, and should allow the reader to dip in at will, without approaching the material in any particular order.

1.1.1. What is a forecast?

A forecast is any statement about the future. Such statements may be well founded, or lack any sound basis; they may be accurate or inaccurate on any given occasion, or on average; precise or imprecise; and model-based or informal. Forecasts are produced by methods as diverse as well-tested systems of hundreds of econometrically-estimated equations, through to methods which have scarcely any observable basis (such as forecasts of the 2003 Derby winner made on 31 December 2000 – before the entrants are even known). Thus, forecasting is potentially a vast subject. Historically, almost every conceivable method has been tried, with the legacy that there are in excess of 36 different words in English for the activity of "foretelling," in a broad sense, what the future might bring forth.

1.1.2. What can be forecast?

Since it is merely a statement about the future, anything can be forecast, ranging from next month's rate of consumer price inflation, tomorrow's weather patterns, the average rise in sea levels by the end of the third millennium, through the earth's population at the same date, to the value of the Dow Jones index at the start of 2010. We are not claiming that the resulting forecasts are necessarily *useful* in any sense: consider, for example, a forecast that the first Extra Terrestrial to land on Earth will be six meters tall, blue, and will arrive in New York on July 4th, 2276 to celebrate the quincentenary of the U.S.A. Even if such a claim were to prove correct, it would be of no value for the next 250 years; and of course, it is anyway essentially certain to be incorrect.

1.1.3. How confident can we be in forecasts?

Clearly, our confidence will depend on how well-based the forecasts are. Mere guesses should not inspire great confidence; forecasts from well-tested approaches may be viewed more hopefully. Unfortunately, even the latter is not enough. The trouble is that the future is uncertain. There are two distinct senses in which this applies, expressed by Maxine Singer in her "Thoughts of a Nonmillenarian" (*Bulletin of the American Academy of Arts and Sciences*, 1997, **51**, 2, p. 39) as:

> Because of the things we don't know [that] we don't know, the future is largely unpredictable. But some developments can be anticipated, or at least imagined, on the basis of existing knowledge.

Little can be done in advance about uncertainty stemming from "things we don't know we don't know." However, the apparent randomness of outcomes within the realms we do understand, which we will call "measurable uncertainty," can often be usefully communicated to the user of a forecast. This usually takes the form of a forecast interval around a "point" forecast, the latter then being viewed as the central tendency, or "most likely" outcome. For example, the statement that "the moon is *exactly* 5,000 miles away" is very precise (but wholly inaccurate), and taken literally would be associated with a forecast interval of length zero. On the other hand, the statement that "the moon lies between 1,000 and 1 billion miles away" is correct, but very imprecise, having a huge forecast interval. More sophisticated presentations of measurable uncertainty include density forecasts; namely, estimates of the probability distribution of the possible future outcomes. The Bank of England tries to present its *Inflation Report* forecasts in this last form, using a "fan chart" where uncertainty fans out into the future in ever wider bands of lighter color (unfortunately, they chose red for the inflation forecasts and green for output, so these were called "rivers of blood" and "rivers of bile" respectively: see Coyle (2001) for an amusing discussion.

1.1.4. How is forecasting done generally?

There are many ways of making forecasts. These include formal model-based statistical analyses, statistical analyses not based on parametric models, informal "back-of-the-envelope" calculations, simple extrapolations, "leading indicators," "chartist" approaches, "informed judgment," tossing a coin, guessing, and "hunches." It is difficult to judge the frequency with which each of these methods is used in practice, but most occur regularly in our everyday lives. In earlier times, tea leaves, entrails, movements of the stars, etc., all were tried – without great success so far as we can ascertain – but some (such as astrology) remain in use today. This book, for better or worse, will focus on formal statistical approaches.

1.1.5. How is forecasting done by economists?

In economics, methods of forecasting include:

1 guessing, "rules of thumb," or "informal models";
2 expert judgment;
3 extrapolation;
4 leading indicators;
5 surveys;
6 time-series models; and
7 econometric systems.

Guessing and related methods only rely on luck. While that may be a minimal assumption compared to the other methods we will discuss, guessing is not generally a useful method: "good" guesses are often reported, and bad ones quietly ignored; and the uncertainty attaching to each guess is usually impossible to evaluate in advance. If many individuals guess, some will be "right" by chance, but that hardly justifies the approach (otherwise economists will start producing thousands of forecasts and claiming success whenever any one of them is accurate).

Expert judgment is usually part of a forecasting approach, but lacks validation when it is the sole component, even if at any point in time, some "oracle" manages to have forecasted accurately. Unfortunately, systematic success proves elusive even to experts, and no one can predict which oracle will be successful next (note the recent advice to ignore past performance when choosing a fund manager!).

Extrapolation is fine so long as the tendencies persist, but that is itself doubtful: the telling feature is that different extrapolators are used at different points in time. Moreover, forecasts are most useful when they predict changes in tendencies, which extrapolative methods are likely to miss. Many a person has bought shares, or a house, at the peak of a boom. . . .

Forecasting based on *leading indicators* requires a stable relationship between the variables that "lead" and the variables that are "led." When the reasons for the lead are clear, as with orders preceding production, then the indicators may be useful, but otherwise are liable to give misleading information. Even for such "obvious" leading indicators as housing starts leading to completed dwellings, the record is poor (because the delay can narrow and widen dramatically in housing market booms and busts – or with very severe weather).

Surveys of consumers and businesses can be informative about future events, but rely on plans being realized. Again we see "many a slip twixt cup and lip": adverse changes in the business "climate" can induce radical revisions to plans, since it is less costly to revise a plan than the actuality.

Time-series models which describe the historical patterns of data are popular forecasting methods, and have often been found to be competitive relative to econometric systems of equations (particularly in their multivariate forms). These are the work-horse of the forecasting industry, and several chapters below explain and analyze variants thereof. But like all other methods, they focus on "measurable uncertainty."

Econometric systems of equations are the main tool in economic forecasting. These comprise equations which seek to "model" the behavior of discernible groups of economic agents (consumers, producers, workers, investors, etc.) assuming a considerable degree of rationality – moderated by historical experience. The advantages to economists of using formal econometric systems of national economies are to consolidate existing empirical and theoretical knowledge of how economies function, provide a framework for a progressive research strategy leading to increased understanding over time, help to explain their own failures, as well as provide forecasts and policy advice. Econometric and time-series models are the primary methods of forecasting in economics, but "judgment," "indicators," and even "guesses" may modify the resulting forecasts.

1.1.6. How can one measure the success or failure of forecasts?

A forecast might reasonably be judged "successful" if it was close to the outcome, but that judgment depends on how "close" is measured. Reconsider our example of "guessing" the distance to the moon: it is apparent that accuracy and precision are two dimensions along which forecasts may be judged. To the layman, a very precise forecast that is highly inaccurate might be thought undesirable, as might an accurate but very imprecise forecast: and experts concur – the "gold standard" is an accurate and precise forecast. Failure is easier to discern: a forecast is a failure if it is inaccurate by an amount that is large relative to its claimed precision. Thus, forecasters are squeezed between wanting accurate and precise forecasts, yet not claiming so much precision that they regularly fail.

The notion of "unbiasedness," whereby forecasts are centered on outcomes, is used in technical analyses to measure accuracy; whereas that of small variance, so only a narrow range of outcomes is compatible with the forecast statement, measures precision. In principle, in any specific instance, the costs attached to bias and variance will depend on the purposes to which the forecast is to be put, via the costs of any actions taken on the basis of the forecast. Such costs define the "loss function," though in practice, the loss function may not be explicitly stated. When (squared) bias and variance are combined one-for-one, we obtain the commonly-reported mean square forecast error (MSFE).

Unfortunately, no unique measure of a "winner" is possible in a forecasting competition involving either multi-period or multi-variable forecasts, which are the norm in economics – forecasting say unemployment and inflation up to two years ahead for a "misery index." Surprisingly, that claim remains true even when the metric for "closeness" (such as MSFE) is agreed. Figure 1.1 illustrates the problem. The forecast in the top-left panel (denoted *a*) is awful for the levels of the series shown, but is accurate for the growth rates (top-right panel); conversely, forecast *b* (lower-left panel) is fine for the levels, but dreadful for growth (lower-right panel). Thus, one must decide on which aspect it is important to be close before any choice between methods is possible. Worse still, MSFE itself is not an obvious criterion: a stockbroker probably does not care how good or bad a model is on MSFE if it is the best for making money. Indeed, errors on the sign of a price change may be much more important to her than mistakes with a large magnitude, but the correct sign.

To explore the difficulty of judging the winner of a forecasting competition further, consider two friends passing time while waiting at a bus-stop. Sue challenges Peter to forecast the behavior of a student who is standing inside the bus shelter: every 30 seconds they will both write in their diary a forecast for the next 30 seconds as to whether or not the student will have left. Sue realizes that she

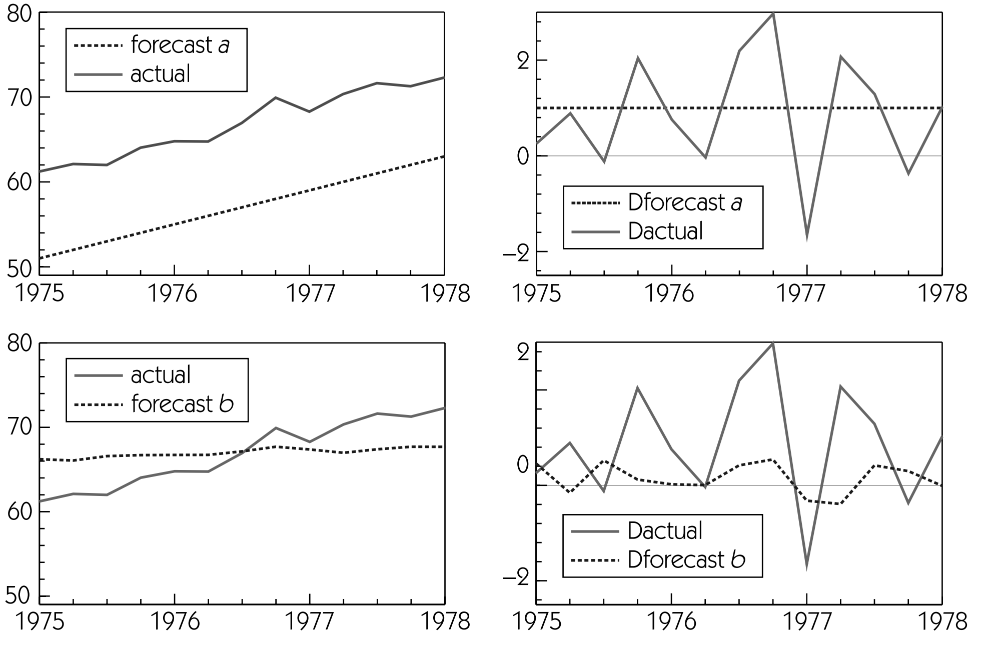

Figure 1.1 Which wins: forecasts of levels or growth?

will only ever make a single mistake if she adopts the following strategy: always forecast that the current state will persist. When the student is there, she forecasts he will still be there in 30 seconds' time; and when he has left, she simply writes that. Thus, in the 5 minutes' wait before the student departs, Sue is correct 10 times, then wrong once, but thereafter correct for ever. Peter, however, is an economist, so he uses a "causal model": students stand at bus stops to get on buses. Thus, if no bus approaches, Peter forecasts the student will stay; but when a bus appears, he forecasts the student will board the bus. Unfortunately, four different buses come by, and the student remains stubbornly at the bus stop – then his girlfriend appears on her motor bike, the student climbs on, and they go away. Peter is wrong four times in the 5 minutes, and if he stuck to his causal model, wrong ever after since the student never got on a bus! Sue wins easily: or does she?

Substitute the phrase "the volcano will not explode" for "will remain at the bus stop," and the vacuous nature of Sue's forecast is clear, even if she did appear to win. On any reasonable loss function, the ability to *anticipate* a volcanic eruption, however imperfectly – which her "no-change" forecast is inherently unable to do – should score highly. The appearance of winning is dependent on agreeing a particular "metric," whereas only awarding a "win" for correctly forecasting the specific event could yield a different winner (of course, that in turn leads to the problem of "forecasting eight of the last three recessions". . .).

In addition to bias and variance considerations, point forecasts are often judged on criteria such as the efficient use of information. Also, forecasts often include forecast intervals, and sometimes of the complete density of outcomes, so these are required to be "well calibrated." Because it is seldom the case that only a single forecast exists of any economic phenomenon of interest, rival forecasts are often available to allow comparisons of one against the other, *ex post*. A natural focus of attention is then whether a combination of one or more forecasts is better than any one forecast alone, or whether one forecast contains all the useful information in another (so that it *encompasses* that forecast). We may wish to test whether, given a particular loss function, one forecast is statistically better than another, and further, whether allowing for the uncertainty inherent in the estimates of a model's parameters affects the inferences made.

1.1.7. How does one analyze the properties of forecasting methods?

The properties of forecasting methods can be investigated in both empirical and artificial settings, using mathematical analysis and computer-intensive numerical methods. For example, we could try the ideas discussed just above for measuring the success or failure of forecasts by testing combinations of forecasting models for encompassing. Forecasting methods can also be compared by Monte Carlo (or stochastic simulation), where an investigator generates artificial "data" on which the models are compared in repeated trials, to calculate how well such methods perform in a controlled environment of their own choosing. However, the empirical relevance of such results depends on whether or not the artificial data "mimic" the relevant properties of the "real world," so may be open to doubt. This method of analysis is most useful when we know the large-sample behavior of statistics of interest for the forecasting methods (say, MSFEs), and wish to investigate the usefulness of these asymptotic results for samples of the size typically available to the applied researcher. This is an example of one of the earliest uses of Monte Carlo in econometrics, namely calculating the small-sample distributions of estimators and tests whose asymptotic behavior is known.

Empirical comparisons, in the form of forecasting competitions, typically look at the performance of different methods for many time series. Because the data-generating process is not under the investigator's control, and will only be imperfectly known, the results of forecast comparisons for any one series could turn on idiosyncratic features of the series, so limit their general applicability. For this reason, many series are compared, and often series are selected which share certain characteristics, with the caveat that the results might only be expected to hold for other series with those characteristics. This highlights a "circularity problem": until we know how empirical economic data are generated, we cannot know the appropriate framework for developing or analyzing methods, so cannot actually know how well they should perform.

1.1.8. What special data features matter most?

Many economic and financial time series possess a number of special features, including, in various combinations: seasonality, business-cycle fluctuations, trend growth, successive dependence, and changing variability. More generally, data in economics is often "nonstationary," namely, has changing means and variances over time. These special data features are potentially important for a number of reasons. Failure to allow for such specific characteristics (say, seasonality) may result in inferior forecasts of aspects of interest (say, turning points, or the underlying trend) especially if, as some recent research suggests, these characteristics are inherently interlinked. More positively, certain of these characteristics may themselves be the focus of attention in the forecasting exercise. For example, one may wish to forecast a business-cycle characteristic, such as the next recession, and be otherwise uninterested in the level, or rate of growth, of the series. For both these reasons, models have been developed which attempt to capture special features, and as will become apparent, many different approaches have been proposed.

Provided that "nonstationary" can be modeled, or incorporated in a systematic way into a forecasting method, having changing means and variances over time is unproblematic. Thus, so-called "unit-root" nonstationary (which leads to trending variances) has been the focus of much research, partly because solutions were clear, although important generalizations were in fact discovered (see, for example, Hendry and Juselius (2000), and Doornik, Hendry, and Nielsen (1998), for recent surveys). However, some other sources of change, particularly in means, have not yielded to modeling as yet; and what "we don't know we don't know" simply cannot be included in any model.

1.1.9. What are the main problems?

One of the main problems with forecasting in economics is that economies evolve over time and are subject to intermittent, and sometimes large, unanticipated shocks. Economic evolution has its sources in scientific discoveries and inventions, which lead to technical progress. This becomes embodied in physical and human capital, and provides the engine for sustained growth in real output. In addition, structural breaks may be precipitated by changes in legislation, sudden switches in economic policy, or political turmoil (examples of breaks relevant to the U.K. include the abolition of exchange controls, the introduction of interest-bearing checking accounts, privatization, the introduction of several radically different monetary-policy regimes, and the steep rises in fuel prices in the 1970s). Thus, erstwhile stable relationships between economic variables are subject to change, and if used to provide forecasts at such times, can result in large and persistent forecast errors.

Moreover, the empirical econometric models used to understand and forecast processes as complicated as national economies are far from perfect representations of behavior. Forecasters may only be dimly aware of what changes are afoot, and even when developments can be envisaged, may find it hard to quantify their likely impacts (for example, the effects of building society demutualizations on consumers' spending in the 1990s). These difficulties entail that economic forecasting is fraught with problems, and in practice, forecast failure – a significant deterioration in forecast performance relative to the anticipated outcome – is common.

Econometric forecasting models are systems of relationships between variables such as GNP, inflation, exchange rates, etc. Their equations are then estimated from available data, mainly aggregate time series. Such models may be viewed as having three main components: deterministic terms introduced to capture averages and steady growth (represented here by intercepts and linear trends, which take the values 1, 1, 1, . . . ; and 1, 2, 3, . . . respectively), and whose future values are known; observed stochastic variables with unknown future values (like consumers' expenditure, prices, etc.); and unobserved errors, all of whose values (past, present, and future) are unknown (though perhaps estimable in the context of a model). The relationships between any of these three components could be inappropriately formulated, inaccurately estimated, or change in unanticipated ways. Each of the resulting nine types of mistake *could* induce poor forecast performance, either from inaccurate (that is, biased), or imprecise (that is, high variance) forecasts. However, it transpires that systematic forecast failure is most likely to depend on the behavior of the deterministic terms, and in particular on unanticipated changes in their values. Such deterministic shifts may reflect changes elsewhere in the economy, interacting with an incomplete or incorrect model specification.

1.1.10. Do these problems have potential solutions?

Forecast intervals seek to measure forecast uncertainty, but can only reflect the "known uncertainties" – deriving from model estimation, assuming future shocks resemble the past – whereas unanticipated deterministic shifts occur intermittently in economics. As we have already stressed, since we don't know what we don't know, it is difficult to account for this "unknown uncertainty."

A simple example may make this clear. Suppose that the variable we wish to forecast is in fact generated by the following equation:

$$Y_t = \delta + \phi Y_{t-1} + a_t, \tag{1.1}$$

where $\{a_t\}$ is a sequence of independent and normally distributed random variables with mean zero, and variance σ_a^2, written $\{a_t\} \sim \mathsf{IN}[0, \sigma_a^2]$. At period T, we know that the value the variable will take next period (in $T + 1$) will be $y_{T+1} = \delta + \phi y_T$ *plus* the realized value of some random shock.[2] Here, we know the distribution of that shock, so we know that Y_{T+1} will be normally distributed around a mean of $\delta + \phi y_T$ with variance σ_a^2. Thus, we can construct intervals for our central projection ($y_{T+1|T} = \delta + \phi y_T$). These will take the form of probability statements that the outcome will fall in certain intervals, say:

$$\Pr(Y_{T+1} \in \{y_{T+1|T} \pm z_{\frac{\alpha}{2}}\sigma_a\}) = 1 - \alpha,$$

where $z_{\frac{\alpha}{2}}$ is the value such that the probability that a standard normal variate is smaller is $\alpha/2$. This interval states that if we were able to witness R realizations of $\{Y_{T+1}\}$, then on $(1 - \alpha) \times R$ of them, the actual outcome would fall within $\pm z_{\frac{\alpha}{2}}\sigma_a$ of the central projection. Since the degree of uncertainty in our point forecast is known *exactl*y, wherein lies the problem?

First, in practice we will never know the values of the parameters of the model, $\{\delta, \phi, \sigma_a^2\}$ so these will have to be replaced by estimates. However, the additional uncertainty this step imparts is a form of "known uncertainty": we know it will arise, and it can be taken into account; see, for example, Clements and Hendry (1998, ch. 4), and the survey of bootstrap techniques in Clements and Taylor (2001). Second, the form of the model cannot be known for certain: there may be further lagged responses than those included in (1.1), or the logs of the variables may be needed rather than the levels, and so on. Uncertainty that may arise from the model specification not being known is less easy to handle: see for example, Chatfield (1993), Chatfield (1995), and Draper (1995). The third major problem, though, is when the future ceases to resemble the past, and this change is unknown to the forecaster. Suppose $\{a_t\} \sim \mathsf{IN}[\mu, \tilde{\sigma}_a^2]$ for $t > T$, say, where either $\mu \neq 0$ and/or $\tilde{\sigma}_a^2 \neq \sigma_a^2$. Notice that $\mu \neq 0$ is equivalent to a shift in δ to $\delta + \mu$. Of course, there is no good reason to suppose that the form of the distribution of the shocks also remains unchanged. The shift in the mean will affect the accuracy of our predictions, and the change in the variance of the disturbances means that we will either over- or under-estimate the uncertainty surrounding the point forecasts. The shift in the error distribution will induce incorrect density forecasts.

Nevertheless, there are ways of avoiding systematic forecast failure in economies that are subject to sudden, unanticipated, large shifts. When shocks are unanticipated, it would take a magician to conjure ways of avoiding large errors in forecasts announced *before* such shocks have occurred. Rather, given an inability to forecast the shock, adaptation is required once a shock has occurred, to avoid a sequence of poor forecasts. Some models adapt more quickly than others. We cannot explain the precise nature of all the relevant models in this introduction, but note, for example, that shifts in the means of stationary combinations of variables appear to be a key cause of mis-prediction in the widely-used "vector equilibrium-correction mechanism" class of model, whereas vector autoregressions in first (and even second differences) are more robust to deterministic shifts, and adapt to the changed environment. Even when a given model type does not itself adjust rapidly, devices exist to speed that process: for example, so-called "intercept corrections" can be used to "set a model back on track" and thereby attenuate systematic failure. Thus, some solutions exist for at least part of the problem.

To illustrate these, suppose that in (1.1), the intercept δ changes to $\delta + \mu$ at $T + 1$, and for simplicity, $\phi = 0$, so the process is

$$Y_t = \delta + \mu 1_{(t>T)} + a_t, \tag{1.2}$$

where $1_{(t>T)}$ takes the value 1 when $t > T$ and zero otherwise. Forecasting period $T + 2$ at time $T + 1$, without realizing the mean had changed, we would on average make an error of μ, whereas if we used a "random walk predictor," namely $\tilde{y}_{T+2|T+1} = y_{T+1}$, the average error is zero! Alternatively, suppose we revise our forecast of $T + 2$ by the amount by which we were wrong in predicting period $T + 1$, then the average error is again zero. But the cost of improved accuracy is reduced precision. It is simple to show that a confidence interval which would contain the actual outcome $(1 - \alpha) \times 100\%$ of the time would be approximately $\{\tilde{y}_{T+2|T+1} \pm z_{\frac{\alpha}{2}} 2\sigma_a\}$, which is twice as large as for a one-step-ahead prediction using the correct model in the absence of the mean shift.

Effort is also being devoted to studying "early warning" signals. Such methods seek to anticipate change by observing it elsewhere in "leading regions" (California, say), or at a higher frequency (such as in weekly data, when the usual frequency is quarterly). However, logical problems begin to arise in a social science. First, say, the IMF could predict the onset of currency crises, then remedial action would prevent the forecasted outcome from materializing (this is rather like a car mechanic replacing brakes suspected of likely failure, so that failure does not eventuate). Second, say *you* could forecast a crisis, and kept it secret – apart from being believed by a small group of financial backers – then you could create a money machine: your actions might alter the outcome from what was originally predicted, but might equally induce a "self-fulfilling prophecy," creating a crisis where none would otherwise occur. Thus, there are limits, albeit wide, to what we may ever hope to forecast.

1.1.11. What is the future of economic forecasting?

To succeed simultaneously in forecasting competitions and in the policy arena, econometric models will have to mimic the adaptability of the best forecasting devices, while retaining their foundations in economic analysis. Despite a relatively poor track record to date, and the many problems that economic forecasts from econometric systems confront, these models offer a vehicle for understanding, and learning from, failures, as well as consolidating our growing knowledge of economic behavior. Moreover, a close interaction between theory and evidence is essential for a successful economic science. Consequently, econometric systems provide the best long-run hope for successful forecasting, especially as suitable methods are developed to improve their robustness to unanticipated breaks.

We have highlighted many difficulties in our introduction: what we have been less able to do is emphasize the tremendous progress achieved in understanding the properties of forecasting models, methods, and measures. Economic forecasting used to be the orphan of the discipline: "those who could, did; those who couldn't, forecast." Its poor track record tainted the messengers and analysts, as well as the models. However, the 1990s saw a resurgence of interest in the theory and practice of economic forecasting that continues as we write. This *Companion* describes the outcomes of much of that intellectual endeavor, and points to a bright future – at least for those who study economic forecasting.

1.2. A REVIEW OF THE *COMPANION* CONTENTS

What can be forecast? And how confident can we be in our forecasts?

The chapters by Neil Ericsson on forecast uncertainty, and the survey on density forecasting in economics and finance by Anthony Tay and Ken Wallis, can be viewed as a response to "how confident can we be in our forecasts?"

Neil Ericsson discusses the sources of both predictable and unpredictable forecast uncertainty in empirical economic modeling. The key features of predictable uncertainty are illustrated by several analytic models, and with empirical models of the U.S. trade account and U.K. inflation and real income.

Anthony Tay and Ken Wallis present a survey of density forecasting in macroeconomics and finance, and cover the production, presentation, and evaluation of density forecasts. A density forecast is an estimate of the probability distribution of the possible future values of the variable.

How is forecasting done generally?

The chapters by Diego Pedregal and Peter Young, and by Tommaso Proietti, exposit general statistical approaches to modeling time series. The techniques, models, and methods discussed are applicable to observations recorded at equally-spaced intervals of time, and have been employed in many disciplines outside economics. Diego Pedregal and Peter Young's chapter is a "multi-disciplinary overview of the current state-of-the-art" of statistical approaches to modeling time series. From the numerous statistical approaches available, they focus on the unobserved components approach, which they favor for the modeling and forecasting of nonstationary data. The practical usefulness is demonstrated with two detailed empirical examples.

The chapter by Tommaso Proietti also considers unobserved components models. He discusses the specification of models for the major components: models for trend, cycle, and seasonal components, as well as recent developments concerning multivariate extensions of the models, and technical aspects concerning the specification of initial conditions for estimation. Illustrations of forecasting Italian GDP and U.S. auto sales series show the efficacy of the approach relative to Box–Jenkins time-series models.

How is forecasting done by economists?

As one might anticipate, a wide variety of models and methods are used. In contrast to the statistical, model-based approaches surveyed by Pedregal and Young, and Proietti, Dilek Önkal-Atay, Mary Thomson, and Andrew Pollock discuss judgmental forecasting. They note that judgment may enter the forecasting exercise at many levels, from the choice of the variables to include, to the adjustment of model predictions for new information. The term judgmental forecasting refers to "the incorporation of forecasters' opinions and experience into the prediction process," and so covers a wide variety of situations. They review the factors that affect the accuracy of judgmental forecasts, studies that have sought to compare the accuracy of judgmental and model-based forecasts, and judgmental adjustments and the combination of the two types of forecast. They discuss why forecast users sometimes prefer judgmental forecasts, irrespective of their accuracy.

Adrian Pagan and John Robertson discuss how forecasting is done in central banks with a view to informing the setting of macroeconomic monetary policy. From their account of the approaches adopted in a number of central banks, it becomes apparent that there is a good deal of diversity. An important dimension is the extent to which formal methods are preferred to expert systems. They conjecture as to why the specific banks they consider align themselves as they do.

Helmut Lütkepohl exposits modeling and forecasting with cointegrated vector autoregressive-moving average (VARMA) processes. VAR models have been popular in empirical macroeconomic research and in forecasting since Sims (1980) and Doan, Litterman, and Sims (1984), and VARMA models can be viewed as the natural multivariate successors to the univariate models of Box and Jenkins (1976). The moving-average component of the multivariate models has often been ignored in empirical applications, but Lütkepohl provides a comprehensive and clear treatment of the more general model, and shows how the results specialize for the more familiar VAR models. It is important to allow for cointegration in models of this sort because of the existence of long-run relationships between some economic variables; see, for example, Engle and Granger (1987) and Johansen (1995).

The chapter by Raj Bhansali on multi-step forecasting warns that the standard approach to prediction is only optimal when the model coincides with the data generating process. In general, more accurate predictions may result by selecting a new model and/or reestimating the model parameters separately for each forecast horizon of interest.

How can one measure the success or failure of forecasts?

Herman Stekler surveys the literature on testing the rationality and efficiency of individuals' forecasts, and reports on applications to inflation and output-growth forecasts in both the U.S. and U.K. He conjectures as to why some forecasts are not rational, and draws out the connections between some tests of rationality, and tests of forecast encompassing and combination. He concludes that we do not know why "irrational" forecasts appear to have been made at some times but not others, but that some systematic errors may result from fundamental changes in market structure or in the economy which could not be foreseen, so the resulting errors could not in some sense be reasonably avoided, although there is little evidence on this.

Hashem Pesaran and Spyros Skouras provide an overview of quantitative and qualitative methods for evaluating forecasts when there exists *a priori* information regarding the use to which the forecasts will be put. They contrast the decision-based approach to the evaluation and comparison of forecasts with purely statistical approaches, and show how such an approach can provide a unifying theme for recent developments in the forecast-evaluation literature – namely the use of generalized cost-of-error functions, probability event, and density forecast evaluation and the evaluation of market-timing skills. The problem of testing the "equivalence" of two forecast distributions in a decision-based context is also addressed briefly.

The motivation for the chapter by Paul Newbold and David Harvey on forecast combination and encompassing is that two or more forecasts of the same phenomenon may sometimes be available. Empirically, it is often the case that a combination of the forecasts is preferable to any individual forecast, as might seem reasonable when the forecasts draw on different sources of information. One forecast is said to encompass another when the optimal combination of the two assigns zero weight to the second forecast.

Roberto Mariano considers ways in which *tests* of models' out-of-sample performance can be constructed. This goes beyond the *comparison* of models' forecasts, by attempting to determine whether differences in performance are significant once the stochastic nature of the problem is taken into account. Early efforts toward the construction of appropriate statistical tests of forecast accuracy made a number of stringent assumptions concerning the properties of the forecast errors and loss functions that belied their usefulness, though recent contributions relax these conditions.

Even so, these tests of a model's out-of-sample predictive accuracy often proceed as if the parameters of the models from which the forecasts come are known. Michael McCracken and Kenneth West show that parameter estimation uncertainty can affect the construction of asymptotically-valid tests of predictive ability. They explain the circumstances in which the impact of parameter estimation uncertainty needs to be allowed for in making inference on predictive ability, and suggest how this can be done.

How do we analyze the properties of forecasting methods?

Robert Fildes and Keith Ord review the role of forecasting competitions in improving forecasting practice and research. They consider what has been learnt from the major competitions over the past 30 years. They first describe the components of a competition and the criteria for making comparisons across competing methods. The principal competitions are then reviewed, along with the criticisms that have been levelled against them. They argue that it is possible to draw valid conclusions about the performance of different methods, based upon the evidence provided by forecasting competitions.

Øyvind Eitrheim, Tore Anders Husebϕ, and Ragnar Nymoen report on empirical forecast accuracy comparisons between models that feature equilibrium-correction mechanisms, and models in differences (or growth rates), in the context of inflation targeting. As they note, one- or two-year-ahead forecasts of inflation are a vital input to monetary policy decision-making in countries which have adopted inflation targeting. They show that rival models of the inflation process have different forecasting properties over a period that covers a change in regime in the Norwegian economy, and that models in differences are in some instances preferable.

The third contribution on comparing forecasting methods is by Gonzalo Camba-Mendez, George Kapetanios, Martin Weale, and Richard Smith, who consider the forecasting performance of leading indicators for predicting industrial production in four major European countries. They present a way of evaluating the forecasting ability of composite leading indicator variables of industrial economic activity, and apply it to the four European countries.

What special data features matter most?

Seasonality is an important feature of many economic and financial time series. Traditionally, seasonality has often been viewed as being essentially fixed over time, and a "nuisance" that masks the underlying movements in the original series. As such, researchers interested in, say, the long-run properties of economic time series, have felt justified in analyzing seasonally-adjusted data, although as Denise Osborn notes, seasonality may be viewed as exhibiting a type of unit-root behavior, which has consequences for the long-run properties of the series. Her chapter looks at the issues concerned with whether seasonal patterns exhibit unit roots or are deterministic. Philip Hans Franses and Richard Paap consider forecasting univariate seasonal time-series data using periodic autoregressive models. Periodic models allow the parameters of the model to vary with the season. They discuss a concept of integration that arises in such models, and the appropriate treatment of deterministic terms. They provide an empirical comparison of the out-of-sample forecast performance of a number of models of seasonality for quarterly U.K. consumption series.

There has been much recent interest in time-series models that are able to capture perceived asymmetries in the phases of business-cycle fluctuations in activity. Ruey Tsay provides a general review of nonlinear time-series models and forecasting. He discusses testing for nonlinearity, modeling procedures, and forecasting for some nonlinear time-series models, and illustrates with a number of empirical applications. This is an area of on-going intensive research activity, and Tsay's chapter offers the reader the opportunity to gain a better understanding of nonlinear models and an appreciation of their applications.

Stefan Lundbergh and Timo Teräsvirta focus on a particular class of nonlinear time-series model – smooth transition autoregressive (STAR) models. They discuss modeling procedures, techniques for obtaining multi-period predictions, and the generation and display of forecast densities. For an empirical example, they estimate and forecast STAR models for two quarterly unemployment series.

Terence Mills reviews a number of the special features possessed by many financial time series, such as that price changes (or returns) may be largely uncorrelated, but that the conditional variances of price changes are likely to be forecastable. His chapter covers the long-memory or long-range dependence thought to be a feature of some financial time series, the "fat-tailedness" of the unconditional distributions of returns, nonlinear models of returns, trading rules, calendar effects, and forecasting returns using other variables, and equilibrium-correction models of returns. He conjectures that future developments in the econometric analysis of financial time series will continue to exploit the explosion in computer power of the last decade, focusing on nonlinear models and computer-intensive evaluation techniques such as the bootstrap. Both linear and nonlinear multivariate cointegration techniques are expected to become more widespread.

What are the main problems? Do these have potential solutions?

Clements and Hendry argue that the main problems that afflict economic forecasting arise from "the things we don't know we don't know," and of these, shifts in deterministic terms might be the most pernicious. Other possible sources of forecast errors – such as misspecifying the stochastic components or uncertainty due to estimating their parameters – are likely to be less important. Potential solutions such as updating the models parameters, differencing to exploit the rapid adaptability of a random walk process, "intercept corrections," and modeling the intercept shifts are briefly reviewed by Clements and Hendry. Øyvind Eitrheim, Tore Anders Husebϕ, and Ragnar Nymoen consider whether "differencing," that is, using forecasts from models in the first differences of the variables, may yield better forecasts than models with fully articulated long-run solutions in the presence of possible nonconstancies.

In general, the relative efficacy of these methods will depend on the nature and frequency of the shifts. When a time series exhibits a sudden change in mean over the sample period, one possible course of action is to include appropriate dummy variables (impulse or shift, depending on whether the change is immediately reversed) to capture the effects of outliers or "one-off" factors, without which the model may not be constant over the past. This strategy is popular in econometric modeling. However, to the extent that these "one-off" factors could not have been foreseen *ex ante* and may recur, the model standard error is an under-estimate of the true uncertainty inherent in explaining the dependent variable, and prediction intervals may be similarly misleading.

The nonlinear time-series models reviewed by Ruey Tsay, and Stefan Lundbergh and Timo Teräsvirta, arguably offer a more accurate picture of the uncertainty surrounding the model predictions, because they explicitly build into the probabilistic structure of the model the possibility that further regime changes may occur. For example, if one regime change of a particular type was observed historically in the last 30 years, then the model could be set up in such a way that a typical sample of 30 years generated by the model would include one such episode of that type. However, unless the shifts are regular and to some extent predictable, the best that one might be able to do is to rapidly adapt to the changed circumstances following the break, to prevent sequences of large errors of the same sign occurring. A similar point relates to the unobserved component, or structural time-series, models reviewed by Diego Pedregal and Peter Young, and by Tommaso Proietti. These models would appear to be most useful when there is *gradual* change in, say, the slope and level of the trend.

General treatments of economic forecasting are provided by, in roughly increasing order of difficulty, Hendry and Ericsson (2001), Diebold (1998), Allen and Fildes (2001), Whitley (1994), Box and Jenkins (1976), Granger and Newbold (1986), Harvey (1989), Clements and Hendry (1998), Clements and Hendry (1999), Hackl and Westlund (1991), and Engle and White (1999).

Acknowledgments

Financial support from the U.K. Economic and Social Research Council under grant L138251009 is gratefully acknowledged by both authors.

Notes

1 This chapter is based on Hendry (2001).
2 Lower case letters denote realizations of the random variables given by the upper case letter. $y_{T+1|T}$ is used to denote the forecast of period $T + 1$ made at time T.

References

Allen, P.G. and R.A. Fildes (2001). Econometric forecasting strategies and techniques. In J.S. Armstrong (ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Kluwer Academic Press, USA.

Box, G.E.P. and G.M. Jenkins (1976). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day. First published, 1970.

Chatfield, C. (1993). Calculating interval forecasts. *Journal of Business and Economic Statistics,* 11, 121–35.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society,* A 158, 419–66. With discussion.

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

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

Clements, M.P. and N. Taylor (2001). Bootstrapping prediction intervals for autoregressive models. *International Journal of Forecasting*. 17, 247–67.

Coyle, D. (2001). Making sense of published economic forecasts. In D.F. Hendry and N.R. Ericsson (eds.), *Understanding Economic Forecasts*, pp. 54–67. Cambridge, Mass.: MIT Press.

Diebold, F.X. (1998). *Elements of Forecasting*. Cincinnati: South-Western College Publishing.

Doan, T., R. Litterman, and C.A. Sims (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3, 1–100.

Doornik, J.A., D.F. Hendry, and B. Nielsen (1998). Inference in cointegrated models: U.K. M1 revisited. *Journal of Economic Surveys*, 12, 533–72.

Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society*, B 57, 45–97. With discussion.

Engle, R.F. and C.W.J. Granger (1987). Cointegration and error correction: Representation, estimation and testing. *Econometrica*, 55, 251–76.

Engle, R.F. and H. White (eds.) (1999). *Cointegration, Causality and Forecasting*. Oxford: Oxford University Press.

Granger, C.W.J. and P. Newbold (1986). *Forecasting Economic Time Series*, 2nd edn. New York: Academic Press.

Hackl, P. and A.H. Westlund (eds.) (1991). *Economic Structural Change, Analysis and Forecasting*. Berlin: Springer-Verlag.

Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Hendry, D.F. (2001). How economists forecast. In D.F. Hendry and N.R. Ericsson (eds.), *Understanding Economic Forecasts*. Cambridge, MA.: MIT Press, 15–41.

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

Hendry, D.F. and K. Juselius (2000). Explaining cointegration analysis: Part I. *Energy Journal*, 21, 1–42.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Sims, C.A. (1980). Macroeconomics and reality. *Econometrica*, 48, 1–48. Reprinted in Granger, C.W.J. (ed.) (1990). *Modeling Economic Series*. Oxford: Clarendon Press.

Whitley, J.D. (1994). *A Course in Macroeconomic Modeling and Forecasting*. London: Harvester Wheatsheaf.