1 Introduction

SEAN HOLLY and MARTIN WEALE

In this book we bring together papers prepared by many of the groups supported by the ESRC's UK Macroeconomic Modelling Consortium, together with a number of other contributions. The papers span the issues of forecasting and macroeconomic policy analysis as well as discussing new developments in modelling itself.

One of the puzzles of macroeconomic forecasting is why 'bad' forecasters, i.e. those with no coherent and properly articulated view of the functioning of the macroeconomy, do so well. Clements and Hendry (chapter 2) provide an explanation of this: the assumption that the future will be like the past may, in normal circumstances, provide a reasonably good means of forecasting. But it is obviously useless as a means of understanding either the effects of policy or of macroeconomic shocks on the economy. Their study is a cautionary tale against judging the quality of an economist from the performance of forecasts and, implicitly, a criticism of the argument that a theory can be assessed purely from its explanatory power.

Serious forecasters, like HM Treasury, nevertheless keep a careful eye on their record. Their finding, like that of other researchers, is that forecasts constructed with proper macroeconomic models do perform better than the naïve models of the sort that Clements and Hendry describe. An interesting twist which emerges from the Treasury paper (chapter 3) is that the forecasts presented in the Spring, which are mostly budget forecasts[1] and therefore the Chancellor of the Exchequer's rather than the Treasury's have, between 1986 and 1995, a record worse than those for which the Civil Servants take responsibility.

The middle section of the book deals with modelling itself. Bhattarai and Whalley set out a general equilibrium model and describe its use to look at tax issues. Like many general equilibrium models of this sort, it is purely static. As dynamic models of this type develop, so the existing gap between these and 'traditional' macroeconomic models is likely to become increasing blurred. Nevertheless, as the Bank of England explains, a number of different models is needed in order to ensure that the right tools are available to address the full range of policy issues.

Garratt *et al.* set out an alternative means of reconciling short-term dynamics with economic theory. They take the view that theory has more to say about long-run relations in the economy than about the short run; they set out a multivariate system in which long-term cointegrating relationships between variables are taken from economic theory (with the theory being tested and not always accepted) while the short term is driven by a vector autoregression. Their approach should be expected to provide a model which has the good forecasting properties that Clements and Hendry find in very simple models, but at the same time has the constraints of economic theory and a role for policy variables which should make it robust to the criticism of simple time-series models.

The received, but poorly defended logic that macroeconomic policy is to do with monetary policy, is reflected in the section on economic policy (chapters 8–12). All of the papers focus on monetary rather than on fiscal issues. All five papers in this section also consider the conduct of monetary policy in the context of econometric models. Moreover, they all evaluate monetary policy with the use of a variety of *rules*, both simple or 'hand-crafted', and/or optimal (Blake, Weale and Young). Finally all but the contribution of Fair incorporate forward-looking expectations as a matter of course.

Note

1 The budget moved to November in 1993 and returned to March in 1998.

2 Economic forecasting in the face of structural breaks

DAVID F. HENDRY and MICHAEL P. CLEMENTS[1]

1. Introduction

The 'conventional' approach to the theory of economic forecasting is based on the well-established result that the conditional expectation given the available information delivers the minimum mean-square forecast error, so forecasts formed in this way should be unbiased, with independent (non-systematic) errors that have the smallest attainable variances: see *inter alia* Clements and Hendry (1994). The practical experience is very different: econometric model-based forecasts are usually modified by their proprietors, and are often systematically wrong with and without adjustments. Conversely, simple, extrapolative predictors often win forecasting competitions (see, for example, Fildes and Makridakis, 1995). This chapter seeks to explain why such outcomes emerge, and what implications should be drawn.

We have established definitions of the fundamental concepts (unpredictability and forecastability), and explored their properties when forecasting in a non-stationary, evolving economy subject to structural breaks, using mis-specified, data-based models. That has enabled us to clarify the role of 'causal' variables, and enunciate a taxonomy of forecast errors to elucidate the determinants of forecast failure: see Clements and Hendry (1998b). The taxonomy reveals that forecast-period shifts in deterministic factors are the dominant source of systematic failure. Model mis-specification, poor data, collinearity, a lack of parsimony, changes in parameters of zero-mean variables, and even inconsistent estimation do not by themselves induce systematic forecast failure. However, these can all interact with structural changes in deterministic factors to exacerbate failure: see Hendry and Doornik (1997) and Clements and Hendry (1998a).

When forecasting in the face of structural breaks, therefore, the key aim is to avoid systematic forecasting errors due to deterministic shifts. Various approaches are possible, including intercept corrections, differencing, co-breaking and regime-switching models. We are currently examining their properties empirically and in simulation studies, emphasising the distinction between equilibrium correction (based on cointegration) and error correction (automatically offsetting past errors).

Of the many ways of making economic forecasts (including guessing; 'informal models'; extrapolation; surveys; leading indicators – see Emerson and Hendry, 1996; time-series models; and econometric systems), we will only consider time-series and econometric models. There is a vast literature on both, and we merely note some key publications on the former by Kalman (1960), Box and Jenkins (1976), and Harvey and Shephard (1992). The autoregressive integrated moving-average model (ARIMA) is a dominant class, based on the Wold (1938) decomposition theorem; the corresponding multivariate time-series form is the vector autoregression (VAR: see for example, Doan, Litterman and Sims, 1984). Economic forecasting based on econometric models of multivariate time-series will be our primary focus, since such systems consolidate 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 relevant to policy analyses.

However, the likely success of any model-based forecasts depends upon there being regularities which are informative about the future, which the proposed method captures without being swamped by non-regularities. The characteristics of the economic system determine the former so only the last two are under the control of the forecaster. The history of economic forecasting in the UK shows some regularities informative about future events, but also major irregularities as well (see for example, Burns, 1986, Wallis, 1989, Pain and Britton, 1992, and Cook, 1995). The dynamic, cointegrated systems with intermittent structural breaks that are formalised below are consistent with such evidence. Capturing regularities without suffering from non-regularities motivates a different interpretation of parsimony and collinearity, as well as a re-examination of the role of causal information when forecasting models are mis-specified. Several results transpire to be misleading once model mis-specification interacts with non-stationary data (denoting thereby the general sense of processes whose first two moments are not constant over time). Conversely, it becomes feasible to account for the empirical success of procedures that difference data, or use intercept corrections (see for example, Theil, 1961, Klein, 1971, and Wallis and Whitley, 1991), although these methods have no rationale when models are correctly specified. Many potential routes to improving model-based forecasts merit investigation: here, methods of robustifying forecasts by intercept corrections and co-breaking are considered, although we also note pooling devices related to forecast encompassing, and non-linear (regime-switching) models.

The plan of the chapter is as follows. Sections 2 and 3 review and contrast some of the results we have obtained for forecasting in cointegrated-stationary processes, and processes subject to breaks, respectively. We then discuss models and generating mechanisms, before developing a taxonomy of forecast errors in section 4 to highlight the impact of shifts in deterministic terms. Section 5 demonstrates why differencing can be an effective strategy for avoiding systematic forecast failure when deterministic shifts occur in cointegrated time series. Section 6 considers the removal of regime-shift non-stationarity by co-breaking, namely the cancellation of breaks across linear combinations of variables, analogous to cointegration removing unit roots. Such an outcome would allow a subset of variables to be forecast as anticipated despite a break occurring. Finally, section 7 concludes.

2. Forecasting in cointegrated-stationary processes

In Clements and Hendry (1998b), we expound a theory of forecasting applicable to economic time series that can be transformed to stationarity by differencing and cointegration. A general theory of macroeconomic forecasting must allow for non-stationary processes which are subject to intermittent structural breaks, using models that do not coincide with the mechanism that generated the data, and that are selected from (possibly inaccurate) data evidence. Nevertheless, many useful insights come from an analysis of the cointegrated-stationary case, even though it will transpire that there are also important differences when there are breaks. In this section, we review some of the main results for the cointegrated-stationary case, before reporting on the structural-break case in section 3.

The 'conventional' approach

We begin at a high level of abstraction – the theory of economic forecasting assuming that the econometric model coincides with the mechanism generating the data in a (difference) stationary world: see, for example, Klein (1971) and Granger and Newbold (1986). The economic system consists of an n-dimensional stochastic process $\mathbf{x}_t$ with density $\mathsf{D}_{\mathbf{x}_t}(\mathbf{x}_t|\mathbf{X}_{t-1}, \boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq^k$, which is a function of past information $\mathbf{X}_{t-1} = (\ldots \mathbf{x}_1 \ldots \mathbf{x}_{t-1})$. A statistical forecast $\tilde{\mathbf{x}}_{T+h}$ for period $T+h$, conditional on information up to period T is given by $\tilde{\mathbf{x}}_{T+h} = \mathbf{f}_T(\mathbf{X}_T)$, where $\mathbf{f}_T(\cdot)$ indicates that a prior estimate of $\boldsymbol{\theta}$ may be needed (simpler still would be to assume $\boldsymbol{\theta}$ is known). Forecasts calculated as the conditional expectation $\hat{\mathbf{x}}_{T+h} = \mathsf{E}[\mathbf{x}_{T+h}|\mathbf{X}_T]$ are unbiased, and no other predictor conditional on only $\mathbf{X}_T$ has a smaller mean-square forecast error (MSFE) matrix:

$$\mathsf{M}\left[\hat{\mathbf{x}}_{T+h} \mid \mathbf{X}_T\right] = \mathsf{E}\left[\left(\mathbf{x}_{T+h} - \hat{\mathbf{x}}_{T+h}\right)\left(\mathbf{x}_{T+h} - \hat{\mathbf{x}}_{T+h}\right)' \mid \mathbf{X}_T\right].$$

However, the relevance of this result is suspect when the model is mis-specified for the mechanism in an unknown way, and requires both selection and estimation from available data generated by a non-stationary economy subject to unanticipated structural breaks. In such a setting, not only is it difficult to correctly model the underlying processes – the costs of failing to do so can be large. Consequently, we consider what can be established, based on research reported in Clements and Hendry (1995b, 1996a, 1996b) and Hendry and Clements (1994a, 1994b).

Conceptual basis

It is helpful to draw a distinction between two basic concepts which are often used interchangeably – (un)predictability and (un)forecastability. Unpredictability refers to the relationship between a random variable and an information set – a variable is unpredictable if the conditional distribution (given that information set) coincides with the unconditional distribution. Predictability is necessary but not sufficient for forecastability: for the latter, we require a systematic relationship, and also need to know the form of the conditional density of the variable, namely, how the information set enters the data generation process (DGP). Thus, although the conditional expectation delivers the minimum mean-square forecast error, its optimality properties are not a useful basis for forecasting when the *form* of the conditional expectation is either unknown or changes over time.

Non-causal variables may be more relevant than (previously) causally-relevant variables if the model is mis-specified for the DGP and the DGP undergoes structural breaks. This result will be reviewed below given its importance in explaining the roles of practical procedures that would otherwise seem unjustifiable. Since a non-constant DGP, and a misspecified model thereof, may occur regularly in economics, forecasting with an empirical model may be fundamentally different from predicting using the DGP.

That we cannot forecast the unpredictable is a truism. But the scope and applicability of this statement is often not appreciated: more aspects of reality may be unpredictable than just the stochastic error on postulated equations, which is all that forecast-error variance formulae sometimes reflect. Rather, the regular occurrence of forecast failure reveals that other unanticipated changes do occur over the forecast horizon, a theme explored in Hendry and Doornik (1997).

Measuring forecast accuracy

Forecast accuracy can only be assessed once a metric is agreed, and the choice of metric may have a greater influence on the success or failure of a forecasting exercise than is often imagined. Although context-specific cost functions defining mappings between forecast errors and the costs of making those errors may exist in some instances, MSFE-based measures have been the dominant criteria for assessing forecast accuracy in macroeconomic forecasting. However, for multi-step forecasts or multivariate models, such measures are not invariant to non-singular, scale-preserving linear transforms, even though linear models are. Further, unpredictability is not invariant under intertemporal transforms, so uniquely acceptable measures of predictability do not exist. Consequently, different rankings across models, methods or horizons can be obtained from various measures by choosing alternative yet isomorphic representations of a given model. Thus, MSFE rankings can be an artefact of the transformation selected. A generalised forecast-error second moment criterion (denoted GFESM) is invariant, but cannot resolve all problems relating to model choice across different forecast horizons (see Clements and Hendry, 1993), particularly for asymmetric loss functions. Although it is desirable that forecasts be unbiased and efficient, in practice, performance relative to rival forecasts determines the worth of any forecasting procedure.

Forecasts and their confidence intervals derived from linear autoregressive models depend crucially on the time-series properties of the variables. In practice, it may be difficult to discriminate between a trend-stationary and difference-stationary DGP, although the implications of the two for how accurately the process can be forecast are quite different in terms of the 'limit to forecastability' (the horizon up to which forecasts are informative).

Forecasting with systems of integrated (I(1)) variables is a non-trivial extension of the univariate analysis of forecasting with an integrated variable because of cointegration, whereby a linear combination of individually integrated variables may be non-integrated (I(0)). We have established representations of integrated-cointegrated systems relevant for forecasting, derived asymptotic forecast-error variances for multi-step forecasts (which transpire to be useful guides to the finite-sample outcomes), and addressed the implications for forecast accuracy of small-sample parameter estimation uncertainty: see Clements and Hendry (1995a). In bivariate systems, imposing too few cointegration vectors seemed to impose greater costs in forecast accuracy than including levels terms which did not cointegrate. Why this might be the case can be seen by comparing forecasts from a correctly-specified model with those from a model specified solely in differences (the DV model), where we abstract from parameter-estimation uncertainty. The correctly-specified model is the limiting case (infinite sample) of imposing too many levels terms and the DV model imposes too few. Figure 2.1 (reproduced from Clements and Hendry, 1993) plots the ratio of the trace MSFE (TMSFE) for the DV to that for the correctly-specified model, against the forecast horizon. The models are compared in terms of their ability to predict the levels of the variables (the solid line in the figure), their first-differences (the line with squares), and a cointegrating combination (the line with circles). It is apparent that the size of the forecast gains to allowing cointegration depend on the transformation selected, and moreover, for levels and differences evaluation, these gains are greater at short, rather than long, horizons. In fact, whether cointegration is imposed or not makes no difference to the rate at which the MSFEs (or forecast confidence intervals) increase in the forecast horizon. For the levels of the variables these are $O(h)$ in the horizon, h, and for the differences and cointegrating combination they are $O(1)$, for both models. The GFESM (not shown) unambiguously indicates gains to imposing cointegration. However, when breaks occur, differencing (in the sense of imposing too few cointegration vectors) may play a 'robustifying role', as discussed below, potentially reversing the above implications, and instead leading to the strategy of imposing too few cointegration vectors being preferred.

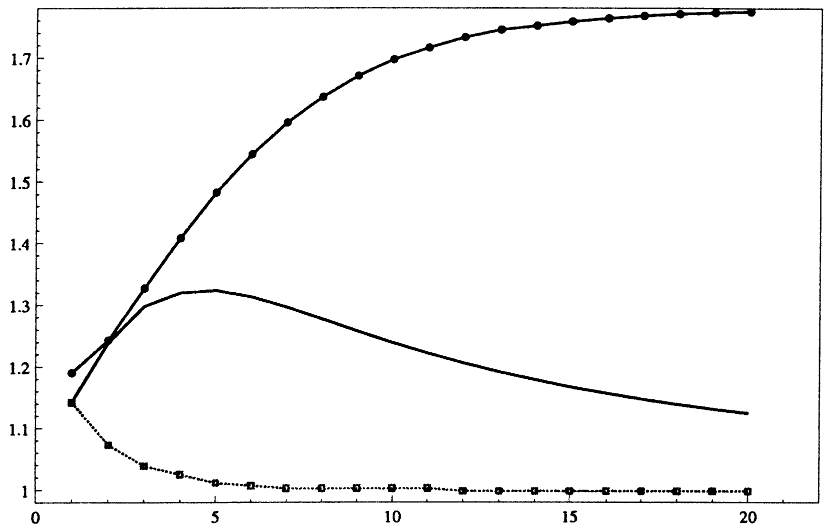

Figure 2.1 *Ratio of* TMSFE *of the DV model to the correctly-specified case. The solid line is for the levels of the* I*(1) data, the line with squares their first differences, and the line with circles an* I*(0) transform (a difference and the cointegrating combination).*

Model adjustments

Forecasts from large-scale macroeconometric models usually reflect adjustments made by their proprietors, and the value-added of such adjustments has long been recognised. We have developed in Hendry and Clements (1994b) a general framework for analysing the adjustments typically made to model-based forecasts, based on the relationships between the DGP, the estimated econometric model, the mechanics of the forecasting technique, the data accuracy, and any information about future events held at the beginning of the forecast period. This suggested various rationales for intercept corrections (non-zero values for a model's error terms over the forecast period), which dove-tail with the sources of forecast error in the taxonomy of Clements and Hendry (1994). Intercept corrections were shown to offer some protection against structural breaks (see Clements and Hendry, 1996b), a point to which we return below. A number of taxonomies of forecast errors can be developed as organising tools for analysing forecasts from large-scale macroeconometric models (see for example, Hendry, 1997, and section 4 below). Generally, we decompose forecast errors into five major categories: parameter non-constancy, model mis-specification, sampling variability, variable mis-measurement, and error uncertainty. Subsequent investigations have suggested that parameter change may have been responsible for many of the more dramatic historical episodes of forecast failure, and that research is the focus of section 3.

Forecast combinations

A combination of forecasts may be superior (on MSFE) to each of the constituents. However, forecast combination runs counter to encompassing: a test for forecast encompassing is the same as that for whether there is any benefit to combination. When models do not draw on a common information pool, so are of an essentially different type, or are differentially susceptible to structural breaks, then a case can be made for combination. Nevertheless, if the aim of an econometric modelling exercise is discovering forecasting models that can be used reliably, then combination is at best a stop-gap measure.

Non-linear models

Two well-known classes of non-linear model which have been used to model (univariate) economic time series are SETAR (self-exciting threshold autoregression: see for example, Tong, 1983) and MS-AR (Markov-switching autoregression: see for example, Hamilton, 1989): Krolzig (1997) provides an overview. The SETAR model is an example of a piece-wise linear model, in that the model is linear within a regime but moves between regimes depending upon the realised value of the process a number of periods previously. The MS-AR model is also linear, conditional upon the regime that the process is in, but now the regime-determining variable is an unobservable variable assumed to follow a Markov chain.

While such models may possess some attractive features for characterising the history of economic processes, their forecast performance is not clearly superior: see, for example, Clements and Smith (1997b, 1999), Clements and Krolzig (1998). The linear autoregressive model is a relatively robust forecasting device against a range of non-linear DGPs. The ability to exploit non-linearities for forecasting may turn on whether forecasts are evaluated conditional upon the regime, reflecting the ability of non-linear models to forecast well in certain states of nature, but not always sufficiently well to score better than linear models on average