Nuffield College
UNIVERSITY OF OXFORD

ECONOMICS DISCUSSION PAPERS

Paper No: 2025-W02

Looking Back to 1991 Economic Forecasting: A Report to the Treasury and Civil Service Select Committee

By David F. Hendry

February 2025

Looking Back to 1991 Economic Forecasting: A Report to the Treasury and Civil Service Select Committee

David F. Hendry*
Nuffield College and Institute for New Economic Thinking
at the Oxford Martin School, University of Oxford, UK

February 4, 2025

Abstract

Originally written in 1991 to advance the formal analysis of forecasting models and methods, the report considers alternative forecasting devices including guessing, extrapolating, leading indicators, surveys, time-series models and econometric systems. Conditional and unconditional forecasts are reviewed and the issue of data accuracy is discussed. The main focus is on macro-econometric model forecasts and their forecast errors, so forecast variances are described. Forecast comparisons using mean square errors across models are criticized, as are methods of pooling disparate forecasts, which violate forecast encompassing. The non-stationarity of economic data is discussed in terms of unit roots and stochastic trends, technical progress and regime shifts. The various sources of forecast error are delineated including: uncertainty about parameters and non-modelled variables; cumulative innovation errors in forecasting endogenous variables; lagged feedbacks onto exogenous variables, parameter non-constancy; incorrect initial values; and model mis-specifications. The report concludes with some recommendations.

JEL classifications: B22, B23.
KEYWORDS: Macroeconomic Modelling; Macro-econometric models; Forecasting, Economic policy.

Contents

*Financial support from the U.K. Economic and Social Research Council under grant R231184 is gratefully acknowledged. Many of the results cited below are the fruits of a research programme reported in Chong and Hendry (1986), Hendry and Richard (1989), Hendry, Neale and Srba (1988), Hendry and Ericsson (1991), Hendry and Mizon (1991), and Lu and Mizon (1991) inter alia. I am indebted to Mike Clements for discussions about his unpublished research on forecasting with potentially cointegrated systems. This report draws heavily on unpublished work in Hendry (1995) and Hendry and Mizon (1993).

Prologue

This paper was originally written in 1991 as a report in my role as Special Advisor to the Treasury and Civil Service Select Committee on Economic Forecasting. The overall reports of such committees are usually published by Her Majesty's Stationery Office (HMSO), and I had thought my report had been included, and indeed had recorded it as such till recently when Neil Ericsson discovered that it had not appeared. At his invitation, I have recreated the report from his scanned copy of the original, as a precursor to my research on forecasting over the next 25 years, and recording much of what was known about forecasting at the time.[1] Sections 9–18 were originally a technical appendix, and have been left after the original main text, but are now numbered sequentially. The conclusion has been moved to the end of the original report. The Epilogue comments on what led to useful developments, what I missed and what I misunderstood at the time. Publications in the body of the text dated later than July 1991 were either working papers or forthcoming at the time of the report, and have been updated; publications in the Epilogue since the report was originally written are under 'Added References'.

1 Introduction and summary

This report first discusses alternative forecasting methods used in economic forecasting. The term forecast is used to denote a statement about a future event or set of events; prediction is used to denote an implication of a model, so forecasts are a sub-class of predictions. Of the possible forecasting methods available, macro-econometric models are the focus of interest here, and section 2 explains why. Next, the two possible forms of forecast, namely conditional on some other events

[1] I am grateful to Jennifer L. Castle and Neil R. Ericsson for many corrections arising from this recreation. Footnotes have been added to clarify some issues in the original, though are somewhat anachronistic.

occurring (e.g., no change in oil prices) or unconditional, are described in Section 3. Most H.M. Treasury forecasts are conditional, so the analysis will be restricted to those. In section 4, the issue of data accuracy is discussed. The basic formulae for forecasting, forecast errors and forecast variances are described in section 5. The properties of these are important in appraising both the limits to forecastability and possible improvements in forecasting. Section 6 discusses mean square forecast errors and forecast comparisons across models. This introduces the topics of pooling of forecasts and the obverse of forecast encompassing, which concerns whether the forecasts of one model can explain the forecast errors made by another. Any need to pool forecasts from distinct sources entails model mis-specification.

Section 7 considers the effects of various forms of non-stationarity in economics data. The purpose of this section is to highlight when forecasts are feasible and contain genuine information, and when in an appropriate transformation, forecasts have no information content. An example based on computer generated data is used to illustrate the effects of certain kinds of non-stationarity. The key concept of cointegration is explained and a range of graphs show the need to forecast non-integrated components to correctly appraise forecasts. Section 8 describes and analyses sources of forecast error and suggests a number of potential improvements to current practices. There are eight major sources of error, none of which can be removed in its entirety, but all of which can be partially mitigated by good practice. The limits to forecastability and the roles of initial conditions and intercept corrections are also discussed. The problem that forecast confidence intervals are non-monotonic in the forecast horizon is noted, and forecast frequency and the use of annual average forecasts are discussed. Some criticisms are made of existing practices.

A formal treatment of forecasting by linear macro-econometric models is presented in what was originally the technical appendix. Key issues discussed include non-stationarity (§9), exogeneity (§10), structural econometric models (§11), cointegration in systems and models (§12), regime shifts and parameter changes (§13), encompassing (§14), and system and conditional dynamic forecasts (§15–§16) including the calculation of forecast confidence intervals. Section 17 provides brief comments on the role of non-linearity and a description of the artificial example used is in section 18. Section 19 concludes the report and notes a number of recommendations.

Issues of model size and aggregation are not investigated here as there are no useful points I can make. My intuition suggests that for a restricted budget allocation, the larger the model beyond the minimum size needed to forecast the main macroeconomic aggregates, the poorer its equations and hence the poorer the forecasts. Historical track records are not discussed either, for the same reason: moreover the Warwick Bureau is in a much better position than myself to comment on such matters. Finally seasonality is not considered: for recent results, see Ericsson et al. (1994). The funding of research into macroeconomic models recently has been severely curtailed by budget restrictions at the Economic and Social Research Council. The latest round of the Macroeconomics Consortium, who supported most of the research underpinning this report, has meant the termination of finance for all researchers in this area other than the actual modelling teams. I believe this to be a disastrous degeneration of an already parlous situation. Had the Select Committee met a year later, I would have had no support to help their enquiry.

2 Alternative forecasting methods

There are many ways of making forecasts ranging from [a] guessing; [b] extrapolating recent tendencies; [c] leading indicators; [d] surveys of plans and anticipations by consumers and business; [e] time-series models; through to [f] formal econometric systems. Irrespective of the method selected, three ingredients are essential to success: that there are regularities to capture; that the selected method captures them; and that such regularities are informative about the future.

[a] **Guessing**

This approach relies on luck, and can be ruled out as a generally useful method, even if at any point in time, some 'oracle' claims to be able to forecast accurately. The flaw of this method is that one cannot predict which oracle will be successful.

[b] **Extrapolation**

This method is fine so long as the perceived tendencies do indeed persist, but the likelihood of that is itself in doubt. Here the telling feature is that different exrapolators are used at different points in time.

[c] **Leading indicators**

Forecasting based on indicators is undergoing a surprising revival of interest: see Stock and Watson (1989). Unfortunately, it remains unreliable unless the reasons for the apparent link of the past to the future are clear, as with, say, orders preceding production. Also, most indicators undergo regular changes in composition. The best - worst? - examples are the Harvard Barometer A-B-C curves and models which try to predict from nonsense relationships (see Coen et al., 1969).

[d] **Surveys**

Sample surveys of plans and anticipations by consumers and business might be informative, but much more research is needed on the extent to which plans are realized, and if not, what explains the departures from the plans. Presently, they seem at best an adjunct to formal methods, not a substitute, and below we comment on linking distinct forecasting methods.

[e] **Time-series models**

These usually are of the form proposed by Box and Jenkins (1976). Historically, they have performed well relative to econometric methods. Much of this apparent success is due to dynamic mis-specification in econometric models, and such a source of error was greatly reduced during the 1980s as modern methods were adopted. The multivariate successor to Box–Jenkins is the Vector Autoregressive Representation (VAR), and in the U.S.A., this approach has claimed success. In the appendix, we demonstrate a role for VARs in model evaluation, but dismiss them as forecasting devices per se. Because economics data are integrated – as discussed in section 7 – VARs must be expressed in changes and hence omit crucial error-correction feedbacks.

[f] **Econometric systems**

Formal econometric systems of national economies seem the only hopeful road. These consolidate existing empirical and theoretical knowledge of how economies function, provide a framework for a progressive research strategy for improving forecasting, and help explain their own failures. Moreover, they are open to adversarial scrutiny and so offer a scientific basis for research. In particular, only [e] and [f] allow well based derivations of measures of forecast uncertainty, and associated tests of forecast adequacy. Nevertheless, to be useful, macro-econometric models must:

(i) be well specified, which requires an evaluation of their congruence, namely, how closely the models embody the available data information;

(ii) dominate alternatives, which is their encompassing ability, namely their ability to account for the results obtained by rival explanations;

(iii) remain constant even when economic policies change, which is their invariance to regime shifts and structural change.

Sections 9–16 consider these aspects for [f], criticize some present practices, point out important developments in forecasting theory, and briefly describe how (i)-(iii) can be evaluated relative to the requirements for success in forecasting.

3 Conditional versus unconditional forecasting

A conditional forecast is a statement about future events which takes other events as given. For example, 'England will win the fourth test if Ambrose is ill'. An unconditional forecast averages across alternative possible future states of nature, taking only the past as given. For example, if conditionally GNP is expected to grow at 2% if no war occurs, and 4% if one does, and the probability of a war is 0.1, then the unconditional forecast is: 0.9x 2% + 0.1x4% = 2.2%. This is a sensible forecast of what would happen on average across many such scenarios; it is a silly forecast of the outcome in any specific instance, since either 2% or 4% will eventuate, not something in between. Thus, an unconditional forecast is akin to the statement that the average family size is 2.4 individuals.

If the purpose of a forecast is as an input to a governmental policy decision (e.g., whether to lower interest rates), then several conditional forecasts are required, namely with and without the policy change implemented. Such exercises are often called scenario studies. At most one of the outcomes will eventuate, and so conditional forecasts are needed. However, for an investment in a new mine, say, an unconditional forecast averaged across a wide range of states of nature may be preferred.

Most Treasury forecasts are conditional. An important issue, therefore, is the choice of the conditioning variables. In particular, invalid conditioning can occur where variables which will be altered by a policy change resulting from a forecast are nevertheless treated as given. This invalidity derives from feedbacks of the forecast variables onto the supposedly given values of the future variables: see section 8 below.

4 Data accuracy

The accuracy of the data influences forecast performance in many ways. Poor samples of data may lead to poor model choices, poor parameter estimates, and increased forecast uncertainty from the errors on the model. Poor baseline data may lead to inaccurate forecasts by indicating a false state of the initial conditions of the economy. Later poor data may lead to weak model evaluation and perhaps invalid inferences about model adequacy.

These difficulties are compounded by data revisions, index rebasings and inconsistencies between series. The first can reverse decisions about model adequacy, especially when they are as large as the current regular revisions to the national accounts. Rebasing price indices in an evolving world can induce failures to reveal important links between time series: specifically, cointegration between series in constant prices may be masked by price revisions being integrated. Data inconsistencies, however, allow the possibility that econometric models could forecast more accurately than the data are measured. Most of the rest of this report treats the data as a given, and of sufficient accuracy not to seriously contaminate forecasts. A sufficient justification is that the data errors are no larger on average than the measured growth rates of the variables: see section 8 below.

5 Forecasts, forecast errors and forecast variances

Among the most basic regularities to be captured in economics are the means and the variances of suitably transformed variables. Such entities are constant and well defined only if the time series of the variables are stationary, which is to say that they 'look alike' in different epochs. Stationarity played a dominant role in econometrics until the late 1970s despite criticisms dating from 1901.[2] To understand that concept, some formalization is essential. The obverse of non-stationarity will later help to clarify the notion of stationarity. Consider an efficient financial market, where if a forecast of a change were feasible, then large profits could be made by suitable speculation. For example, imagine knowing that the £/$ exchange rate, one week ahead, was going to be 10% higher than its present value. Denote a change in a variable x_t at time t by $\Delta x_t = x_t - x_{t-1}$, then we would not expect that change to be forecastable in a financial market (indeed contradictions can ensue otherwise). Technically, we denote an expected value or average by E, so that an inability

[2] Although these references were known at the time, they have been added here in 2018: see e.g., Hooker (1901) and Yule (1926).

to predict a change is expressed as $\mathsf{E}[\Delta x_t] = 0$: in words, the expected change is zero. The actual change is random, being upwards or downwards due to the market participants' views, so that:

$$\Delta x_t = \epsilon_t \tag{1}$$

so $x_t = x_{t-1} + \epsilon_t$ where ϵ_t is a random variable with a zero mean, assumed to be independent in successive periods (or else it could be predicted), and approximately normally distributed. This set of assumptions is denoted by $\mathsf{IN}[0, \sigma_\epsilon^2]$ where σ_ϵ^2 is the variance of ϵ_t. A normally distributed variable has a 68% chance of lying within $\pm\sigma_\epsilon$ of the mean, a 95% chance of lying within $\pm 2\sigma_\epsilon$ of the mean, and is almost certain to lie within $\pm 4\sigma_\epsilon$ of the mean, which is the interval $(-4\sigma_\epsilon, +4\sigma_\epsilon)$. When x_t is the exchange rate, σ_ϵ is about 5% of x on quarterly data, so the previous sentence entails that the exchange rate almost never changes by more than $\pm 20\%$ between quarters.

On several widely accepted criteria, the expected value is the best point forecast of a random variable with a finite mean. In particular, no other linear unbiased predictor exists which is always better. Thus, and again hardly startling, the best forecast of ϵ_{t+1} in (1), given the history of the series, is its mean of zero.

A process like (1) is a difference, or dynamic, equation and is said to have a unit root because the coefficient of x_{t-1} is unity. This makes it non-stationary, in that the mean and variance of x_t are not constant over time. To develop this idea, consider the slightly more general model:

$$x_t = d_0 + d_1 x_{t-1} + \nu_t \tag{2}$$

where $\nu_t \sim \mathsf{IN}[0, \sigma_\nu^2]$. For simplicity, $\{\nu_t\}$ is taken to be an autonomous normal random variable with zero mean and constant variance σ_ν^2, where the successive νs (nus= 'news') are independent. Here, autonomous means that the error is not affected by the economic events being modelled, but does affect them. Such an error is also called an innovation since it is unpredictable given available information (otherwise it would not be news).

The values of d_0 and d_1 determine the properties of x_t and very different sequences of $\{x_t\}$ result as the parameter values change. The stationary case occurs if $-1 < d_1 < 1$. Then x_t is a stationary process, which fluctuates around a mean of $\mathsf{E}[x_t] = d_0/(1 - d_1)$ with a variance of $\mathsf{V}[x_t] = \sigma_\nu^2/(1 - d_1^2)$. The correlation between successive xs is determined by d_1. If the future starts at time $T + 1$, a forecast of x_{T+1} can be calculated by:

$$\widehat{x}_{T+1} = \widehat{d}_0 + \widehat{d}_1 x_T \tag{3}$$

using estimated parameters in practice. This is the conditional mean of next period's x given x_T denoted by $\mathsf{E}[x_{T+1}|x_T] = d_0 + d_1 x_T$. The forecast error for estimated parameter values is:

$$\widehat{\nu}_{T+1} = x_{T+1} - \widehat{x}_{T+1} = \nu_{T+1} + (d_0 - \widehat{d}_0) + (d_1 - \widehat{d}_1)x_T \tag{4}$$

The variance of the forecast error $\mathsf{V}[\widehat{\nu}_{T+1}]$ is:

$$\mathsf{V}[\widehat{\nu}_{T+1}] = \sigma_\nu^2 + \mathsf{V}[\widehat{d}_0] + \mathsf{V}[\widehat{d}_1]x_T^2 + 2\mathsf{C}[\widehat{d}_0, \widehat{d}_1]x_T \tag{5}$$

Four components play a role in (5): the variance σ_ν^2 of the innovation; the variances and covariances of the estimated parameters $\mathsf{V}[\widehat{d}_0]$, $\mathsf{V}[\widehat{d}_1]$ and $\mathsf{C}[\widehat{d}_0, \widehat{d}_1]$; and the initial condition x_T and its square x_T^2. These are all discussed in more detail in section 8. When the parameters are known, $\mathsf{V}[\widehat{\nu}_{T+1}] = \sigma_\nu^2$, which is also the conditional variance $\mathsf{V}[\widehat{x}_{T+1}|x_T]$, is less than the unconditional data variance:

$$\mathsf{V}[x_{T+1}] = \sigma_\nu^2/(1 - d_1^2) \tag{6}$$

Otherwise, for estimated parameters, (5) could be larger than (6): see Chong and Hendry (1986).

To forecast further ahead, say h periods, (2) must be solved backwards from time $T+h$:[3]

$$\widehat{x}_{T+h} = d_0 + d_1 x_{T+h-1} \tag{7}$$

Assuming the expectations of future νs are all zero, the 'best' forecast of x at time $T+h$ in terms of the information available at the start of the forecast period T is given by:

$$\widehat{x}_{T+h} = d_0(1-d_1^h)/(1-d_1) + d_1^h x_T \tag{8}$$

This is the conditional mean of x at time $T+h$ given x at time T. As h increases, d_1^h tends to zero since d_1 is smaller than unity in absolute value. Thus, eventually, the forecast becomes the unconditional mean $d_0/(1-d_1)$. In this sense, stationary processes are well behaved.

Even for known parameter values, the multi-period forecast error variance formula becomes more complicated, but it merits note because of the important role it plays in the feasibility of forecasting:

$$\widehat{\nu}_{T+h} = x_{T+h} - \widehat{x}_{T+h} \tag{9}$$

so that:

$$\mathsf{V}[\widehat{\nu}_{T+h}] = \sigma_\nu^2(1-d_1^{2h})/(1-d_1^2) \tag{10}$$

$\mathsf{V}[\widehat{\nu}_{T+h}]$ is also $\mathsf{V}[x_{T+h}|x_T]$, and from (10), it increases with h, sometimes very rapidly: if d_1 is 0.95, then $\mathsf{V}[\widehat{\nu}_{T+8}]/\mathsf{V}[\widehat{\nu}_T] = 3.5$ rising to over 10-fold as h increases. Alternatively, if d_1 is zero, future values of x cannot be predicted any better than their unconditional mean so that larger values of d_1 ensure more predictability. However, if d_1 is unity, the change in x cannot be predicted beyond its unconditional mean, since (2) then coincides with (1).

These formulae all assume that the forecast uncertainty itself does not alter over time. There are many models of error processes that allow changing variances to occur, of which the most popular is the autoregressive conditional heteroscedastic (ARCH) model due to Engle (1982). However, the empirical evidence in any given case must be the arbiter of the need for such an additional phenomenon, and no issue of principle is involved. The formulae above are basic to forecasting, but in practice must be extended to allow for a variety of complications, as follows: forecasting is usually undertaken from a system; there are non-modelled variables which must be separately forecast; the system is usually non-linear and involves longer lagged reactions than just one period; parameter estimates may be rather uncertain; the initial conditions are subject to error and revision; the model may not characterize the data to within an innovation error; parameter values may change in the forecast period; and operational forecasts may be modified judgementally. Nevertheless, most of the salient issues arise in the simple case just discussed. A discussion of the various sources of forecast errors is provided in section 8, and sections 9–16 offer a formal analysis.

Econometric systems derive their predictive power from the inertial dynamics of the economy being regular. This power is enhanced or attenuated by the information on outside variables and any judgement. Section 7 discusses factors which reduce predictability, of which various non-stationarities are the most important, and section 9 also considers that issue formally.

6 Forecast comparisons

Forecasts from large models reflect the judgments, corrections and inputs of their proprietors as much as the models' properties, so that forecast track records reveal little about model validity. When pure forecast information is available, measures of forecast uncertainty are needed to test model validity, but are rarely provided. Thus, forecast comparisons are often based on mean square forecast errors.

[3] Assuming known parameters here.

[a] Mean square forecast error

The mean square forecast error (MSFE) from a model for a time $t+h$ made at time $t < T$ (to allow its calculation) is $\mathsf{E}[(x_{t+h} - \widehat{x}_{t+h})^2]$, which combines the bias in forecast errors, namely $\mathsf{E}[x_{t+h} - \widehat{x}_{t+h}] = \mathsf{E}[\widehat{\nu}_{T+h}]$ with the forecast error variance, namely $\mathsf{V}[\widehat{\nu}_{T+h}] = \mathsf{E}[(\widehat{\nu}_{T+h} - \mathsf{E}[\widehat{\nu}_{T+h}])^2]$. MSFE is used in both forecast evaluations and for forecast comparisons across models. At first sight, it seems a demanding criterion for a minimum MSFE in a class. Unfortunately, this is not so: it is neither necessary nor sufficient that the model have valid exogeneity, constant parameters, nor provide accurate forecasts, all of which one might deem vital concerns for a model to be used in policy. Encompassing ensures MSFE dominance, but not conversely, and together with tests of parameter constancy and the validity of conditioning help to check for model validity. We now briefly consider forecast encompassing evaluations.

[b] Pooling of forecasts and forecast encompassing

Tests for forecast encompassing concern whether the forecasts of one model can explain the forecast errors made by another. They were proposed by Chong and Hendry (1986) as a feasible way to evaluate large-scale econometric models. Ericsson (1992) and Lu and Mizon (1991) discuss the relative merits of forecast encompassing test statistics and parameter constancy tests. They establish that forecast encompassing entails MSFE dominance so that any need to pool forecasts from disparate sources is indicative of model mis-specification.

Generally if $\widehat{x}_t$ denotes the forecast from model 1, denoted M_1, and $\widetilde{x}_t$ is a forecast from M_2, then the pooled forecast is:

$$\bar{x}_t = (1 - a_1)\widehat{x}_t + a_1\widetilde{x}_t \tag{11}$$

It cannot be shown that this is a sensible way to handle information: two bad models are unlikely to combine on a systematic basis to produce good forecasts. It is much better to sort out the models. Indeed, if the forecast error from M_1 is given by $x_t - \widehat{x}_t = \widehat{u}_t$, and from M_2 by $x_t - \widetilde{x}_t = \widetilde{u}_t$ then:

$$x_t - \bar{x}_t = x_t - \widehat{x}_t + a_1(\widehat{x}_t - \widetilde{x}_t) = \widehat{u}_t + a_1(\widehat{u}_t - \widetilde{u}_t) = \bar{u}_t \tag{12}$$

(say), so that $\mathsf{V}[\bar{u}_t] < \mathsf{V}[\widehat{u}_t]$ only if $a_1 \neq 0$. If so, then the difference in the forecast errors of the two models can help explain the forecast errors of model 1. Thus, M_1 is a poor model.

7 Non-stationarity

There are three distinct and important kinds of non-stationarity in economics: [a] Unit roots and stochastic trends, which potentially can be offset by cointegration; [b] Technical progress and structural change; [c] Regime shifts and structural breaks.

[7.a] Unit roots and stochastic trends

In this case, relative to (2), $d_1 = 1$ and x_t is a random walk, with drift if $d_0 \neq 0$:

$$x_t = d_0 + x_{t-1} + \nu_t \tag{13}$$

which implies that $\Delta x_t = d_0 + \nu_t$. First consider $d_0 = 0$, so that $\Delta x_t = \nu_t$ as in (1). We can express x_t as a function of past νs:

$$x_t = x_0 + \sum_{j=1}^{t} \nu_j \tag{14}$$

Thus, past shocks do not die out and appear to persist for ever. The variance of x_t increases linearly over time and x_t 'wanders' everywhere eventually. Over millennia, this seems an unlikely description of economic variables, many of which have bounds, and most of which do not need to have constant variance inputs, but locally it is a fair first approximation to many (e.g., asset prices). If $d_0 \neq 0$ then x_t also has a constant drift, since in place of (14) we find:

$$x_t = x_0 + d_0 t + \sum_{j=1}^{t} \nu_j \tag{15}$$

This class is often called difference stationary since its first difference is a stationary process. Now, $\mathsf{E}[x_t] = x_0 + d_0 t$, and so trends over time. This contrasts with the conditional mean of x_{t+1} given x_t, which is always $d_0 + x_t$.

The conditional variance for known parameters is $\mathsf{V}[x_{t+1}|x_t] = \sigma_\nu^2$. However, the unconditional variance is $\mathsf{V}[x_t] = \sigma_\nu^2 t$, which trends making distant forecasts of the level completely unreliable. The correlation between successive x_ts is almost unity. The h step ahead forecast from time T is:

$$\begin{aligned} x_{T+h} &= d_0 + x_{T+h-1} + \nu_{T+h} = d_0 + (d_0 + x_{T+h-2} + \nu_{T+h-1}) + \nu_{T+h} \\ &= 2d_0 + x_{T+h-2} + \nu_{T+h-1} + \nu_{T+h} = \cdots = d_0 h + x_T + \sum_{i=0}^{h-1} \nu_{T+h-i} \end{aligned} \tag{16}$$

Thus, the deviation of the variable from the initial condition has a local trend, and a cumulative error. Conversely, from (13), one-step ahead changes cannot be forecast more accurately than their unconditional mean d_0.

It is important to note that $\sigma_\nu^2/\mathsf{V}[x_t]$ will tend to zero over time, so that compared to the level of the economic variable, distant forecasts will appear to be useful, even when there is no forecast accuracy for changes.

The model in (13) is non-stationary, and is the focus of much current research in econometrics: see inter alia, Dickey and Fuller (1979)(1981), Phillips and Durlauf (1986), Park and Phillips (1988)(1989) and Sims et al. (1990). The behaviour of estimation methods and statistical tests is greatly affected by whether variables are stationary or not, more specifically, whether or not unit roots occur, and if so, how many are present. Since x_t in (14) cumulates, or integrates, past shocks, but does so only once, it is usually denoted $\mathsf{I}(1)$, called integrated of order one. If x_t is $\mathsf{I}(1)$, then Δx_t is $\mathsf{I}(0)$, so first differences remove the unit root or integrator. Thus, x_t in (2) is $\mathsf{I}(0)$. Economic processes could potentially be $\mathsf{I}(d)$, and need Δ^d to remove all of their unit roots. Equally, their degree of integration is not an inherent characteristic and could change over time (e.g., prices in hyper-inflations). However, many economic time series seem to be $\mathsf{I}(1)$.

Cointegration

An important possibility is that linear combinations of $\mathsf{I}(1)$ variables could be $\mathsf{I}(0)$ because of 'cointegration': see Engle and Granger (1987). A simple example is if consumption C_t and income Y_t are $\mathsf{I}(1)$ but saving $S_t = Y_t - C_t$ is $\mathsf{I}(0)$, then consumption and income are cointegrated. This example is not empirically reasonable, merely illustrative. Cointegrating relationships are those that hold in the long run: despite levels of variables changing without bound, combinations of variables move together.

To illustrate the main ideas and concepts, a set of (quarterly) data was generated on a computer to mimic the time-series properties of macroeconomics data, artificially dated as 1950(1)–1991(1). The mechanism by which the data were created is, therefore, known to the author. The data are named 'consumption', 'income' and 'saving' to provide a concrete analog, but have no direct relationship to the series of these names in the U.K. Section 18 records the computer-based model and some of its properties. The main points of interest for this section are that the changes in the 'income' series satisfy a model like (2) with $d_1 = 0.5$; that 'consumption' is determined by 'income' and lagged 'saving' such that 'consumption' and 'income' are cointegrated; and that 'saving' is stationary.

Figure 1 shows the time series for the level of 'income', together with the fitted values from the estimated econometric system. The fit appears to be spectacularly good, with both the trend and the minor cycle being accurately tracked.

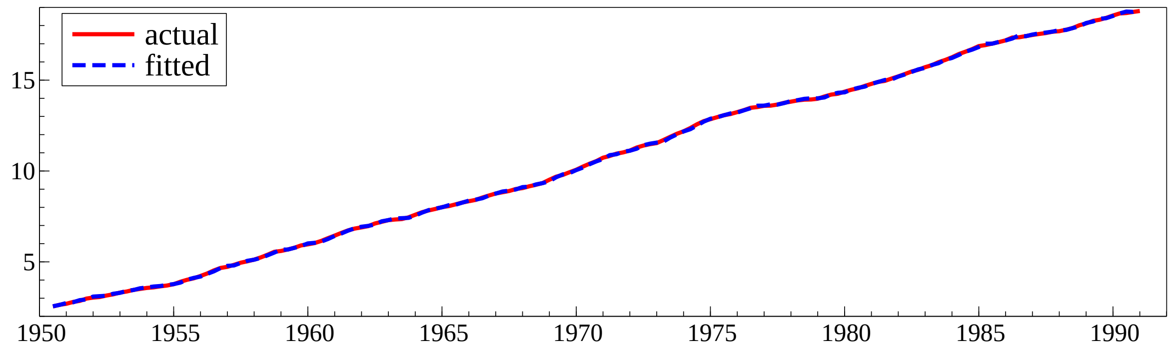

Figure 1: Fitted and actual values for 'income'.

Figure 2 transposes the same information to a graph of the change in 'income', denoted Δ'income'$_t$: clear departures between the fitted and actual values are now apparent and although the track remains good, the correlation between fitted values and outcomes is only 0.5, consistent with the variance of the error.

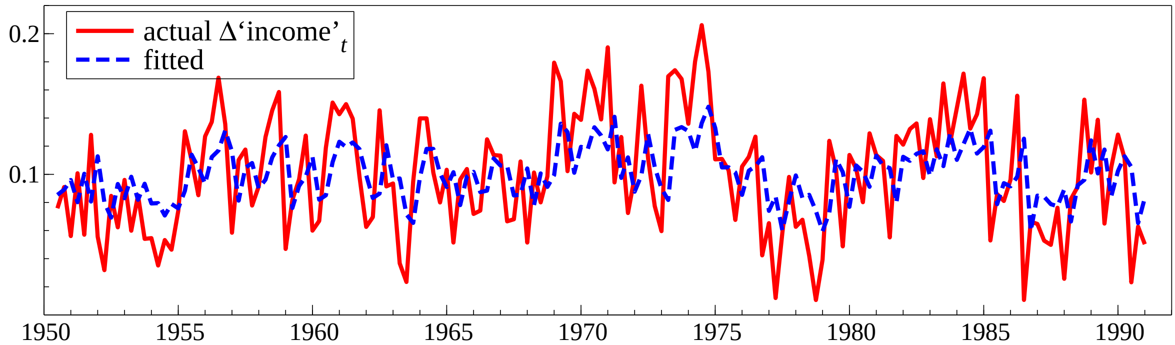

Figure 2: Fitted and actual values for Δ'income'$_t$='income'$_t-$'income'$_{t-1}$.

Figure 3 records a similar graph for 'saving': again the tracking is excellent, and the 'look alike' feature at different points in the sample is reasonably clear.

Given the data on the levels of these three time series, empirical modelling determines one cointegrating link, approximately that 'saving' = 'income'-'consumption'; one non-stationary combination, which represents the stochastic trend; and one identity. The first two are shown in Figure 4, matched by their mean values for clarity. The large range of the latter from -3 to $+3$ swamps the former, which has a range of -0.2 to $+.25$, as seen in Figure 3.

One-step ahead forecasts from the model, based on a succession of steps like those in (3), are shown for 'saving' in Figure 5 over the final eight observations. The 95% interval forecasts are shown by the vertical error bars, centered on the forecasts. These measures correctly reflect the uncertainty due to the equation's innovation error and to the fact that the parameter values have been estimated. All of the realized outcomes lie within their associated confidence interval, and there is little change in the size of the vertical bars as the forecast period advances.

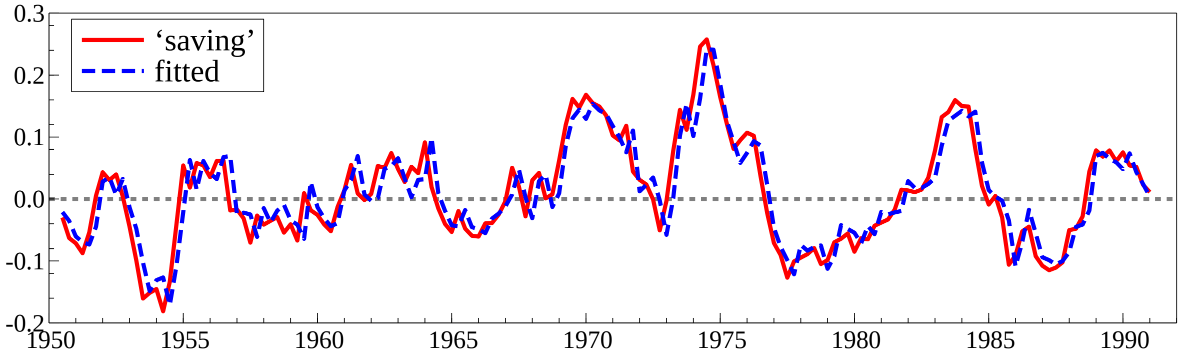

Figure 3: Fitted and actual values for ‘saving’.

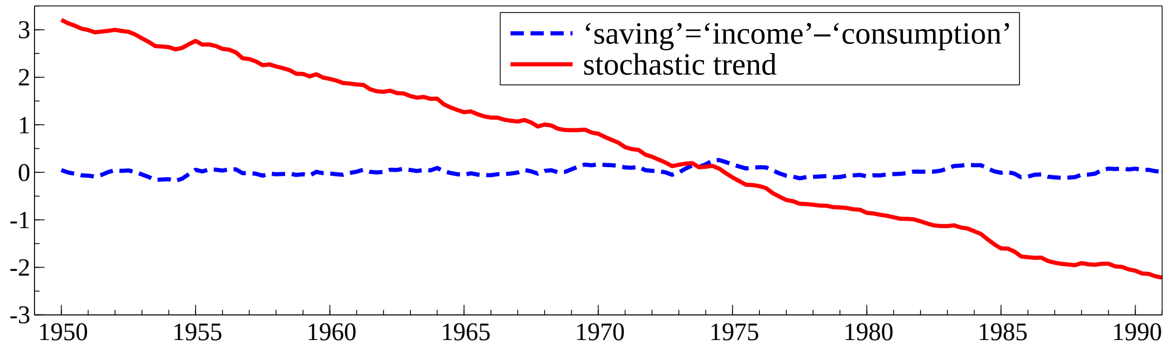

Figure 4: Time series of the cointegrating linear combination (approximately ‘saving’) and the stochastic trend (non-cointegrating combination).

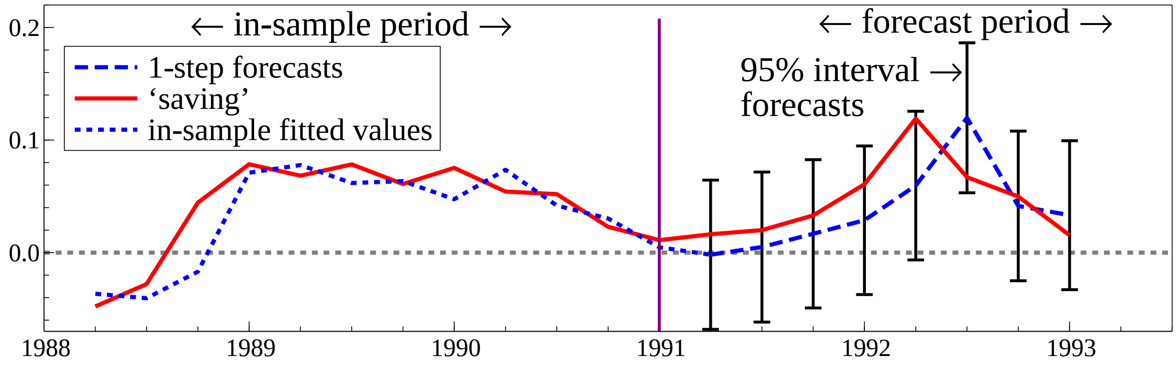

Figure 5: Eight 1-step ahead forecasts of ‘saving’ over 1991(2)–1993(1) with the artificially generated outcomes.

In many respects, this is an ideal illustration since the model passes all the tests of the adequacy of its specification, explains most of the variation in the data, and does not suffer from any of the practical difficulties of measurement errors, parameter change or unmodelled variables. Thus, Figure 6 might be an unpleasant surprise.

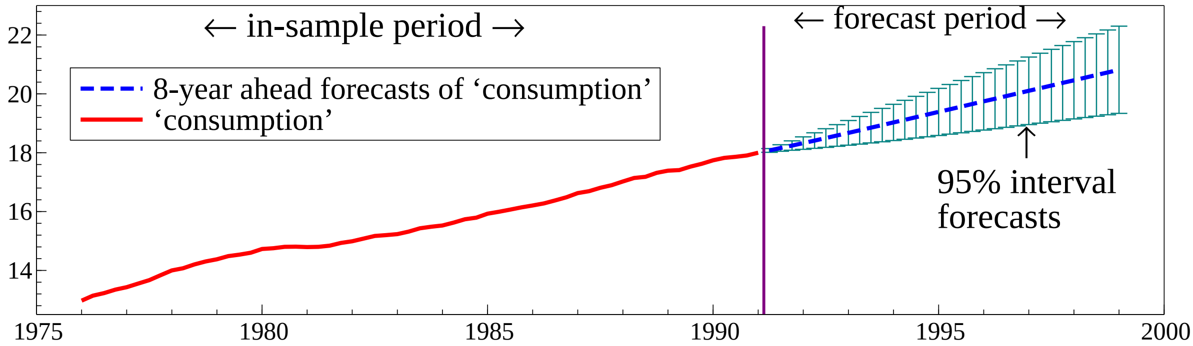

Figure 6: Eight-year ahead forecasts of 'consumption' over 1991(1)–1999(1) with 95% interval forecasts.

Figure 6 reports the multi-step, or dynamic, forecasts for 'consumption' analogous to (8) above, together with the 95% interval forecast bars, which reflect mainly the cumulative uncertainty due to the innovation error. The huge increase in the interval forecasts is obvious: they trend upwards, and at 32 periods ahead – corresponding to 8 years of quarterly data – span a range almost as large as that of the previous 60 data observations. That range is about 3.2 units, whereas 'saving' never varies outside ± 0.3. The mean forecast quickly becomes a trend since the data are non-stationary, and the forecasts are almost uninformative after 10 periods due to the large variances.

Either a large recession or a major boom would be compatible with the interval forecasts calculated. Such a statement holds despite these intervals seriously underestimating the uncertainty likely to be present in any realistic setting, as the model coincides with the mechanism which generated the data. Figure 7 reports eight-year ahead forecasts for a recession scenario of 'consumption' induced by a fall increasing to 5% at the end of the period. This leads to a marked fall in final period 'consumption' relative to the central forecast, but nevertheless lies entirely within the 95% interval forecast bars. The social welfare of the two outcomes is very different (cumulating to a loss of almost half-a-year of 'consumption'), although statistically the two paths are not significantly distinct.

However, the usual reporting mode for a central forecast is shown in Figure 8, namely the central trend without any indices of uncertainty. It may look impressive, but Figures 6 and 7 reveal how uncertain the real environment is. A reasonable idea of the uncertainties inherent in predictions could help produce better policy, and more informed discussion thereof.

Figure 9 records the forecast information in terms of the change in 'consumption'. The forecast variances rapidly converge to a constant, spanning about 0.2, which matches the range of the observed changes in consumption over the previous sample. The forecasts reveal a return of the growth rate to its unconditional mean after a few periods, where it then settles. This is in line with the calculations of forecastability presented in section 18.

Finally, Figure 10 shows the forecast behaviour for 'saving', which is the difference between the two non-stationary series, rather than the change in either. Nevertheless, the outcome is similar to that in Figure 9: the forecast variances rapidly stabilize; there is some information up to about 6 periods ahead; but thereafter, conditional forecasts are no better – at best – than the unconditional, or long-run, mean and variance.

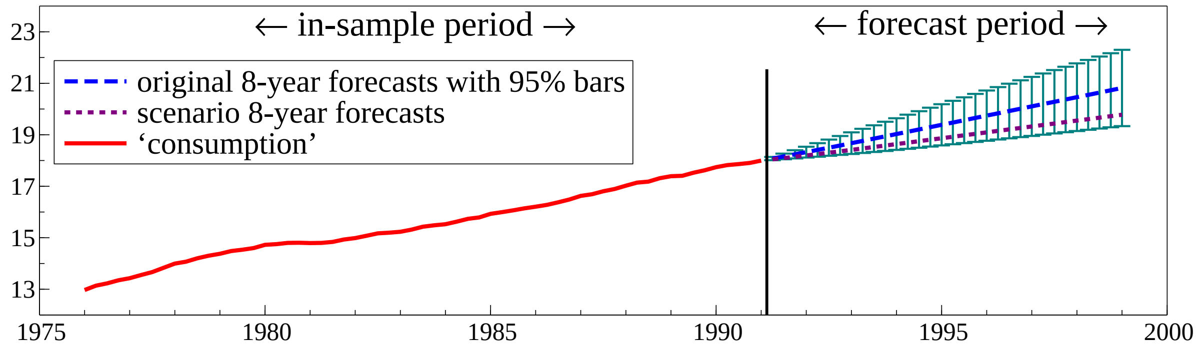

Figure 7: Eight-year ahead forecasts of 'consumption' with 95% interval forecasts and an alternative future trajectory over 1991(1)–1999(1).

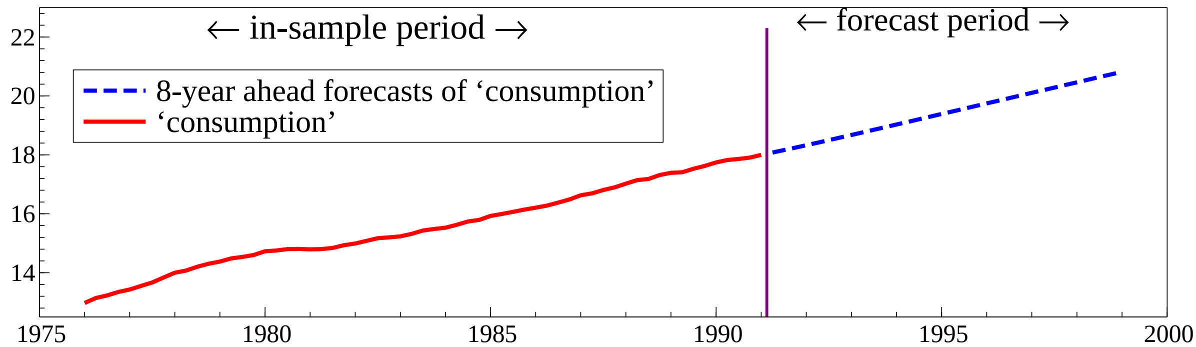

Figure 8: Eight-year ahead forecasts of 'consumption' over 1991(1)–1999(1) with no interval forecasts.

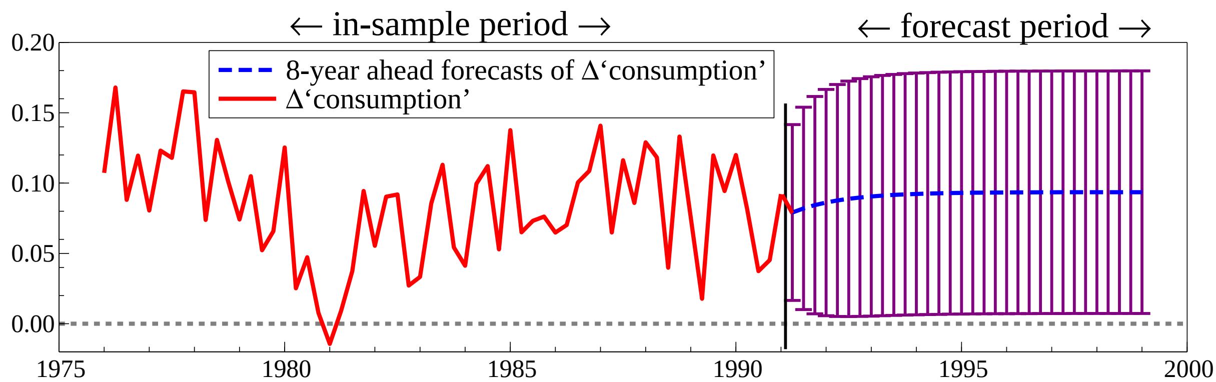

Figure 9: Eight-year ahead forecasts of Δ'consumption' over 1991(1)–1999(1) with 95% interval forecasts.

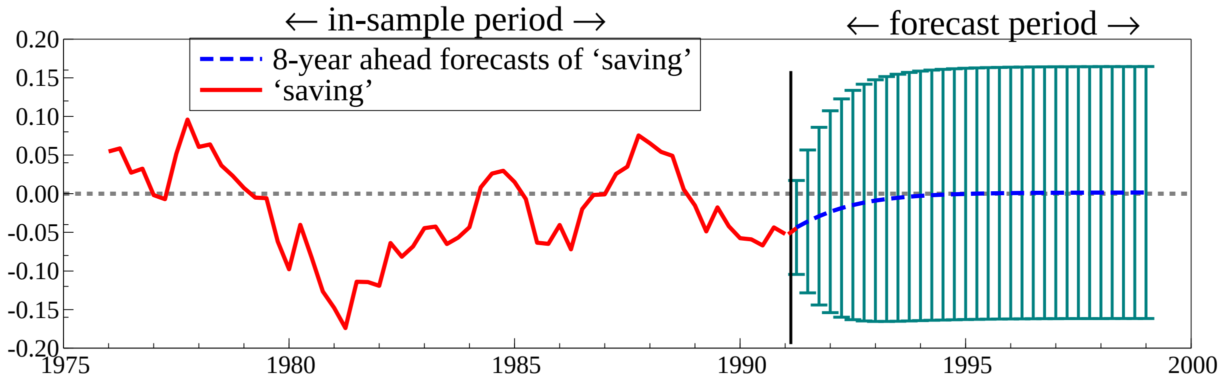

Figure 10: Eight-year ahead forecasts of 'saving' over 1991(1)–1999(1) with 95% interval forecasts.

We return to the system case in section 18, but otherwise it is too technical to analyze further here. However, that section does record the formulae needed to calculate the statistics reported in the graphs (the data were originally generated by PC-NAIVE–see Hendry et al., 1991–and were analyzed and graphed by PC-GIVE–see Hendry, 1989: here OxMetrics was used, Doornik, 2018, with PcGive generating and analyzing the data, Doornik and Hendry, 2018).

[7.b] Technical progress and structural change

The model in equation (13) also describes the phenomenon of economic growth, and is one method of modelling technical progress, when it is random but cumulative. What is not captured is the learning and adjustment processes associated with inventions and the implementation of innovations. The main engines of economic growth are the cumulation of knowledge, and its embodiment in human and physical capital. These are inherently non-stationary processes like those in [7.a]. Econometric systems must also model such features to the extent that they occurred in the sample data, and be modified post-sample for any anticipated changes.

For example, the introduction of interest-bearing sight deposits was a major financial innovation which radically altered the opportunity cost of keeping idle balances in chequing accounts. Appropriate adjustments, taking account of learning, could have been made ex ante to achieve vastly superior forecasts of £M1: see Hendry and Ericsson (1991) for discussion.

[7.c] Regime shifts and structural breaks

Phenomena such as changes in monetary, fiscal or exchange rate policy, oil crises, major innovations like personal computers, the creation of new types of financial assets etc. are better viewed as structural breaks. The occurrence and timing of these is hard to predict but their effects can be modelled, and sometimes anticipated. Sufficiently major changes in one sector which induce predictive failure in a different sector of a model reveal it to have been mis-specified. The so-called 'Lucas critique' (see Lucas, 1976, originally due to Frisch, 1938) is discussed in section 14.

8 Sources of forecast error

There are eight sources of forecast error discussed in detail in section 16. To minimize the technicality of the discussion, some of the verbal discussion is left without the mathematical caveats that would be present in an academic article, but I do not think that any very misleading implications arise thereby.

a] The innovation error;
b] cumulative mistakes in forecasting endogenous variables;
c] parameter uncertainty;
d] non-modelled variables;
e] endogenous feedbacks onto variables treated as given;
f] parameters which are assumed constant but alter;
g] incorrectly measured initial values within sample;
h] incorrect specification of the model for the economic mechanism.
We discuss these in turn.

[a] *The innovation error*

The innovation error in a model is not that of the economy, but is created by the design of the model. Thus, it can be reduced, but only so far as the information set allows, and in the limit, to that inherent in the world economy. In practice, many models have errors which are not even innovations against their own information, and hence they can be improved by better modelling (e.g., by not imposing 'theoretically plausible' but data-rejected constraints); better methodology (e.g., general-to-simple) and larger data bank sources of information. Many useful economic time series for econometric modelling were thoughtlessly cut out by the Rayner review, and many other potentially useful data series are not collected. We greatly underspend on data collection relative to the likely benefits of better information.

[b] *Cumulative mistakes in forecasting endogenous variables*

The cumulative errors due to successively forecasting endogenous variables cannot be overcome in principle; it is unreasonable to imagine that the future evolution of non-stationary dynamic mechanisms can be predicted beyond a few periods ahead. The best that can be achieved here is the accuracy attainable from knowledge of the mechanism itself, which is in effect what section 16 below shows

For $\mathsf{I}(0)$ transforms, such an error source is bounded by the unconditional variance. However, for the level of the economy (i.e., its scale), the variance is unbounded. In words, anything is possible in the indefinite future: imagine the level and composition of GNP today had the Industrial Revolution not occurred. Fortunately, only a few periods ahead are usually necessary for economic policy of a stabilization form (e.g., 4–6 quarters), and only medium-term averages are needed for growth and investment strategies such as education and infrastructure. Again, this is illustrated in section 16.

[c] *Parameter uncertainty*

Reducing parameter uncertainty requires better measured and longer samples of data; better methods of estimation; better economic insights; better specifications; and more parsimonious explanations. However, better data alone are not a panacea. When economics variables are non-stationary, their variances grow over time as shown in section 7. It is also unreasonable to model measurement errors as growing relative to measured growth rates: if we measure GNP as growing at 4%, it may be growing at 2% or even 6%, but it is not likely to be growing at 4000%. Thus, in a cointegrated system, measurement error variances will be bounded relative to long-run relationships, so they mainly affect the short-run dynamics. This still matters for forecasting, and hence any attempted stabilization policy, but is less relevant for medium-term planning.

When parameter values are uncertain, in $\mathsf{I}(0)$ processes, correct forecast confidence intervals may be non-monotonic in the horizon: as the horizon of the forecast increases, the confidence band first expands then contracts. Since it must end up at the unconditional variance, conditional dynamic forecasts can have larger variances than unconditional. This allows some scope for pooling or averaging of forecasts, especially between unconditional and conditional predictors.

[d] *Non-modelled variables*

Non-modelled variable uncertainty can only be reduced by devoting greater resources to modelling the 'exogenous' and policy variables; having a better feel for what the future will bring forth; and taking account of extraneous information on their likely evolution. Survey information has a potentially useful role in improving 'off-line' forecasting. Of course, the situation is very different for the Treasury than for outside agencies as regards information about the future path of economic policy variables, so the Treasury should be able to out perform most outside forecasters, *ceteris paribus*. We now turn to what section 16 calls the hidden sources of error.

[e] *Endogenous feedbacks onto variables treated as given*

Lagged feedbacks of the endogenous variables on to the non-modelled variables, namely those which are treated as given, is called Granger causality in economics (see Granger, 1969). Such feedbacks violate conditioning on future non-modelled variables, but are testable from sample information. Very misleading forecasts can result in the worst case: consider forecasting nominal wages assuming a fixed future trajectory for prices.

[f] *Parameters which are assumed constant but alter*

Parameters are almost automatically assumed constant in forecasting but may alter. Two possibilities are that agents alter the way they form expectations (the Lucas critique discussed in section 13, and noted above); and that a regime shift occurs out of sample. Examples of how financial innovation in the U.K. and the U.S.A. can be handled to allow reasonable parameter constancy are provided in Hendry and Ericsson (1991) (also see Baba et al., 1992).

[g] *Incorrectly measured initial values*

The initial values for the forecast are usually incorrectly measured and this is undoubtedly an important source of error in short-term forecasts.[4] Several possibilities suggest themselves for improving current practice. First, the econometric model itself could be used to predict the initial conditions for the current forecast, given the better established initial values from a few years earlier. Adjustments can be made where discrepancies arise suggestive of incorrect initial values. Second, 'intercept corrections' to reflect extraneous information can be used. This is about the only justifiable use of such *ad hoc* factors. As shown in Chong and Hendry (1986) the impact of larger measurement errors at the end of the sample can be derived analytically, and entail appropriate corrections. Thirdly, survey information can help confirm or amend other measurements on the state of the economy.

[h] *Incorrect specification of the model for the economic mechanism*

When, as always, the econometric model is not a correct specification of the economic mechanism, a variety of mistakes can ensue. Much effort in technical econometric theory is devoted to this issue. Conversely, to obtain useful forecasts, it is not necessary for the model to be a facsimile of the economy, merely to model it up to an innovation error on the available information. Section 16 considers some of the specification issues; Hendry (1995) provides an extensive treatment.

9 Systems for non-stationary data

We first need a notation for writing down large systems. Let $\mathbf{x}_t$ be an $N \times 1$ vector of time-series variables such that $\mathbf{x}_t = (x_{1t}, x_{2t}, ..., x_{Nt})'$. For exposition, the data generation process (DGP) is the dynamic linear system:

$$\mathbf{A}(L)\mathbf{x}_t = \mathbf{\Psi}\mathbf{D}_t + \nu_t \;\text{ where }\; \nu_t \sim \mathsf{IN}_n\left[\mathbf{0}, \mathbf{\Sigma}\right], \tag{17}$$

for $t = 1, \ldots, T$ where $\mathbf{D}_t$ is $N \times k$ and contains k deterministic components (constant, trend, any shift indicators and centered seasonal dummy variables), and ν_t is an independent normal error with expectation $\mathsf{E}[\nu_t] = \mathbf{0}$ and variance matrix $\mathsf{V}[\nu_t] = \mathbf{\Sigma}$. In (17):

$$\mathbf{A}(L) = \sum_{j=0}^{p} \mathbf{A}_j L^j = \mathbf{I}_N + \mathbf{A}^*(L) \tag{18}$$

[4]Depending on the variable, as some could be accurately measured, such as domestic interest rates.

which is a pth order matrix polynomial in the lag operator L with $\mathbf{A}_0 = \mathbf{I}_N$. Also, $\mathbf{\Sigma}$ is an unrestricted $N \times N$ covariance matrix. The initial values $(\mathbf{x}_{1-p}, \mathbf{x}_{2-p}, ..., \mathbf{x}_0)$ are fixed, and p is finite, so that moving average components are excluded. These assumptions, together with those about independence, normality and homoscedasticity, are made to simplify the analysis, and none of them is a fundamental restriction. Their relaxation would complicate the algebra without yielding valuable additional insights. However, the assumption that $\{\mathbf{A}_1, ..., \mathbf{A}_p, \psi, \mathbf{\Sigma}\}$ are constant is fundamental. Equation (17) can be reparametrized as:

$$\Delta \mathbf{x}_t = \sum_{i=1}^{p} \mathbf{\Pi}_i \Delta \mathbf{x}_{t-i} + \mathbf{\Pi} \mathbf{x}_{t-p} + \mathbf{\Psi} \mathbf{D}_t + \nu_t \tag{19}$$

where $\mathbf{\Pi}_i = -(\mathbf{I}_N + \sum_{j=1}^{i} \mathbf{A}_j)$ and $\mathbf{\Pi} = -(\mathbf{I}_N + \sum_{j=1}^{p} \mathbf{A}_j) = -\mathbf{A}(1)$ and so $\mathbf{\Pi}$ is the matrix of long-run responses: see e.g. Johansen and Juselius (1990). Although $\nu_t \sim \mathsf{IN}_n[\mathbf{0}, \mathbf{\Sigma}]$, and hence is stationary, the N variables in $\mathbf{x}_t$ need not all be stationary. For example, if $p = 1$ and $\mathbf{\Pi} = \mathbf{0}$, then (19) consists of N random walks, possibly with drifts and quadratic trends. Ignoring deterministic non-stationarities, the rank of $\mathbf{\Pi}$ determines how many levels variables are stationary, and how many are integrated of higher order. If the long-run matrix $\mathbf{\Pi}$ is full rank then $\mathbf{x}_t$ is stationary.

It is assumed throughout that none of the roots of $\det(\mathbf{A}(L)) = 0$ are inside the unit circle, so that explosive variables are excluded. Letting r denote the rank of $\mathbf{\Pi}$, there are three cases: (i) $r = N$ so that all N variables in $\mathbf{x}_t$ are I(0) and hence stationary if there are no deterministic regime shifts or trends in $\mathbf{D}_t$; (ii) $r = 0$ so that all N variables in $\mathbf{x}_t$ are I(1), and $\Delta \mathbf{x}_t$ is stationary (apart from the possibility of a deterministic trend); and (iii) $0 < r < N$ in which case there are $(N - r)$ linear combinations of $\mathbf{x}_t$ which act as common stochastic trends, and r cointegrated linear combinations of $\mathbf{x}_t$ that lead to error-correction mechanisms (ECMs): see e.g., Hylleberg and Mizon (1989).[5]

It is crucial that the nature of the deterministic variables in $\mathbf{D}_t$ be established if valid inferences are to be drawn, particularly when $\mathbf{x}_t$ is I(1). The systems-based Maximum Likelihood Estimator (MLE) of Johansen (1988) determines the number of cointegrating vectors. His approach allows tests of linear restrictions on cointegrating vectors so that alternative theories of long-run outcomes can be tested directly without conditioning too strongly on any particular model of short-run dynamics. Cointegration occurs if linear combinations of the $\mathbf{x}_t$ are I(0), denoted by $\beta' \mathbf{x}_t \sim$ I(0): see Granger (1986) and Engle and Granger (1987). The number of cointegrating vectors equals the rank of β, denoted r, but r will almost always be unknown and first has to be determined from the data. If r is underestimated, then empirically relevant ECMs will be omitted, possibly misleading policy decisions. Conversely, if r is overestimated, the distributions of statistics will be non-standard, so incorrect inferences will result from using conventional critical values in tests, and forecasts may be badly behaved. The test for r cointegrating vectors in Johansen (1988) is equivalent to testing whether $\mathbf{\Pi} = \alpha \beta'$, where α and β are $N \times r$, and hence are tests of $\mathbf{\Pi}$ for reduced rank. If $\mathbf{\Pi}$ were unrestricted, a conventional regression estimator would result.

Once the degree of cointegration has been established, the estimated cointegrating combinations are denoted by $\widehat{\beta}' \mathbf{x}_t$, and these linear combinations of the data are the estimated ECMs. Moreover, α reveals the importance of each cointegrating combination in each equation, and is related to the speed of adjustment after a disequilibrium. If a given ECM enters more than one equation, the parameters are inherently cross-linked between such equations, and hence their dependent variables cannot be weakly exogenous in the related equations (see Engle et al., 1983). This entails that joint estimation is needed for fully efficient estimation. It must be stressed that cointegrating vectors are the only invariants of linear systems under linear transformations, and that they are preserved in all linear representations of the system.

[5]ECM was the name given by Davidson et al. (1978) and also used by Engle and Granger (1987), but as discussed in the Epilogue, in reality, ECMs do not error correct across regime shifts, but correct back to the previous equilibrium, which can be inimical to forecasting, but was not understood at the time of the report.

To analyze forecasting, we re-express the system in $\mathsf{I}(0)$ space in stacked form with a single lag:

$$\mathbf{s}_t = \boldsymbol{\Upsilon}\mathbf{s}_{t-1} + \boldsymbol{\Psi}^*\mathbf{D}_t^* + \epsilon_t \text{ where } \mathbf{s}_t = \begin{pmatrix} \Delta\mathbf{x}_t \\ \vdots \\ \Delta\mathbf{x}_{t-p+2} \\ \beta'\mathbf{x}_{t-p+1} \end{pmatrix}, \ \mathbf{D}_t^* = \begin{pmatrix} \mathbf{D}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \text{ and } \epsilon_t = \begin{pmatrix} \nu_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \tag{20}$$

and:

$$\boldsymbol{\Upsilon} = \begin{pmatrix} \boldsymbol{\Pi}_1 & \boldsymbol{\Pi}_2 & \cdots & \boldsymbol{\Pi}_{p-1} & \alpha \\ \mathbf{I}_N & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \beta' & \mathbf{I}_N \end{pmatrix}, \text{ and } \boldsymbol{\Psi}^* = \begin{pmatrix} \boldsymbol{\Psi} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{21}$$

In (20), $\epsilon_t \sim \mathsf{IN}[\mathbf{0}, \boldsymbol{\Omega}]$ where $\boldsymbol{\Omega}$ is non-negative definite, but possibly singular and $\mathbf{s}_t$ is $\mathsf{I}(0)$. The system can be reformulated by partitioning $\mathbf{x}_t$ into $(\mathbf{x}'_{a,t}, \mathbf{x}'_{b,t})'$ where $\Delta\mathbf{x}_{a,t}$ and $\beta'\mathbf{x}_t$ are $\mathsf{I}(0)$, the former corresponding to the equations with unit roots. Letting $\mathbf{w}_t = (\Delta\mathbf{x}'_{a,t}, \beta'\mathbf{x}'_t)'$ then the long-run solution for the system can be defined in terms of $\mathsf{E}[\mathbf{w}_t]$: this is a major contrast with $\mathsf{E}[\mathbf{x}_t]$, which is not constant. In what follows, we ignore deterministic terms $\boldsymbol{\Psi}^*\mathbf{D}_t^*$ to simplify the algebra.

Equation (20) defines both conditional and unconditional means and variances of all the $\mathsf{I}(0)$ variables:

$$\mathsf{E}[\mathbf{s}_t|\mathbf{s}_{t-1}] = \boldsymbol{\Upsilon}\mathbf{s}_{t-1} \text{ and } \mathsf{V}[\mathbf{s}_t|\mathbf{s}_{t-1}] = \boldsymbol{\Omega} \tag{22}$$

and:

$$\mathsf{E}[\mathbf{s}_t] = \mathbf{0} \text{ and } \mathsf{V}[\mathbf{s}_t] = \mathbf{M} = \boldsymbol{\Upsilon}\mathbf{M}\boldsymbol{\Upsilon}' + \boldsymbol{\Omega} \tag{23}$$

In $\mathsf{I}(0)$ space, the VAR can be interpreted as a model of conditional means and variances.

Given that background, simultaneous equations models (SEMs) have the form:

$$\boldsymbol{\Phi}\mathbf{s}_t = \mathbf{u}_t \tag{24}$$

where $\boldsymbol{\Phi}$ is a $m \times N^*$ matrix of structural parameters with $N^* = N(p-1) + r$ and $m < N$, subject to normalization and sufficient a priori identifying restrictions. Equations (20) and (24) imply that:

$$\boldsymbol{\Phi}\mathbf{s}_t = \boldsymbol{\Phi}\boldsymbol{\Upsilon}\mathbf{s}_{t-1} + \boldsymbol{\Phi}\epsilon_t = \mathbf{u}_t \tag{25}$$

The errors $\mathbf{u}_t$ are usually assumed by investigators to be serially uncorrelated, but this will only be the case if $\boldsymbol{\Phi}\boldsymbol{\Upsilon} = \mathbf{0}$, so that $\mathbf{u}_t = \boldsymbol{\Phi}\epsilon_t$. If $\boldsymbol{\Phi}\boldsymbol{\Upsilon} = \mathbf{0}$, then (24) is dynamically well specified, but such a condition cannot be known to hold a priori and so has to be tested relative to (20). When $\boldsymbol{\Phi}\boldsymbol{\Upsilon} \neq \mathbf{0}$, (24) is dynamically mis-specified in that $\mathbf{u}_t$ will be serially correlated, and hence potentially all other inferences about the elements of $\boldsymbol{\Phi}$ are invalid.

10 Exogeneity

For an open system to be well defined in terms of $\mathsf{I}(0)$ modelled variables, all of the initial non-stationarity must be inherited from the non-modelled variables. If the locations of the unit roots and cointegrating vectors are known, the system as a whole can be mapped into $\mathsf{I}(0)$ space where conventional asymptotics apply. Modify the transformed partition of $\mathbf{x}_t$ into $\mathbf{w}_t$ as $(\mathbf{y}_t : \mathbf{z}_t)$, where the $\mathbf{z}_t$ are to be treated as weakly exogenous for the parameters of interest, denoted μ, deemed to include α and β (see Engle et al., 1983). Factorize the joint sequential density of $\mathbf{w}_t$ accordingly as:

$$\mathsf{D}_{w_t}(\mathbf{w}_t|\mathbf{s}_{t-1},\theta) = \mathsf{D}_{y_t|z_t}(\mathbf{y}_t|\mathbf{z}_t,\mathbf{s}_{t-1},\theta_1)\mathsf{D}_{z_t}(\mathbf{z}_t|\mathbf{s}_{t-1},\theta_2) \tag{26}$$

noting that information on the cointegrating vectors is retained by $\mathbf{s}_{t-1}$. Then, $\mathbf{z}_t$ is weakly exogenous for μ if μ depends on θ_1 alone, and θ_1 and θ_2 are variation free. Such a requirement is violated if both θ_1 and θ_2 depend on the same β_i, which is testable from the procedure in Johansen (1988) and Johansen and Juselius (1990), whereas linking short-run parameters with β_i need not matter for some inferences, although it may do for forecasting.

11 Structural econometric models

From (26), $\mathbf{y}_t$ is the vector of m endogenous variables to be explained by the open system, and $\mathbf{z}_t$ is the vector of $N-m$ non-modelled variables, both of which could enter the model with lags of up to p periods. We only consider linear stochastic models in this section, and write the model (24) in extensive form as:

$$\mathbf{B}_0\mathbf{y}_t + \mathbf{B}_1\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p\mathbf{y}_{t-p} = \sum_{i=0}^{p}\mathbf{C}_i\mathbf{z}_{t-i} + \xi_t \tag{27}$$

so that $\mathbf{B}_0$ is the $m\times m$ matrix of coefficients of all the endogenous variables, $\mathbf{B}_1$ is the coefficient matrix of the 1-lagged endogenous variables ..., $\mathbf{C}_0$ is the $(N-m)\times(N-m)$ matrix of the current dated non-modelled variables, etc. Providing the model is a congruent representation of the system, ξ_t has a well-defined error distribution, taken here to be an independent normal distribution with mean zero and covariance matrix $\mathbf{\Sigma}_\xi$.

The structural form (27) often claims to model the decision equations of the relevant agents. The reduced, or solved, form finds $\mathbf{y}_t$ as a function of the givens, i.e., the parameters, non-modelled variables and lagged variables, and corresponds to the top block of the VAR in (17), in $\mathsf{I}(0)$ space:

$$\mathbf{y}_t = \mathbf{f}\left(\mathbf{y}_{t-1},\ldots,\mathbf{y}_{t-p},\mathbf{z}_t,\ldots,\mathbf{z}_{t-p},\xi_t\right) \tag{28}$$

If the system is linear, a necessary condition to be able to solve it for the outcomes is that $\mathbf{B}_0$ is non-singular, in which case, a solution is given by:

$$\begin{aligned}\mathbf{y}_t &= (\mathbf{B}_0)^{-1}\sum_{i=0}^{p}\mathbf{C}_i\mathbf{z}_{t-i} - (\mathbf{B}_0)^{-1}\mathbf{B}_1\mathbf{y}_{t-1} - \cdots - (\mathbf{B}_0)^{-1}\mathbf{B}_p\mathbf{y}_{t-p} + (\mathbf{B}_0)^{-1}\xi_t \\ &= \mathbf{\Gamma}_0\mathbf{z}_t + \mathbf{\Gamma}_1\mathbf{z}_{t-1} + \cdots + \mathbf{\Gamma}_p\mathbf{z}_{t-p} + \mathbf{G}_1\mathbf{y}_{t-1} + \cdots + \mathbf{G}_p\mathbf{y}_{t-p} + \eta_t \end{aligned} \tag{29}$$

Thus, given the data up to and including time $t-1$ on $(\mathbf{y},\mathbf{z})$, together with $\mathbf{z}$ and all the parameters $\left(\mathbf{\Gamma}_0,\mathbf{\Gamma}_1,\ldots,\mathbf{\Gamma}_p,\mathbf{G}_1,\ldots,\mathbf{G}_p\right)$, the next value of $\mathbf{y}_t$ is generated by equation (29).

We have shown in (20) that one lag is fully general and is the only case we need to consider. Hereafter we write the system as:

$$\mathbf{B}_0\mathbf{y}_t + \mathbf{B}_1\mathbf{y}_{t-1} = \mathbf{C}_0\mathbf{z}_t + \mathbf{C}_1\mathbf{z}_{t-1} + \xi_t \tag{30}$$

with reduced form:

$$\mathbf{y}_t = \mathbf{\Gamma}_0\mathbf{z}_t + \mathbf{\Gamma}_1\mathbf{z}_{t-1} + \mathbf{G}_1\mathbf{y}_{t-1} + \eta_t \tag{31}$$

The model also holds one period earlier, so that from (31) and its value at time $t-1$:

$$\begin{aligned}\mathbf{y}_t &= \mathbf{\Gamma}_0\mathbf{z}_t + \mathbf{\Gamma}_1\mathbf{z}_{t-1} + \mathbf{G}_1\left(\mathbf{\Gamma}_0\mathbf{z}_{t-1} + \mathbf{\Gamma}_1\mathbf{z}_{t-2} + \mathbf{G}_1\mathbf{y}_{t-2} + \eta_{t-1}\right) + \eta_t \\ &= \mathbf{\Gamma}_0\mathbf{z}_t + \left(\mathbf{\Gamma}_1 + \mathbf{G}_1\mathbf{\Gamma}_0\right)\mathbf{z}_{t-1} + \mathbf{G}_1\mathbf{\Gamma}_1\mathbf{z}_{t-2} + \mathbf{G}_1^2\mathbf{y}_{t-2} + \eta_t + \mathbf{G}_1\eta_{t-1}\end{aligned} \tag{32}$$

If the autoregressive component of the model is dynamically stable, $\mathbf{G}_1^h \to \mathbf{0}$ as $h \to \infty$ so that:

$$\begin{aligned}\mathbf{y}_t &= \mathbf{\Gamma}_0\mathbf{z}_t + (\mathbf{\Gamma}_1 + \mathbf{G}_1\mathbf{\Gamma}_0)\,\mathbf{z}_{t-1} + \mathbf{G}_1(\mathbf{\Gamma}_1 + \mathbf{G}_1\mathbf{\Gamma}_0)\,\mathbf{z}_{t-2} + \cdots + \eta_t + \mathbf{G}_1\eta_{t-1} + \cdots \\ &\to \mathbf{\Gamma}_0\mathbf{z}_t + \sum_{i=0}^{\infty}\mathbf{G}_1^i(\mathbf{\Gamma}_1 + \mathbf{G}_1\mathbf{\Gamma}_0)\,\mathbf{z}_{t-i-1} + \sum_{i=0}^{\infty}\mathbf{G}_1^i\eta_{t-i}\end{aligned} \tag{33}$$

Thus, due to the dynamics, $\mathbf{y}_t$ is the cumulation of all past $\mathbf{z}_{t-h}$ with declining weights (since $\mathbf{G}_1^h \to \mathbf{0}$). The shocks cumulate as well, again with declining weights if the system is stable. Define the static equilibrium (unconditional expectation) for $\mathsf{E}[\mathbf{z}_t] = \mathbf{z}^* \; \forall t$. Then:

$$\mathsf{E}[\mathbf{y}_t] = \mathbf{y}^* = \mathbf{\Gamma}_0\mathbf{z}^* + \sum_{i=0}^{\infty}\mathbf{G}_1^i(\mathbf{\Gamma}_1 + \mathbf{G}_1\mathbf{\Gamma}_0)\,\mathbf{z}^* = \mathbf{\Xi}\mathbf{z}^*, \tag{34}$$

From (31), an alternative expression for $\mathbf{\Xi}$ is:

$$\mathbf{y}^* = (\mathbf{I} - \mathbf{G}_1)^{-1}(\mathbf{\Gamma}_0 + \mathbf{\Gamma}_1)\,\mathbf{z}^* = \mathbf{\Xi}\mathbf{z}^* \tag{35}$$

so that the long-run multiplier is:

$$\frac{\partial \mathbf{y}^{*\prime}}{\partial \mathbf{z}^*} = \mathbf{\Xi} \tag{36}$$

This shows the total effect of a change in $\mathbf{z}$ on $\mathbf{y}$.

If η_t is not autonomous to the economic system, but is treated as exogenous, incorrect results will be obtained. An example of this mistake is the treatment of incomes policies as exogenous shocks when in fact they are an endogenous response to developments in the economy. More subtle examples occur in the analyses of closed systems with no $\mathbf{z}$ variables so that (31) becomes a VAR, which seeks to attribute the effects of shocks to different endogenous variables.

12 Model cointegration

In practice, since investigators do not know which transformations of the original data are $\mathsf{I}(0)$, initial (and sometimes final) formulations of econometric models have neither $\mathbf{y}_t$ nor $\mathbf{z}_t$ stationary, but evolving. Since economic growth is attributed to such factors as technical progress, knowledge accumulation, population growth etc., closed macroeconomic systems are rare. Thus, $\mathbf{G}_1^h \to \mathbf{0}$ in (33) is reasonable if growth is explained by $\mathbf{z}_t$. However, that requires both that $\mathsf{E}[\Delta\mathbf{z}_t] = \mathbf{g}$ is also reasonable and that sufficient elements of $\mathbf{C}_0$ and $\mathbf{C}_1$ are non-zero. Thus, the non-modelled variables alone have the unit roots.

Unit roots can affect the consistency of (30) in that combinations of orders of integration for $\mathbf{y}_t$ and $\mathbf{z}_t$ need not be mutually consistent with the assertion that η_t is $\mathsf{I}(0)$. For example, if $\mathbf{y}_t$ is $\mathsf{I}(1)$ (say), the unit roots must cancel between $\mathbf{B}_0\mathbf{y}_t$ and $(\mathbf{B}_1\mathbf{y}_{t-1} + \mathbf{C}_0\mathbf{z}_t + \mathbf{C}_1\mathbf{z}_{t-1})$. One way of doing so is if $\mathbf{B}_1 = \mathbf{B}_0$ and $\mathbf{C}_1 = -\mathbf{C}_0$, so that (30) really holds in differences:

$$\mathbf{B}_0\Delta\mathbf{y}_t = \mathbf{C}_0\Delta\mathbf{z}_t + \xi_t \tag{37}$$

Now all terms are $\mathsf{I}(0)$ if the original data are $\mathsf{I}(1)$, and no issue of balance arises. However, with no lags, no forecasting power remains either unless $\Delta\mathbf{z}_t$ is forecastable. Moreover, (37) has a more serious drawback: the growth rates of economic variables are made independent of their historical levels and of any past disequilibria, which seems counter factual. Falsely assuming that (37) is true can adversely affect policy by neglecting the feedbacks of past disequilibria on to present outcomes: for example, in the housing market, the history of mortgage rationing led to a large backlog to be cleared.

When $\mathbf{y}_t$ and $\mathbf{z}_t$ are $\mathsf{I}(1)$ and cointegrated, the static solution derived above characterizes the cointegrating equations so that $(\mathbf{y}_t - \Xi\mathbf{z}_t)$ is $\mathsf{I}(0)$. Thus, the model has the form:

$$\mathbf{B}_0\Delta\mathbf{y}_t = \mathbf{C}_0\Delta\mathbf{z}_t + \mathbf{D}_0\left(\mathbf{y}_{t-1} - \Xi\mathbf{z}_{t-1}\right) + \xi_t \tag{38}$$

where every term is $\mathsf{I}(0)$ even though the levels are $\mathsf{I}(1)$. Past disequilibria influence present growth, and if $\mathbf{D}_0$ has the requisite properties (i.e., induces a stable $\mathbf{G}_1$), the system is continually being driven towards the dynamic temporary equilibrium $\mathbf{y}_t = \Xi\mathbf{z}_t$. Terms of the form $(\mathbf{y}_{t-1} - \Xi\mathbf{z}_{t-1})$ are the ECMs at the system level. The Oxford Economic Forecasting (OEF) model has this form, and can be solved statically for the cointegrating equations as $\mathbf{y}_t = \Xi\mathbf{z}_t$, or dynamically for growth paths.

13 Regime shifts

To clarify notation, let θ_t be the parameters of the joint density at t, $\lambda_t = \mathbf{g}(\theta_t)\ \forall t$, and consider the density formulation over the sample where $\mathbf{Y}_{t-1} = (\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots)$ and similarly for $\mathbf{Z}_{t-1}$:

$$\prod_{t=1}^{T}\mathsf{D}_{y_t|z_t}\left(\mathbf{y}_t|\mathbf{z}_t, \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}, \lambda_{1,t}\right)\prod_{t=1}^{T}\mathsf{D}_{z_t}\left(\mathbf{z}_t|\mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}, \lambda_{2,t}\right) \tag{39}$$

where $\lambda'_t = (\lambda'_{1,t}, \lambda'_{2,t})$ The first term is the conditional model of $\mathbf{y}_t$ and the second the marginal model for $\mathbf{z}_t$. If agents form expectations about $\mathbf{z}_t$ when planning $\mathbf{y}_t$, then $\lambda_{1,t}$ will depend on $\lambda_{2,t}$ and so alter when $\lambda_{2,t}$ changes. If $\mathbf{z}_t$ is not observable by the agent at the time the plan is formulated, then the first density must be marginalized with respect to $\mathbf{z}_t$, or $\mathbf{z}_t$ must be modelled. Both conditional and expectational models can be derived from the joint density of $(\mathbf{y}_t, \mathbf{z}_t)$.

Reconsider (39) but factorized in the alternative direction as:

$$\prod_{t=1}^{T}\mathsf{D}_{z_t|y_t}\left(\mathbf{z}_t|\mathbf{y}_t, \mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}, \gamma_{1,t}\right)\prod_{t=1}^{T}\mathsf{D}_{y_t}\left(\mathbf{y}_t|\mathbf{Y}_{t-1}, \mathbf{Z}_{t-1}, \gamma_{2,t}\right) \tag{40}$$

where $\gamma'_t = (\gamma'_{1,t}, \gamma'_{2,t})$. Because both $\gamma_{1,t}$ and $\gamma_{2,t}$ depend on all elements of λ_t, neither model in (40) can have constant parameters; nor can both marginal models. Tests of this dependence are discussed in Engle and Hendry (1993) and Favero and Hendry (1992). In practice, there is almost no evidence for such dependencies being due to ignoring expectations processes. Moreover, the power of tests for the Lucas critique may be low, suggesting that it lacks force. Conversely, there is plenty of evidence of model failure due to mis-specification, so resources may be better allocated to improving equation forms than to re-interpreting equations as expectational.

14 Encompassing

Consider two rival SEMs of the form:

$$\mathsf{H}_a : \mathbf{\Phi}_a\mathbf{s}_t = \mathbf{u}_{a,t} \ \text{ and } \ \mathsf{H}_b : \mathbf{\Phi}_b\mathbf{s}_t = \mathbf{u}_{b,t} \tag{41}$$

which are overidentified relative to the congruent statistical system (24) where $\mathbf{\Phi}_a$ and $\mathbf{\Phi}_b$ are $m_a \times N^*$ and $m_b \times N^*$ respectively. The SEMs are assumed to be complete (i.e. m_a and m_b equal m) and are $\mathsf{I}(0)$. This ensures that the correct critical values of tests are used in practice (e.g. not conventional ones if the VAR is $\mathsf{I}(1)$ and the SEM is $\mathsf{I}(0)$); we assume a mapping to $\mathsf{I}(0)$ prior to testing encompassing hypotheses. Subtle issues arise when the SEMs are incomplete as they must be explaining the same set of variables (see Hendry and Richard, 1989). Here, we restrict attention to complete systems.

The error covariance matrices of each specification in (41) are given by $\mathbf{\Sigma}_{aa} = \mathbf{\Phi}_a \mathbf{M} \mathbf{\Phi}'_a$ and $\mathbf{\Sigma}_{bb} = \mathbf{\Phi}_b \mathbf{M} \mathbf{\Phi}'_b$ respectively, so that H_a dominates H_b in generalized variance (or likelihood) if $\det(\mathbf{\Sigma}_{aa}) \leq \det(\mathbf{\Sigma}_{bb})$. Although generalized variance dominance provides a natural criterion for model comparison, it is only a necessary condition for H_a to encompass H_b (denoted $\mathsf{H}_a \mathcal{E} \mathsf{H}_b$). When the DGP for $\mathbf{x}_t$ is not known, the congruent statistical system, in this case the VAR (17), rewritten as in (20), provides a common probability space within which the rival models can be compared.

Let γ_b denote the vector of parameters in $\mathbf{\Phi}_b$ which characterizes H_b, noting that the value of γ_b is implied by (20), and hence is a function of $\mathbf{\Upsilon}$ and $\mathbf{\Omega}$. Let γ_p denote what H_a predicts γ_b to be when H_a is assumed to be the DGP. Then:

$$\mathsf{H}_a \text{ encompasses } \mathsf{H}_b \quad (\mathsf{H}_a \mathcal{E} \mathsf{H}_b) \text{ if and only if } \quad (\gamma_b - \gamma_p) = \mathbf{0} \tag{42}$$

Both H_a and H_b are nested within the VAR (20)—ignoring deterministic terms as being in common—which we denote by $\mathsf{H}_i \subset$VAR. It is clear that the VAR automatically encompasses H_i by virtue of each H_i being a reduction of the VAR. The issue here is whether either of H_a or H_b encompasses the VAR. If so, since a simple model is accounting for the results of a more general model within which it is nested, we call this parsimonious encompassing of the VAR, denoted $\mathsf{H}_a \mathcal{E}_p$VAR, or $\mathsf{H}_b \mathcal{E}_p$VAR. For further details on encompassing, see Mizon (1984), Mizon and Richard (1986), and Hendry and Richard (1989).

The main difficulty in establishing encompassing theorems about SEMs is that their exogeneity assumptions may differ, so that the models are not in a common probability space. This problem does not arise with closed systems if either H_a or H_b asserts that $\mathsf{E}_i[\mathbf{u}_{i,t}\mathbf{s}'_{t-1}] = \mathbf{0}$ for $i = a, b$ where E_i denotes expectations with respect to model i. Consider the case when $i = a$. Since $\mathsf{E}_a[\mathbf{u}_{a,t}\mathbf{s}'_{t-1}] = \mathbf{0}$ implies that the derived error $\mathbf{u}_{a,t}$ is an innovation against the past of the process, from (20):

$$\mathsf{E}[\mathbf{\Phi}_a \mathbf{s}_t \mathbf{s}'_{t-1}] = \mathbf{\Phi}_a \mathbf{\Upsilon} \mathsf{E}[\mathbf{s}_{t-1}\mathbf{s}'_{t-1}] \tag{43}$$

which must be zero $\forall t$, and hence:

$$\mathbf{\Phi}_a \mathbf{\Upsilon} = \mathbf{0} \tag{44}$$

This is precisely the earlier condition for the absence of dynamic mis-specification. If (44) holds, then H_a is a valid reduction of the VAR and hence H_a parsimoniously encompasses the VAR. Conversely, if (44) does not hold, H_a cannot encompass the VAR since H_a can be rejected against the VAR in that elements of $\mathbf{s}_{t-1}$ must be relevant to H_a in addition to the variables included through its definition in (20). If neither modeller accepts the validity of the instruments used by the other or there are insufficient instruments to identify the joint or nesting hypothesis, considerable problems can ensue and cross-model encompassing may not be testable. Thus, conditional on the proprietor of H_a accepting lagged variables as legitimate potential explanatory variables:

Theorem 1: $\mathsf{H}_a \mathcal{E}$VAR if and only if $\mathbf{\Phi}_a \mathbf{\Upsilon} = \mathbf{0}$.

A consequence of Theorem 1, is:

Theorem 2: If $\mathsf{H}_a \mathcal{E}_p$VAR and $\mathsf{H}_i \subset$VAR then $\mathsf{H}_a \mathcal{E} \mathsf{H}_i$.

Thus, the VAR is a catalyst for allowing encompassing comparisons between SEMs. The converse of Theorem 2 is false in that $\mathsf{H}_a \mathcal{E} \mathsf{H}_i$ could hold without $\mathsf{H}_a \mathcal{E}_p$VAR (e.g., because $\mathsf{H}_i \subset \mathsf{H}_a$), but by requiring congruence, H_a would be rejected against the VAR in our framework.

In practice, for models the size of that used in H.M. Treasury, system parameter, or even variance, encompassing is hard to implement. Thus, the alternative of forecast encompassing (see Lu and Mizon, 1991) is useful, namely can H_a explain the forecast errors of H_b? This is discussed in section 6.

15 System forecasting

Consider the slightly simplified reduced form system from (31) defined by:

$$\mathbf{y}_t = \boldsymbol{\Gamma}_0 \mathbf{z}_t + \mathbf{G}_1 \mathbf{y}_{t-1} + \eta_t \tag{45}$$

where $\mathbf{w}'_t = (\mathbf{z}'_t : \mathbf{y}'_{t-1})$ rather than as in §10 for the conditional distribution in (26). The $\mathbf{z}_t$ are the non-modelled and policy variables (including their lags), where $\mathbf{y}_t$ and $\mathbf{z}_t$ are $\mathsf{I}(1)$ but cointegrated. One-step forecasts from (45) at time T are given by:

$$\widehat{\mathbf{y}}_{T+1} = \widehat{\boldsymbol{\Gamma}}_0 \mathbf{z}_{T+1} + \widehat{\mathbf{G}}_1 \mathbf{y}_T \tag{46}$$

with forecast errors *ex post* when $\mathbf{z}_{T+1}$ is known (as in a policy context):

$$\begin{aligned} \mathbf{y}_{T+1} - \widehat{\mathbf{y}}_{T+1} &= \left(\boldsymbol{\Gamma}_0 - \widehat{\boldsymbol{\Gamma}}_0\right) \mathbf{z}_{T+1} + \left(\mathbf{G}_1 - \widehat{\mathbf{G}}_1\right) \mathbf{y}_T + \eta_{T+1} \\ &= \left(\mathbf{G} - \widehat{\mathbf{G}}\right) \mathbf{w}_{T+1} + \eta_{T+1} = \widehat{\eta}_{T+1} \end{aligned} \tag{47}$$

For testing for predictive failure, 1-step forecast errors are calculated as:

$$\widehat{\eta}_{T+1} = \left(\mathbf{I} \otimes \mathbf{w}'_{T+1}\right) \left(\mathbf{G} - \widehat{\mathbf{G}}\right)^v + \eta_{T+1} \tag{48}$$

where $(\cdot)^v$ denotes vectoring. Then to a first approximation when $\widehat{\mathbf{G}}$ is unbiased for $\mathbf{G}$ (see e.g., Hendry and Trivedi, 1972):

$$\mathsf{E}\left[\widehat{\eta}_{T+1}\right] = \mathbf{0} \text{ and } \mathsf{V}\left[\widehat{\eta}_{T+1}\right] = \boldsymbol{\Omega} + \left(\mathbf{I} \otimes \mathbf{w}'_{T+1}\right) \mathsf{V}\left[\widehat{\mathbf{G}}\right] \left(\mathbf{I} \otimes \mathbf{w}_{T+1}\right) = \boldsymbol{\Theta}_1 \tag{49}$$

so that (see Calzolari, 1981, and Chong and Hendry, 1986):

$$\widehat{\eta}_{T+1} \approx \mathsf{IN}\left[\mathbf{0}, \boldsymbol{\Theta}_1\right] \tag{50}$$

where $\boldsymbol{\Theta}_1$ reflects both parameter and innovation uncertainty.

In an operational context, *ex ante* statements about an economy require knowledge of all of the conditioning variables when the forecast horizon H exceeds unity. This necessitates a closed system, or 'off-line' forecasts of $\mathbf{z}_{T+1}$. From (20):

$$\mathbf{s}_t = \boldsymbol{\Upsilon} \mathbf{s}_{t-1} + \epsilon_t \tag{51}$$

where $\epsilon_t \sim \mathsf{IN}[\mathbf{0}, \boldsymbol{\Omega}]$ which is an innovation against $\mathbf{s}_{t-1}$. Then (51) is a stacked version of the system in $\mathsf{I}(0)$ space so $\boldsymbol{\Upsilon}$ depends on all the parameters of that system, which we denote by the vector of parameters δ to distinguish from the structural parameters. Since MLEs are equivariant to functional transformations, the MLE of $\boldsymbol{\Upsilon}$ is given by $\widehat{\boldsymbol{\Upsilon}} = \boldsymbol{\Upsilon}(\widehat{\delta})$ from which $\mathsf{V}[\widehat{\boldsymbol{\Upsilon}}]$ can be derived.[6] The optimal predictor of a mean-zero martingale difference sequence is zero, so $\mathsf{E}[\epsilon_{T+i}] = \mathbf{0}$. Consequently, once $\mathbf{s}_T$ is known, the best predictor of $\mathbf{s}_{T+i}$ for $i > 0$ is:

$$\widehat{\mathbf{s}}_{T+i} = \widehat{\boldsymbol{\Upsilon}}^i \mathbf{s}_T \tag{52}$$

with forecast errors $\widehat{\epsilon}_{T+i} = \mathbf{s}_{T+i} - \widehat{\mathbf{s}}_{T+i}$ The derivation of the variance matrix $\mathbf{V}_i = \mathsf{V}[\widehat{\epsilon}_{T+i}]$ is complicated so is not presented here. The model must continue to characterize the economy as well in the forecast period as it did over the estimation sample if the prediction errors are to be from the same distribution as that assumed, namely $\mathsf{IN}[\mathbf{0}, \boldsymbol{\Omega}]$, ignoring the sampling variation due to estimating $\boldsymbol{\Upsilon}$. This is a strong requirement, and is unlikely to be met unless $\boldsymbol{\Upsilon}$ is constant within sample: the further condition that $\boldsymbol{\Upsilon}$ is invariant to any regime changes out-of-sample is required. Even if $\boldsymbol{\Upsilon}$ is both constant and invariant, $\mathbf{V}_i$ increases with i, and the forecast errors will be heteroscedastic and serially correlated, so care is required in interpreting forecast errors.

[6] $\boldsymbol{\Psi}^*$ in (20) also needs estimated in practice jointly with $\boldsymbol{\Upsilon}$.

16 Conditional dynamic forecasts

To see where all sources of uncertainty enter, we explicitly express the model in a conditional form. When $\mathbf{y}_t$ is generated dynamically, multi-period forecasts (denoted by $\widetilde{}$) must mimic the dynamics by successively calculating $\mathbf{y}_{t+i}$, feeding back previously generated $\mathbf{y}_{t+i-j}$ as in (ignoring parameter estimation uncertainty to focus on the dynamics):

$$\widetilde{\mathbf{y}}_{T+i} = \mathbf{\Gamma}_0 \mathbf{z}_{T+i} + \mathbf{G}_1 \widetilde{\mathbf{y}}_{T+i-1} \tag{53}$$

Subtracting (53) from the conditional reduced form (45) yields:

$$\mathbf{y}_{T+i} - \widetilde{\mathbf{y}}_{T+i} = \mathbf{G}_1 \left(\mathbf{y}_{T+i-1} - \widetilde{\mathbf{y}}_{T+i-1} \right) + \eta_{T+i} \tag{54}$$

Thus, even for known $\mathbf{z}_{T+i}$, $\widetilde{\mathbf{y}}_{T+i}$ can differ from $\mathbf{y}_{T+i}$ sometimes drastically. From (54), the stochastic simulation error $\widetilde{\eta}_{T+i} = \mathbf{y}_{T+i} - \widetilde{\mathbf{y}}_{T+i}$ follows the same autoregressive process as the model. Thus, iterating:

$$\widetilde{\eta}_{T+i} = \sum_{j=0}^{i-1} \mathbf{G}_1^j \eta_{T+i-j} \tag{55}$$

so the multi-step forecast error process $\widetilde{\eta}_{T+i}$ is autocorrelated and heteroscedastic.

In any operational environment of dynamic forecasting, parameters must be estimated; the future values of the $\mathbf{z}_{T+i}$ will be unknown and must be forecast as well (e.g., world trade, oil prices); and initial conditions will be incorrect. We denote parameter estimates by $\widehat{}$, off-line forecasts by $\overline{}$, and endogenous variables forecasts by $\widetilde{}$. Thus, for $j = 1, \ldots, H$, dynamic forecasts must iteratively solve:

$$\widetilde{\mathbf{y}}_{T+j} = \widehat{\mathbf{\Gamma}}_0 \overline{\mathbf{z}}_{T+j} + \widehat{\mathbf{G}}_1 \widetilde{\mathbf{y}}_{T+j-1} \tag{56}$$

commencing from the initial conditions $\widetilde{\mathbf{y}}_T$ and $\overline{\mathbf{z}}_{T+1}$, which may differ from the true values $\mathbf{y}_T$ and $\mathbf{z}_{T+1}$

For forecasting more than one period ahead outside of the historical sample, the errors on (56) behave like those in (55). If $\widehat{\mathbf{G}}_1$ has a root of unity, the forecast errors will cumulate and will have a trending variance as in section 7. In I(0) space, however, the forecast error variance will be bounded. Since model forecasts are statistics, subject to sampling fluctuations, they have standard errors. In the simplest case when parameter values, initial conditions and future $\mathbf{z}_{T+i}$ are known and correct, the forecast errors cumulate as in (55). Since forecasts from a correctly specified linear model are essentially unbiased, the forecast error variance from (55) is given by:

$$\mathsf{E}\left[\widetilde{\eta}_{T+j}\widetilde{\eta}'_{T+j}\right] = \mathsf{E}\left[\sum_{i=0}^{j-1}\sum_{k=0}^{j-1} \mathbf{G}_1^i \eta_{T+j-i} \eta'_{T+j-k} (\mathbf{G}_1^k)'\right] \tag{57}$$

As before, forecast standard errors increase with the horizon ahead. If all the latent roots of $\mathbf{G}_1$ lie inside the unit circle, the expression in (57) will converge to the unconditional variance of $\mathbf{y}_t$, given current and past $\mathbf{z}_t$. Otherwise, it will diverge.

It is not always wise to focus on high frequency forecasts even if the forecast period is frequent. For example, when a period t denotes a quarter, an annual average forecast may be more useful in many circumstances. Denote these annual average forecasts by $\widehat{}$. For the simplest case of complete knowledge, except for the innovation error, where J denotes the 4th quarter of year T, so $J = T/4$, let:

$$\widehat{\eta}_{T,J} = \frac{1}{4} \sum_{i=0}^{3} \widetilde{\eta}_{J-i} \tag{58}$$

Thus, the variance of the annual mean forecast error will be one sixteenth of the sum of the variances plus the covariances, which will usually lie between one half and one quarter of the final period variance.

When the future $\mathbf{z}_{T+i}$ have to be forecast, initial conditions are uncertain and parameters are estimated, more complicated results are obtained:

$$\begin{aligned}\mathbf{y}_{T+j} - \widetilde{\mathbf{y}}_{T+j} &= \mathbf{\Gamma}_0\mathbf{z}_{T+j} - \widehat{\mathbf{\Gamma}}_0\overline{\mathbf{z}}_{T+j} + \mathbf{G}_1\mathbf{y}_{T+j-1} - \widehat{\mathbf{G}}_1\widetilde{\mathbf{y}}_{T+j-1} + \eta_{T+j} \\ &= \left(\mathbf{\Gamma}_0 - \widehat{\mathbf{\Gamma}}_0\right)\mathbf{z}_{T+j} + \widehat{\mathbf{\Gamma}}_0\left(\mathbf{z}_{T+j} - \overline{\mathbf{z}}_{T+j}\right) + \left(\mathbf{G}_1 - \widehat{\mathbf{G}}_1\right)\mathbf{y}_{T+j-1} \\ &+ \widehat{\mathbf{G}}_1\left(\mathbf{y}_{T+j-1} - \widetilde{\mathbf{y}}_{T+j-1}\right) + \eta_{T+j} \end{aligned} \tag{59}$$

There are four obvious sources of error in (59), and four non-obvious ones, namely:
a] Parameter uncertainty, in the form of $\left(\mathbf{\Gamma}_0 - \widehat{\mathbf{\Gamma}}_0\right)$ and $\left(\mathbf{G}_1 - \widehat{\mathbf{G}}_1\right)$;
b] Non-modelled variable uncertainty, in the form $(\mathbf{z}_{T+j} - \overline{\mathbf{z}}_{T+j})$;
c] Cumulative mistakes in forecasting endogenous variables, from $(\mathbf{y}_{T+j-1} - \widetilde{\mathbf{y}}_{T+j-1})$;
d] The innovation error, η_{T+j};
e] Treating the $\overline{\mathbf{z}}_{T+j}$ as given, when in fact $\mathbf{y}_{T+j-1}$ affects $\mathbf{z}_{T+j}$;
f] Assuming $(\mathbf{\Gamma}_0, \mathbf{G}_1)$ to be constant, when they alter;
g] Incorrectly measuring the initial values, so $\widetilde{\mathbf{y}}_T \neq \mathbf{y}_T$ and $\overline{\mathbf{z}}_{T+1} \neq \mathbf{z}_{T+1}$;
h] The model is not a correct specification of the economic mechanism so the η_{T+j} are not independent homoscedastic Normal.
These eight sources of forecast error are discussed in section 8.

17 Non-linearity

Many of the features established above for linear models carry over to non-linear systems, although some do not. In particular, deterministic solutions do not coincide with the average of stochastic replications, and explicit model solutions are often not feasible for a non-linear system, so that system congruence is harder to check and systems have to be solved numerically. We can write the general case as:

$$\mathbf{f}\left(\mathbf{y}_t, \mathbf{y}_{t-1}, \mathbf{z}_t, \mathbf{z}_{t-1}; \eta_t | \theta\right) = \mathbf{0} \tag{60}$$

where θ denotes the vector of parameters. There are many algorithms for solving models, but Gauss–Seidel is the most common. To calculate a multiplier analog, care is needed since the numerical sizes of multipliers depend on the state of the system. Let one z, say $z_{j,t}$, change by an amount $\delta_{j,t}$. Then, for an impact multiplier, using a first-order Taylor approximation:

$$\frac{\partial \mathbf{y}_t}{\partial \mathbf{z}_{j,t}} = \frac{\mathbf{f}\left(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{z}_t + \delta_{j,t}\iota, \mathbf{z}_{t-1}; \theta\right) - \mathbf{f}\left(\cdot\right)}{\delta_{j,t}} \tag{61}$$

where ι is a zero vector except for unity in the jth position. Thus, the procedure is: solve for $\mathbf{y}_t$, change or perturb $\mathbf{z}_t$, solve again for the changed $\mathbf{y}_t$, and use $\Delta\mathbf{y}_{i,t}/\delta_{j,t}$ as the multiplier. In practice, large models seem close to log-linearity locally other than having equations in logs but linear identities, so this method works reasonably well. Figure 7 illustrates by a graph for a linear system, where the sequence of multipliers can be calculated from the differences between the two trajectories, divided by the perturbation.

Since deterministic forecasts are obtained by solving the model out of sample, unbiased forecasts and forecast standard errors can be found by randomly drawing perturbations in the solution process and averaging across the results for the mean forecast. The empirical variance around that mean then estimates the forecast uncertainty, determined by whatever factors are perturbed. Although this process is computationally intensive, forecast standard errors are so useful that effort should be devoted to obtaining and reporting them.

18 The artificial data model

Denote the three variables corresponding to "consumption", "income" and "saving" by C, Y and S respectively. Then the system used to generate the data was:

$$\Delta C_t = 0.05 + 0.5\Delta Y_t + 0.2S_{t-1} + \epsilon_{1,t} \tag{62}$$

$$\Delta Y_t = 0.05 + 0.5\Delta Y_{t-1} + \epsilon_{2,t} \tag{63}$$

$$S_t = Y_t - C_t \tag{64}$$

where $\epsilon_{i,t} \sim \mathsf{IN}[0, \sigma_{ii}]$ with $\mathsf{E}[\epsilon_{1,t}\epsilon_{2,s}] = 0 \ \forall t, s$, and $\sigma_{11} = 0.02$ and $\sigma_{22} = 0.05$. Solving (62) and (63) for the reduced form consumption equation yields:

$$\Delta C_t = 0.075 + 0.25\Delta Y_{t-1} + 0.2S_{t-1} + \epsilon_{1,t} + 0.5\epsilon_{2,t} \tag{65}$$

so that the savings equation is:

$$S_t = -0.025 + 0.25\Delta Y_{t-1} + 0.8S_{t-1} - \epsilon_{1,t} + 0.5\epsilon_{2,t} \tag{66}$$

Thus, equations (63) and (66) define the 2-equation model in $\mathsf{I}(0)$ space.

In terms of the $\boldsymbol{\Upsilon}$ matrix, where $\mathbf{s}_t = (\Delta Y_t, S_t)$:

$$\mathbf{s}_t = \begin{pmatrix} 0.050 \\ -0.025 \end{pmatrix} + \begin{pmatrix} 0.50 & 0.0 \\ 0.25 & 0.8 \end{pmatrix} \mathbf{s}_{t-1} + \begin{pmatrix} \epsilon_{2,t} \\ -\epsilon_{1,t} + 0.5\epsilon_{2,t} \end{pmatrix} \tag{67}$$

This representation retains the correct number of differenced variables and cointegrating combinations. If fewer of the latter were kept, policy would suffer since important long-run connections would be ignored; if more of the latter were kept, they must be spurious and actually be $\mathsf{I}(1)$ functions, which would, therefore, have a detrimental effect on forecasting.

If the model is reduced to using only growth rates, to ensure that only $\mathsf{I}(0)$ variables are included, as essentially occurs with VAR models in the U.S.A. (see e.g., Sims, 1980), the two equations become:

$$\Delta C_t = \alpha_{1,0} + \alpha_{1,1}\Delta Y_{t-1} + \xi_{1,t} \tag{68}$$

$$\Delta Y_t = \alpha_{2,0} + \alpha_{2,1}\Delta Y_{t-1} + \xi_{2,t} \tag{69}$$

where $\xi_{1,t}$ contains S_{t-1}. Further lags may be needed to remove the resulting residual autocorrelation. This representation ensures bounded forecast error variances, but at the cost of forecast efficiency, and possibly poor policy advice, since past disequilibria are often important in determining future behaviour.

If the model is left in levels, it can be written as:

$$C_t = 0.075 + 0.45Y_{t-1} - 0.25Y_{t-2} + 0.8C_{t-1} + \epsilon_{1,t} + 0.5\epsilon_{2,t} \tag{70}$$

$$Y_t = 0.05 + 1.5Y_{t-1} - 0.5Y_{t-2} + \epsilon_{2,t} \tag{71}$$

All of the variables in (70) and (71) are $\mathsf{I}(1)$ and have trending means and variances, so inference is suspect unless appropriate critical values are used for tests: these values differ according to whether the test happens to involve $\mathsf{I}(0)$ or $\mathsf{I}(1)$ combinations. Many statistics have non-standard distributions with much larger critical values than conventional ones. The forecast variances of the levels trend as the horizon increases: see figure 6. Concerning the predictability of economic models, the various powers of $\boldsymbol{\Upsilon}$ are as follows:

$$\boldsymbol{\Upsilon}^2 = \begin{pmatrix} 0.25 & 0.00 \\ 0.325 & 0.64 \end{pmatrix} \text{ and } \boldsymbol{\Upsilon}^3 = \begin{pmatrix} 0.125 & 0.00 \\ 0.385 & 0.51 \end{pmatrix} \tag{72}$$

with:
$$\mathbf{\Upsilon}^4 = \begin{pmatrix} 0.063 & 0.00 \\ 0.339 & 0.41 \end{pmatrix} \text{ and } \mathbf{\Upsilon}^5 = \begin{pmatrix} 0.031 & 0.00 \\ 0.287 & 0.33 \end{pmatrix} \tag{73}$$

All ability to predict ΔY_t has vanished by 5 period with the coefficient being just 0.031, and little remains for S_t This matches the illustration presented in the text, and for stationary combinations of series, seems a realistic implication.

19 Conclusions and recommendations

Of the available forecasting methods surveyed in section 2, macro-econometric systems offer the only sustainable and progressive alternative. They are resource intensive and as yet far from perfect, but they consolidate empirical knowledge and allow a rational basis for forecast judgements to be made. An important key to avoiding the most egregious forecast errors from econometric systems is to express them in terms of stationary variables. This is not simply a matter of predicting changes rather than levels, but of having all variables in differences or cointegrated linear combinations. Otherwise, serious forecast errors can arise due to retaining non-stationary functions in the forecasts.

Conditional forecasts form an essential component of economic policy decisions by governmental agencies. Unconditional forecasts will rarely be useful *per se*, but may provide a check on other forecasting procedures. Long-horizon forecasts from macro-econometric systems expressed in an appropriate stationary form converge on the unconditional forecasts. Improvements in data accuracy would obviously help greatly, but are not a panacea since many other areas also need improvement.

Pooling of forecasts offers little direct hope for increased forecast accuracy, and indeed the need to pool is direct evidence of non-encompassing of the models involved. However, forecast pooling may have a useful role in improving the accuracy of initial conditions and the forecasts of non-modelled variables, especially when the second source is an unconditional or survey based forecast. Tests between time-series models and econometric systems are useful in appraising the validity of the latter. Selecting models with minimum mean square forecast errors is an inadequate strategy, since the model need not encompass the forecasts of competing models, nor have constant parameters.

Countering the Lucas critique, with its associated emphasis on modelling expectations as an input to the models' equations, seems to offer little scope for forecast accuracy improvements. Tests for its presence or absence should be conducted first.

Forecasting at a high frequency, close to that of decision taking by economic agents, and reporting longer period averages should reduce forecast uncertainty.

Forecast confidence assessments should accompany forecasts.

Epilogue

I was surprised at how much was understood in 1991, as well as seeing some key misunderstandings. A theory of forecasting applicable to a wide sense non-stationary world, namely with stochastic trends and location shifts, when the forecasting model differs from the data generation process is present in an embryonic and unformalised way, as is the start of taxonomies of sources of forecast error, both later leading to many developments, summarized in Clements and Hendry (1999). Indeed, taxonomies played a crucial role in ascertaining what aspects of forecasting devices led to systematic errors (e.g., location shifts and incorrect forecast-origin data), and what played secondary roles (e.g., changes in the parameters of zero-mean variables).

Thus, deterministic terms transpired to be more crucial in explaining systematic forecast failure than reflected in the report, as was the fact that other parameter changes were less crucial, despite the findings to that effect in Favero and Hendry (1992) (first presented as the A.W. Phillips Lecture at the 1989 Australasian Meeting of the Econometric Society, so known at the time). That lack of connecting ideas also influenced my major failure to distinguish 'error-correction mechanisms' (ECMs above), as named by Davidson et al. (1978) (DHSY) and continued by Engle and Granger (1987), from 'equilibrium-correction mechanisms' (EqCMs)—although Clive contemplated using the latter for cointegrating vectors. The key problem is that while ECMs 'error correct' **within regimes**, in that they act to eliminate departures from the equilibrium mean, they do not do so **across regime** shifts, but instead correct back to the previous equilibrium. That property can be highly inimical to forecasting: not only does systematic failure result, but the forecasts move in the opposite direction to the shift to re-establish the in-built equilibrium mean.

Figure 11 clarifies this effect, which was actually occurring to the Treasury forecasts at the time, fooling the then Chancellor of the Exchequer, Norman Lamont, into seeing 'green shoots of Spring' in the economy that were in fact non-existent. The forecasts are from the 'saving' model matching the in-sample artificial data generation process, and are for 1-step ahead after an increase in equilibrium 'saving', taking the value of the change in 'income' as known to highlight the impact of the location shift. The arrows show that each successive forecast after the first is **below the previous outcome**, hence showing a sequence of forecasts of falls when 'saving' has risen above the previous equilibrium mean.

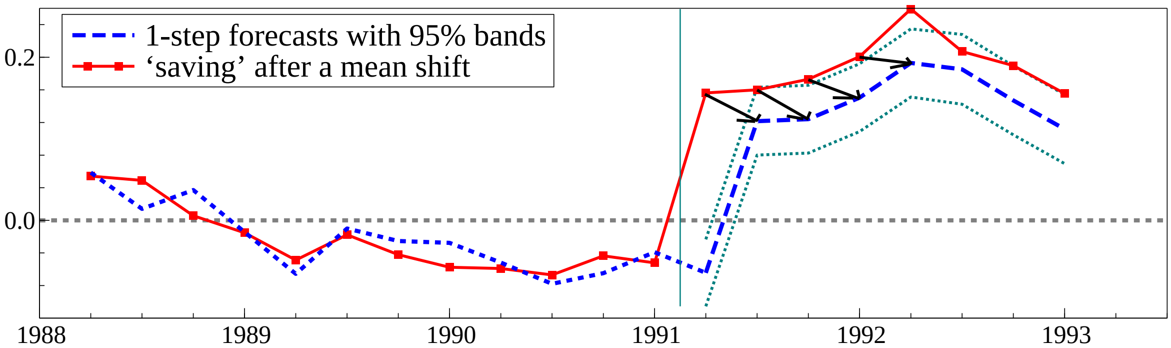

Figure 11: Eight 1-step ahead forecasts of 'saving' with 95% interval forecasts over 1991(1)–1993(1) after a mean shift.

A phone conversation from the Treasury some time after the report telling me that the consumers' expenditure equation based on DHSY had gone off track made me wonder if the major financial innovations of the 1980s, which I had been modelling in terms of shifts in money-demand equations (see e.g., Hendry, 1985), might also have led to a shift in the equilibrium ratio of expenditure to income. A small amount of algebra then revealed the key distinction between ECMs and EqCMs.

Thus, under-emphasising the roles of location shifts, or distributional shifts more generally although mean shifts are particularly problematic, is the most marked lack, and analyzing then resolving their impacts on modelling and forecasting became a major focus, including modelling shifts themselves as in Hendry (1999) leading to the formalization of the vast class of indicator saturation methods (e.g., Hendry et al., 2008, Johansen and Nielsen, 2009, Ericsson, 2012, Castle et al., 2015, and Pretis et al., 2016).

With colleagues I explored a variety of possible 'solutions' to forecasting after a location shift, including intercept corrections in Hendry and Clements (1994) and Clements and Hendry (1996). The comments in section 8g about intercept corrections being *ad hoc* (other than when adjusting poorly measured initial values) needed substantial alteration as one type, namely that based on adding back the previous forecast error, transpired to offset systematic failure after a location shift. This is visible in Figure 11 from the second forecast onwards, as the magnitude of under-forecasting is relatively constant. The recent developments in Castle et al. (2025) show what a huge opportunity to improve forecasting after breaks was missed by not studying intercept corrections more carefully.

We also considered recursive updating, which also helped, but did so by eliminating the EqCM term and converging on an equation in differences. Indeed, such a strategy was noted above, but lost a variable that was potentially valuable for policy, namely the long-run equilibrium feedback. However, we had also shown that adding back the previous forecast error form of intercept correction also worked by differencing, but only the next forecast error not all the variables. That eventually led to the idea of differencing the variables in the estimated cointegrated representation, so retaining the difference of the EqCM and not just differences of variables (see e.g., Hendry, 2006) and from there to a class of forecasting devices that were robust after location shifts (see e.g. Castle et al., 2015). The important distinction is between differencing before (so no EqCM), or after, estimation. Moving estimation windows offered another possible route, but seemed to rely on a good guess of the appropriate length to trade off rapidly dropping past shifts with increased estimation uncertainty.

The alternative of trying to forecast location shifts before they occurred was another possible solution, but the problem of reduced funding for forecasting kicked in, and it was hard to get support for a research program, although the Leverhulme Foundation generously funded a research professorship for me over 1995–2000 to work on that issue, from which several ideas emerged. These included research showing that leading indicators did not seem to work well (see Emerson and Hendry, 1996), that co-breaking for location shifts had close parallels with cointegration for stochastic trends (see Clements and Hendry, 1999, and Hendry and Massmann, 2007), and that a general theory of forecasting from mis-specified models in wide-sense non-stationary process was viable. Castle et al. (2010) later investigated improving forecasting by modelling a break during its trajectory, but found robust devices did as well.

More recently, we highlighted the impacts of location shifts on economic theory (see Hendry and Mizon, 2014, and https://voxeu.org/article/why-standard-macro-models-fail-crises), vitiating proofs of the law of iterated expectations and the unbiasedness of conditional expectations, and revealing that so-called ‘rational expectations’ are irrational when location shifts occur. Indeed, the lack of time dating of expectations implicitly assumed a stationary process.

Another issue where my views have changed greatly concerns the role in system forecasts of unmodelled variables that are forecast off-line. The open-model taxonomy in Hendry and Mizon (2012) establishes that stringent conditions need to be met when facing shifts to ensure such forecasts are an improvement over just omitting unmodelled variables. This is a surprising result, but then so are many others we have discovered about forecasting, such as that all parameters in a model can be greatly changed with essentially no effects on forecasts provided the long-run means are unaltered.

Although MSFE comparisons were criticised above, it took a year or so to also discover their lack of invariance to linear transforms, such as $y_t = \beta_1 z_t + e_t$ to $y_t - z_t = (\beta_1 - 1)z_t + e_t$, and to multi-step forecasts as in Clements and Hendry (1993), eliciting a storm of comment. A measure invariant to both types of transform, the determinant of the generalized forecast-error second-moment (GFESM), was proposed, but seemed non-operational for multiple forecast horizons H and many endogenous variables N, from needing sufficient forecast-error observations to ensure the $NH \times NH$ matrix was non-singular. However, Hendry and Martinez (2017) show how to extend its use to relatively few forecast errors.

The benefits of combining forecasts from different models versus seeking an encompassing framework are also still under debate, as is the choice between parsimonious models and factor models based on huge data sets with more sources of information, capturing both increased T at higher frequencies and larger N. Both issues will probably remain unresolved when different sources and magnitudes of unanticipated location shifts in different parts of economies continue to happen. Much the same can be said of the role of model selection for forecasting, a topic that we are still researching: see Castle et al. (2021).

The hope expressed after Figure 7 that a 'reasonable idea of the uncertainties inherent in predictions could help produce better policy, and more informed discussion thereof' has partly occurred with the Bank of England introducing fan charts to represent their view of the likely range of forecasts, an excellent step forward in emphasising uncertainty. Unfortunately, the Great Recession puts a damper on the idea that doing so would improve policy. Moreover, the uncertainty due to incorrectly measured initial values is rarely included, although the vast increase in research on 'nowcasting', and its use in practice, is a very welcome development, where our approach is summarised in Castle et al. (2018).

Measurement errors and their variances in non-stationary processes have remained a topic of considerable interest to me, partly analyzed in Hendry (1995) and more recently in Duffy and Hendry (2017). The comments in section 8c are formally analyzed in the context of growth and location shifts, and confirm the ability to determine long-run connections with short-run dynamics being contaminated.

References

Baba, Y., D. F. Hendry, and R. M. Starr (1992). The demand for M1 in the U.S.A., 1960–1988. *Review of Economic Studies 59*, 25–61. https://doi.org/10.2307/2297924

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

Calzolari, G. (1981). A note on the variance of ex post forecasts in econometric models. *Econometrica 49*, 1593–1596. https://www.jstor.org/stable/1911422

Chong, Y. Y. and D. F. Hendry (1986). Econometric evaluation of linear macro-economic models. *Review of Economic Studies 53*, 671–690. https://doi.org/10.2307/2297611

Coen, P. G., E. D. Gomme, and M. G. Kendall (1969). Lagged relationships in economic forecasting. *Journal of the Royal Statistical Society A, 132*, 133–163. https://www.jstor.org/stable/2343782

Davidson, J. E. H., D. F. Hendry, F. Srba, and J. S. Yeo (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *Economic Journal 88*, 661–692. https://doi.org/10.2307/2231972

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*, 427–431. https://doi.org/10.1080/01621459.1979.10482531

Dickey, D. A. and W. A. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*, 1057–1072. https://www.jstor.org/stable/1912517

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. *Econometrica 50*, 987–1007. https://doi.org/10.2307/1912773

Engle, R. F. and C. W. J. Granger (1987). Cointegration and error correction: Representation, estimation and testing. *Econometrica 55*, 251–276. https://www.jstor.org/stable/1913236

Engle, R. F., D. F. Hendry, and J.-F. Richard (1983). Exogeneity. *Econometrica 51*, 277–304. https://doi.org/10.2307/1911990

Ericsson, N. R. (1992). Parameter constancy, mean square forecast errors, and measuring forecast performance: An exposition, extensions, and illustration. *Journal of Policy Modeling 14*, 465–495. https://doi.org/10.1016/0161-8938(92)90017-7

Ericsson, N. R., D. F. Hendry, and H.-A. Tran (1994). Cointegration, seasonality, encompassing and the demand for money in the United Kingdom. In C. Hargreaves (Ed.), *Non-stationary Time-series Analysis and Cointegration*, pp. 179–224. Oxford: Oxford University Press.

Favero, C. and D. F. Hendry (1992). Testing the Lucas critique: A review. *Econometric Reviews 11*, 265–306. https://doi.org/10.1080/07474939208800238

Frisch, R. (1938). Statistical versus theoretical relations in economic macrodynamics. Mimeograph dated 17 July 1938, League of Nations Memorandum. Reproduced by University of Oslo in 1948 with Tinbergen's comments. Contained in Memorandum 'Autonomy of Economic Relations', 6 November 1948, Oslo, Universitets Økonomiske Institutt. Reprinted in Hendry D. F. and Morgan M. S. (1995), *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*, 424–438. https://www.jstor.org/stable/1912791

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics 48*, 213–228. https://doi.org/10.1111/j.1468-0084.1986.mp48003002.x

Hendry, D. F. (1985). Monetary economic myth and econometric reality. *Oxford Review of Economic Policy 1*, 72–84. https://doi.org/10.1093/oxrep/1.1.72

Hendry, D. F. (1989). *PC-GIVE: An Interactive Econometric Modelling System*. Oxford: Oxford Institute of Economics and Statistics.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F. and N. R. Ericsson (1991). Modeling the demand for narrow money in the United Kingdom and the United States. *European Economic Review 35*, 833–886. https://doi.org/10.1016/0014-2921(91)90039-L

Hendry, D. F. and G. E. Mizon (1993). Evaluating dynamic econometric models by encompassing the VAR. In P. C. B. Phillips (Ed.), *Models, Methods and Applications of Econometrics*, pp. 272–300. Oxford: Basil Blackwell.

Hendry, D. F., A. J. Neale, and N. R. Ericsson (1991). *PC-NAIVE, An Interactive Program for Monte Carlo Experimentation in Econometrics. Version 6.0*. Oxford: Institute of Economics and Statistics, University of Oxford.

Hendry, D. F., A. J. Neale, and F. Srba (1988). Econometric analysis of small linear systems using Pc-Fiml. *Journal of Econometrics 38*, 203–226. https://doi.org/10.1016/0304-4076(88)90033-4

Hendry, D. F. and J.-F. Richard (1989). Recent developments in the theory of encompassing. In B. Cornet and H. Tulkens (Eds.), *Contributions to Operations Research and Economics. The XXth Anniversary of CORE*, pp. 393–440. Cambridge, MA: MIT Press.

Hendry, D. F. and P. K. Trivedi (1972). Maximum likelihood estimation of difference equations with moving-average errors: A simulation study. *Review of Economic Studies 32*, 117–145. https://doi.org/10.2307/2296866

Hylleberg, S. and G. E. Mizon (1989). Cointegration and error correction mechanisms. *Economic Journal 99*, 113–125. https://doi.org/10.2307/2234074.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12*, 231–254. https://doi.org/10.1016/0165-1889(88)90041-3

Johansen, S. and K. Juselius (1990). Maximum likelihood estimation and inference on cointegration – With application to the demand for money. *Oxford Bulletin of Economics and Statistics 52*, 169–210. https://doi.org/10.1111/j.1468-0084.1990.mp52002003.x

Lu, M. and G. E. Mizon (1991). Forecast encompassing and model evaluation. In P. Hackl and A. H. Westlund (Eds.), *Economic Structural Change, Analysis and Forecasting*, pp. 123–138. Berlin: Springer-Verlag.

Lucas, R. E. (1976). Econometric policy evaluation: A critique. In K. Brunner and A. Meltzer (Eds.), *The Phillips Curve and Labor Markets*, Volume 1 of *Carnegie-Rochester Conferences on Public Policy*, pp. 19–46. Amsterdam: North-Holland Publishing Company.

Mizon, G. E. (1984). The encompassing approach in econometrics. In D. F. Hendry and K. F. Wallis (Eds.), *Econometrics and Quantitative Economics*, pp. 135–172. Oxford: Basil Blackwell.

Mizon, G. E. and J.-F. Richard (1986). The encompassing principle and its application to non-nested hypothesis tests. *Econometrica 54*, 657–678. https://doi.org/10.2307/1911313

Park, J. Y. and P. C. B. Phillips (1988). Statistical inference in regressions with integrated processes. Part 1. *Econometric Theory 4*, 468–497. https//doi.org/10.1017/S0266466600013402

Park, J. Y. and P. C. B. Phillips (1989). Statistical inference in regressions with integrated processes. Part 2. *Econometric Theory 5*, 95–131. https//doi.org/10.1017/S0266466600012287

Phillips, P. C. B. and S. N. Durlauf (1986). Multiple time series regression with integrated processes. *Review of Economic Studies 53*, 473–495. https://doi.org/10.2307/2297602

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica 48*, 1–48. Reprinted in Granger, C. W. J. (ed.) (1990), *Modelling Economic Series*. Oxford: Clarendon Press. https://www.jstor.org/stable/1912017

Sims, C. A., J. H. Stock, and M. W. Watson (1990). Inference in linear time series models with some unit roots. *Econometrica 58*, 113–144. https://www.jstor.org/stable/2938337

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. *NBER Macro-Economic Annual*, 351–409.

Added references

Castle, J. L., M. P. Clements, and D. F. Hendry (2015). Robust approaches to forecasting. *International Journal of Forecasting 31*, 99–112. https://doi.org/10.1016/j.ijforecast.2014.11.002

Castle, J. L., J. A. Doornik, and D. F. Hendry (2021). Selecting a model for forecasting. *Econometrics 9(3)*, https://doi.org/10.3390/econometrics9030026.

Castle, J. L., J. A. Doornik, and D. F. Hendry (2025). A novel approach to forecasting uk inflation over 2021–2024. Economics discussion paper, 2025-w01, Nuffield College, University of Oxford. https://www.nuffield.ox.ac.uk/economics/Papers/2025/2025-W01ForcUKInflationAfterShifts24.pdf.

Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics 3(2)*, 240–264. https://doi.org/10.3390/econometrics3020240

Castle, J. L., N. W. P. Fawcett, and D. F. Hendry (2010). Forecasting with equilibrium-correction models during structural breaks. *Journal of Econometrics 158*, 25–36. https://doi.org/10.1016/j.jeconom.2010.03.004

Castle, J. L., D. F. Hendry, and O. I. Kitov (2018). Forecasting and nowcasting macroeconomic variables: A methodological overview. In EuroStat (Ed.), *Handbook on Rapid Estimates*, pp. 53–107. Brussels: UN/EuroStat.

Clements, M. P. and D. F. Hendry (1993). On the limitations of comparing mean squared forecast errors (with discussion). *Journal of Forecasting 12*, 617–637. https://doi.org/10.1002/for.3980120802

Clements, M. P. and D. F. Hendry (1996). Intercept corrections and structural change. *Journal of Applied Econometrics 11*, 475–494. https://doi.org/10.1002/(SICI)1099-1255(199609)11:5<475::AID-JAE409>3.0.CO;2-9

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

Doornik, J. A. (2018). *OxMetrics: An Interface to Empirical Modelling* (8th ed.). London: Timberlake Consultants Press.

Doornik, J. A. and D. F. Hendry (2018). *Empirical Econometric Modelling using PcGive: Volume I.* (8th ed.). London: Timberlake Consultants Press.

Duffy, J. A. and D. F. Hendry (2017). The impact of near-integrated measurement errors on modelling long-run macroeconomic time series. *Econometric Reviews 36*, 568–587. https://doi.org/10.1080/07474938.2017.1307177

Emerson, R. A. and D. F. Hendry (1996). An evaluation of forecasting using leading indicators. *Journal of Forecasting 15*, 271–91. https://doi.org/10.1002/(SICI)1099-131X(199607)15:4<271::AID-FOR623>3.0.CO;2-7

Engle, R. F. and D. F. Hendry (1993). Testing super exogeneity and invariance in regression models. *Journal of Econometrics 56*, 119–139. https://doi.org/10.1016/0304-4076(93)90103-C

Ericsson, N. R. (2012). Detecting crises, jumps, and changes in regime. Working paper, Federal Reserve Board of Governors, Washington, D.C. https://eesp.fgv.br/sites/eesp.fgv.br/files/file/Neil_Ericsson.pdf

Hendry, D. F. (1999). An econometric analysis of US food expenditure, 1931–1989. In J. R. Magnus and M. S. Morgan (Eds.), *Methodology and Tacit Knowledge: Two Experiments in Econometrics*, pp. 341–361. Chichester: John Wiley and Sons.

Hendry, D. F. (2006). Robustifying forecasts from equilibrium-correction models. *Journal of Econometrics 135*, 399–426. https://doi.org/10.1016/j.jeconom.2005.07.029

Hendry, D. F. and M. P. Clements (1994). On a theory of intercept corrections in macro-economic forecasting. In S. Holly (Ed.), *Money, Inflation and Employment: Essays in Honour of James Ball*, pp. 160–182. Aldershot: Edward Elgar.

Hendry, D. F., S. Johansen, and C. Santos (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics 33*, 317–335. Erratum, 337–339. https://doi.org/10.1007/s00180-007-0054-z

Hendry, D. F. and A. B. Martinez (2017). Evaluating multi-step system forecasts with relatively few forecast-error observations. *International Journal of Forecasting 33*, 359–372. http://dx.doi.org/10.1016/j.ijforecast.2016.08.007

Hendry, D. F. and M. Massmann (2007). Co-breaking: Recent advances and a synopsis of the literature. *Journal of Business and Economic Statistics 25*, 33–51. https://doi.org/10.1198/073500106000000422

Hendry, D. F. and G. E. Mizon (2012). Open-model forecast-error taxonomies. In X. Chen and N. R. Swanson (Eds.), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis*, pp. 219–240. New York: Springer.

Hendry, D. F. and G. E. Mizon (2014). Unpredictability in economic analysis, econometric modeling and forecasting. *Journal of Econometrics 182*, 186–195. https://doi.org/10.1016/j.jeconom.2014.04.017

Hooker, R. H. (1901). Correlation of the marriage rate with trade. *Journal of the Royal Statistical Society 64*, 485–492. Reprinted in Hendry, D. F. and Morgan, M. S. (1995), *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press. https://www.jstor.org/stable/2979868

Johansen, S. and B. Nielsen (2009). An analysis of the indicator saturation estimator as a robust regression estimator. In J. L. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics*, pp. 1–36. Oxford: Oxford University Press.

Pretis, F., L. Schneider, J. E. Smerdon, and D. F. Hendry (2016). Detecting volcanic eruptions in temperature reconstructions by designed break-indicator saturation. *Journal of Economic Surveys 30*, 403–429. https://doi.org/10.1111/joes.12148

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time series (with discussion). *Journal of the Royal Statistical Society 89*, 1–64. Reprinted in Hendry, D. F. and Morgan, M. S. (1995), *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press. https://www.jstor.org/stable/2341482