Macroeconomic Forecasting: A Survey

Kenneth F. Wallis

MACROECONOMIC FORECASTING: A SURVEY*

Kenneth F. Wallis

Developments in macroeconomic forecasting over the last twenty years are surveyed in this paper, which takes the 1969 Presidential Address to the Royal Economic Society (Cairncross, 1969) as its starting point. Sir Alec Cairncross had found no previous occasion on which the Royal Economic Society had discussed 'this new activity', and so selected economic forecasting as the topic for his address. As retiring Chief Economic Adviser to the Treasury he had been preoccupied with forecasting for the preceding few years, and was well placed to reflect on the new kind of economic forecasting that was emerging and the new ways in which it was being organised. At the time these were changing fast, with forecasting becoming in particular more heavily based on computable models. A conference in April 1969 heard that in the Treasury 'a more elaborate fully formalised model is being programmed for a computer' (Roy, 1970); in August 1969, at the National Institute of Economic and Social Research (NIESR), an econometric model of the whole economy became an integral part of the quarterly forecasting exercises that had begun, with the publication of the *National Institute Economic Review*, in 1959. These developments followed the inauguration by the London Business School (LBS) in 1966 of the first series of published forecasts based on the direct application of a complete statistical model of the economy.

In the late 1960s and early 1970s confidence in forecasting was growing. 'Extra resources were put in, and there were hopes that the accumulation of data and more sophisticated techniques would lead to major improvements in accuracy of forecasts and understanding of the economy' (Burns, 1986). This confidence rested in part on the wide acceptance of the neoclassical synthesis as a framework for macroeconomic analysis. The phrase originated with Samuelson, who was largely instrumental in constructing and promulgating the 'grand neoclassical synthesis', which was given considerable prominence in the third edition of his textbook. Here it was noted that economists, instead of being Keynesian or anti-Keynesian, 'have worked toward a synthesis of whatever is valuable in older economics and in modern theories of income determination. The result might be called neo-classical economics and is accepted in its broad outlines by all but about 5 per cent of extreme left wing and right wing writers' (Samuelson, 1955, p. 212). In his *New Palgrave* entry on the neoclassical synthesis, Blanchard (1987) observes that it 'did not expect full employment to occur under *laissez-faire*; it believed however that, by proper use of monetary and fiscal policy, the old classical truths would come back into

relevance. This synthesis was to remain the dominant paradigm for another twenty years, in which most of the important contributions, by Hicks, Modigliani, Solow, Tobin and others, were to fit quite naturally.[1] The neoclassical synthesis enjoyed considerable empirical success in the 1960s: as a result Heller *et al.* (1968, p. 16) could conclude that 'governments have, to a large extent, succeeded in subduing or overcoming the rhythmic fluctuations which used to be called the trade cycle' and Cairncross (1969, p. 805) could 'doubt whether bad theory has played a major part in forecasting errors in this country over the past decade or two'.

Growing confidence in the use of more sophisticated techniques rested on developments in econometric modelling and forecasting that had taken place largely outside the United Kingdom. The construction of economy-wide models had been pioneered by Tinbergen, who had subsequently served as the first director of the Central Planning Bureau of the Netherlands; this had become the official forecasting agency whose activities were the most heavily model-based. An account of the methods of short-term forecasting used by the governments of six member countries was published by the OECD in 1965, and in his introductory survey McMahon reported that 'All the participating countries except the United Kingdom use an econometric model; but except for the Netherlands it is used primarily as a consistency check, rather than as the primary method of making the forecast itself.' Theil's classic *Economic Forecasts and Policy* (1958) and *Applied Economic Forecasting* (1966) represented in part a contribution to the research programme of the Central Planning Bureau. In the United States Klein had led the way, and his brief stay in the United Kingdom had resulted in the first quarterly model of any national economy (Klein *et al.*, 1961). Descendants of the Klein–Goldberger model, built there in the early 1950s, were used for forecasting by the Research Seminar in Quantitative Economics at the University of Michigan (Suits, 1962), and forecasts based on a quarterly model of the US economy were distributed by the Wharton School at the University of Pennsylvania from 1963 onwards (see Klein, 1968 and references therein). Important work on the evaluation of forecasts came from an NBER project on short-term economic forecasting (Zarnowitz, 1967; Mincer and Zarnowitz, 1969). This literature provided a foundation for the initial developments indicated in the opening paragraph, and for subsequent developments that occurred in the period under consideration.

A feature of the last twenty years that is immediately apparent on attempting a survey is that the literature on forecasting has grown apace with forecasting activity. From the three groups in existence in the late 1960s, macroeconomic forecasting activity has grown to the point where the Treasury's monthly compilation *Forecasts for the UK Economy* now covers nineteen forecasts. As for the literature, it remains the case that forecasting has scarcely featured in this JOURNAL, no major article having dealt with the subject in the intervening period, but in this time two journals devoted entirely to forecasting have been established. Notable books on the production and use of macroeconomic forecasts have been published by leading practitioners, for example, Keating

(1985), Klein and Young (1980), and Llewellyn *et al.* (1985), respectively from UK, US and international perspectives. More general textbooks have appeared, for example, Fair (1984), Granger and Newbold (1977), and Pindyck and Rubinfeld (1976). A substantial section of the literature deals with the assessment of forecasts, both in theory and in practice, and in turn with the use of forecasts in model evaluation. An early study of UK forecasts is that by Ash and Smyth (1973); subsequently numerous authors, including the forecasters themselves, have contributed assessments. Various aspects of macroeconomic forecasting are touched on in the international conference volumes edited by Chow and Corsi (1982), Kmenta and Ramsey (1981), and Malgrange and Muet (1984); the UK scene is represented by Hilton and Heathfield (1970), Renton (1975), Ormerod (1979a), and the sequence of reviews by the ESRC Macroeconomic Modelling Bureau (Wallis *et al.*, 1984–7).

The paper proceeds as follows. In the next section the essential ingredients of a forecast are described, by way of background, and some basic data on forecasts and outcomes are presented as a point of reference. The events and consequences of some important episodes are then considered: first 1974–5, when in many countries unprecedented forecast errors were made, which provided a focus for important challenges, both empirical and theoretical, to the prevailing consensus; next 1979–81, which saw a recession, a change of regime and notable forecasting errors, heralding further developments. Forecast evaluation is touched on as much as is necessary throughout this account, but is then treated more thoroughly in a separate section, which includes consideration of a final episode, namely 1986, when the dramatic fall in the world price of oil, fortunately for the forecasters, was not anticipated. Concluding comments follow, the paper closing, like Cairncross (1969), on the issue of the publication of forecasts. The principal orientation of the paper is towards the UK economy, with occasional glances elsewhere, although developments in economic analysis and econometric methods know no frontiers.

I. CONSTRUCTION OF FORECASTS

The need for forecasting in economic policy-making, and the essential ingredients of a forecast, have scarcely changed over twenty years. Thus Cairncross (1969, p. 798) notes that forecasts 'provide a frame of reference for policy decisions' and 'a base against which to judge how policies are working out', and the current Chief Economic Adviser to the Treasury similarly states 'Decisions have to be made against the background of an uncertain current position; they cannot easily be reversed, and many policy changes have consequences stretching years ahead. Some kind of forward look is therefore essential and it is best to do this in a consistent way. Once they are produced, forecasts have an important monitoring function as they provide a basis against which to judge subsequent developments' (Burns, 1986, p. 117). Outside the policy-making context, two further quite different motives for forecasting exist. One is to anticipate events, whether for private gain or for public good, and it

is largely in respect of the former that the growth in forecasting activity has occurred; the other is to put hypotheses about the behaviour of the world to test, which helps explain why independent research groups engage in forecasting.

Although the key ingredients of a forecast have remained quantitative data and a framework for their interpretation and analysis, substantial developments have occurred in respect of the methods of analysis. Early forecasts were based on a limited number of variables, which were analysed in the context of an implicit, perhaps informal model, not necessarily written down. The process relied on the assessment of data and the evaluation of new information by the experienced forecaster. 'A few people possess the extraordinary gift of being able to do this in their heads... The process involved is not well understood, but it seems to involve the checking of data against the predictions and workings of a mental model of how the economy works. The disadvantage of such a method is that the number of people truly gifted in this way is small, and the technique is difficult to explain and transmit to others' (Llewellyn *et al.*, 1985, p. 83) – and to replicate. In the 1960s the use of explicit, more formal models increased, with these models becoming increasingly based on estimated equations. The models distinguish between endogenous and exogenous variables, that is, those determined by the system of equations and those treated as being determined outside the system. In a forecasting context, these latter variables have to be set by projection or assumption, which leads to the further distinction between an unconditional and a conditional forecast. The former represents the conventional understanding of a forecast, namely a prediction of a future event, whereas the latter represents an if–then statement, resting on the occurrence of certain specified conditions. The pure prediction problem is the main concern of this paper, although the distinction is less clear-cut than might appear, since exogenous variables that are policy instruments may be treated in different ways. One possibility is to assume no change in policy instruments, irrespective of the forecast; another is to project a policy response to the forecasted developments.

With greater computerisation the size of the models could be increased, allowing attention to be paid to a greater number of variables, including the components of aggregates previously treated as one. The intervening years have seen such developments in computing that computer capability is no longer a constraint, neither on the size and complexity of the model to be managed, nor on the frequency with which forecasts and policy exercises can be constructed and revised, nor on the econometric estimation and testing procedures applied to the models, these procedures themselves having been substantially developed over the last twenty years.

The removal of the computing constraint and the use of more fully elaborated accounting frameworks should not obscure the fact that some practical limitations, in particular concerning data, remain much as in Cairncross' (1969) description. 'First of all, the forecaster has to contend with the inevitable lag of statistics behind events... when the February forecasts are being prepared in advance of the Budget the latest available national accounts

relate to the third quarter' of the preceding year. This remains the case today. Moreover, the inevitable lags differ from one variable to another, so there is not a clean break between periods for which data are available and those for which they are not; rather, the dataset has a 'ragged edge', presenting additional complications (Wallis, 1986). 'Next comes the issue of reliability', one indication of which is the frequency with which preliminary data are subsequently amended, and another is the frequency of changes in definition of a given series. Some errors are systematic and survive the revision process, the most notable current example being in international trade, where on a balance of payments basis imports considerably exceed exports, globally, by an amount which, as a proportion of total imports, is of the same order of magnitude as the UK current balance. Less frequently, reconstruction of the real national accounts data using revised relative price weights can lead to very substantial revision of data, going back perhaps to the origin of the data series, and affecting both the level of the series and its growth rate over given periods. Last, but not least, is the discrepancy between the statistician's measurement and the economist's concept. The problem of the valuation of stocks, and the decomposition of a given change into price and volume changes, arises in a number of areas, and is part of the well-known problem of the measurement of capital stock; a related problem is that of the measurement of capacity utilisation. In the absence of adequate direct measurements, indirect measurements or 'proxy' variables have to be used. Moreover 'even measures that in principle are equal to one another are often in flat contradiction. To take what may be an extreme example, between the first quarter of 1966 and the first quarter of 1967 the expenditure measure of GNP increased by 2 per cent, the income measure fell by $\frac{1}{2}$ per cent and the output measure was unchanged. Yet all these measure the same thing' (Cairncross, 1969, p. 803). Such contradictions remain in the 1980s. For example, in October 1983 it was estimated that the expenditure measure had risen by $1\frac{1}{4}\%$ between 1980 and 1982, while the output measure had fallen by $\frac{1}{2}\%$ and the income measure had not changed. Just one year later, in October 1984, the expenditure increase was revised to $\frac{1}{2}\%$, that of output to $\frac{1}{4}\%$ and income was estimated to have increased by $1\frac{3}{4}\%$!

The continuing presence of data discrepancies and delays is one reason why there remains a role for informed judgement in forecasting, often expressed through a process of adjustments to model-based forecasts. The use of such adjustments, variously termed constant adjustments, residual adjustments, add factors, etc. is widespread, despite the models having come to represent a more complete framework for analysis. Indeed, the more complete is the model, the greater is the internal consistency with which a given adjustment to a variable or an equation is carried through the forecast calculations. The use of 'conjunctural analysis' to supplement the official macroeconomic data and so to suggest adjustments is well described by Keating (1985); other reasons for making adjustments, based on knowledge of further developments not incorporated in the model or of other model deficiencies, feature in general terms in all practitioners' accounts of their art. An attempt to appraise

the impact of these adjustments, *ex post*, is presented in Section IV.5 below.

As a point of reference for subsequent discussion, data on one-year-ahead forecasts of GDP and inflation published by independent UK forecasting groups, LBS, NIESR and LPL (the Liverpool University Research Group in Macroeconomics) are presented in Figs. 1 and 2. The upper panel of each figure shows the actual outcome, and the lower panel the forecast errors, defined as actual minus forecast. Public discussion usually focuses on the annual (real) GDP growth rate and the annual rate of inflation, so the data are presented in these terms, which also has the merit of reducing the basic GDP and price level variables to (near) stationarity. That is, the proportionate rates of change are trend-free, and their variances and autocovariances can be treated as time-invariant. The forecasts described as one-year-ahead are forecasts of a given year's growth and inflation published early in that year, and
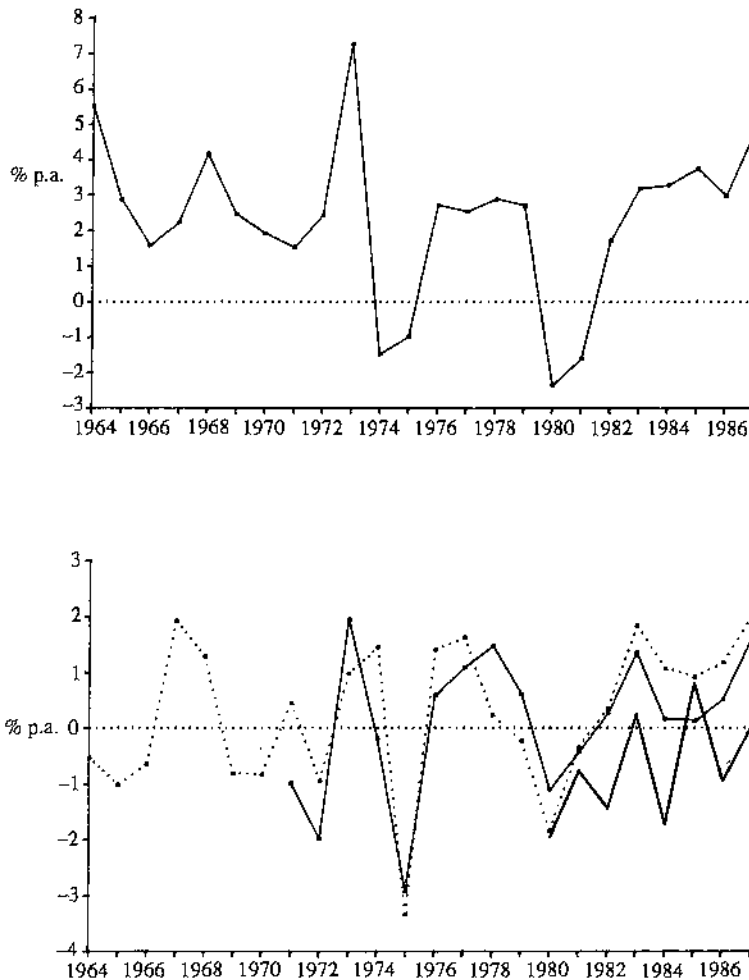
Fig. 1. GDP growth. Upper panel, actual; lower panel, one-year-ahead forecast errors. Key: LBS (—■—), NIESR (··■··) and LPL (——).
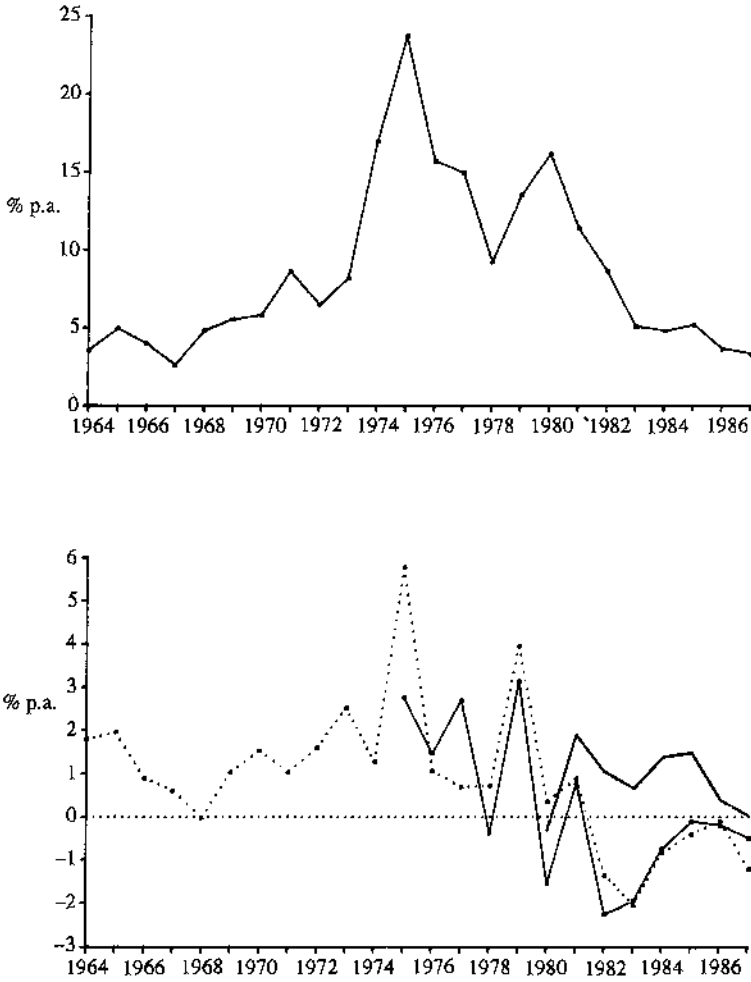
Fig. 2. Inflation. Upper panel, actual; lower panel, one-year-ahead forecast errors.
Key: ——■——, LBS; ··■··, NIESR; ——, LPL.

hence which rest on a formal database ending sometime in the previous year, although other, more timely, conjunctural information may also be used, as noted above.

II. THE 1974–5 RECESSION

It is clear from Figs. 1 and 2 that the largest absolute forecast errors in the period under consideration occurred in 1975. This marked the low point of a period of gradual deterioration in forecast performance. For example Osborn (1979) finds that, compared to its accuracy in forecasting real GDP over the 1965–70 period, the NIESR performance deteriorates in the 1970–5 period, most markedly so in 1974–5. In the United States, McNees (1979a) finds that, for a sequence of quarterly one-year-ahead forecasts of real GNP growth through the 1970s, all of the 'failures', with one minor exception, occurred in

forecasts of the periods ending between 1974:1 and 1976:1, all but one of these nine forecasts being a 'failure': a 'failure' is defined as the median absolute error across five leading US forecasters substantially exceeding its average value over the entire period. For the OECD area, Llewellyn *et al.* (1985, ch. 6) find that, of the seventeen year-ahead forecasts of real OECD GNP growth over the period 1966–82, those for 1974 and 1975 exhibit the two largest errors. In the analytical framework of the time, an overestimate of real growth would be expected to be associated with an underestimate of unemployment, which indeed occurred: McNees (1979*a*) notes the 'extraordinarily large underestimates of unemployment in early 1975'. Emphasis on the absolute forecast errors may underestimate the ability of the forecasters, since, as indicated by Figs. 1 and 2, for example, the situation they were attempting to predict showed a sharp increase in variation. In these circumstances their relative performance, assessed more fully below, might even be considered laudable. It is clear from contemporary writings, however, that this provided little solace, and the absolute errors prompted a vigorous search for forecasting improvements.

The economic scenario in which these forecasting errors occurred was the coexistence, not previously experienced, of rapid inflation and high unemployment, leading into recession. Writing from a US perspective, Lucas and Sargent (1978) in their well-known polemic laid the blame for these reversals very clearly at the door of the neoclassical synthesis. That macroeconomic policies predicated on models in this tradition had not produced the predicted results represented 'econometric failure on a grand scale'. Moreover, not only were the predictions 'wildly incorrect' but 'the doctrine on which they were based is fundamentally flawed'; it was argued 'that the difficulties are *fatal*: that modern macroeconomic models are of *no* value in guiding policy and that this condition will not be remedied by modifications along any line which is currently being pursued'. The models should be replaced by equilibrium models (Lucas, 1975, 1977), which assume that prices and quantities continuously clear markets and that agents continuously optimise, which in turn leads to the imposition of the hypothesis of rational expectations. In such equilibrium or 'new classical' models, fluctuations are caused by agents' reactions to unanticipated shocks.

This recommendation was not followed by the forecasters, whose attention focused on other perceived inadequacies of their models. First among these was the models' oversimplified treatment of supply factors, since the 'first oil price shock', that is, the quadrupling of the price of imported oil in late 1973, was in particular blamed for the 1974–5 recession. 'This unique external supply-side shock provided a severe challenge to the conventional, demand-oriented forecasting techniques' (McNees, 1979*b*) not because it was unanticipated but because its repercussions were not understood and in consequence were underestimated. Thus, in revising their forecasts for 1974 after observing the oil price increase, US forecasters made adjustments to show more inflation and lower growth, but these, while in the right direction, were not nearly large enough, as McNees (1979*b*) notes: 'the upward pressure on prices and

downward pressure on output were far stronger than...anticipated'. Writing at the time, Klein (1974) noted that 'The US models, which have been not only demand oriented, but also overly domestic in character have few provisions for indicating how high import prices (fuel in this case) contribute to domestic inflation. This is not a problem for European and UK model builders, however, because they have generally been alert to problems of imported inflation and allow for that factor in their domestic price formation equations'. Subsequently elaborate energy sectors were introduced into the Wharton model, to allow routine analysis of oil-price changes in particular. Likewise the DRI model moved into its third generation after 1975, to remedy the shortcomings in the existing ('second generation') models that became apparent during 1973–5. Specifically, the design of these models did not 'offer sufficient points of contact with external matters such as raw material prices, oil prices, worldwide booms and recessions, shortages and the financial instability which only became more evident during that period' (Eckstein, 1983, p. 8). They had been fitted over the sample period 1953–73, which enjoyed relatively smooth growth and 'did not reveal the full cyclical vulnerability of the economy'. Appropriate extensions to the model were accordingly made.

For the OECD area as a whole, forecast errors also reflected underestimates of the scale of the response to the oil-price shock. Again, once the oil price increase was known, growth forecasts for 1974 were revised downwards, but insufficiently so. 'Anyone who was involved in managing a macroeconometric model at the time of the 1973 oil-price rise will never forget the strain this supply shock placed on an apparatus largely attuned to coping with demand disturbances' (Higgins, 1988). Since the continuation of the recession into 1975 was not foreseen in forecasts prepared in late 1974, Llewellyn *et al.* (1985) emphasise the lack of understanding not only of the magnitude of the economic responses to the shock but also of their timing. An international analytic framework and an appropriate set of quantified relationships was not available at the time, and·the oil shock provided a strong impetus for the development of globally consistent projections, as Higgins notes. Whereas it was immediately apparent that there would be large transfers of financial resources from oil-importing to oil-exporting countries through the current account of the balance of payments, their continuing consequences were unclear. As the OPEC countries began to react to their new wealth, work began on an international linkage model to focus on the effects of the various transmission mechanisms.

In the United Kingdom the inflationary explosion of 1974–5 was attributed not only to the oil-price shock but also, by monetarists in particular, to the surge in monetary growth in 1972–3. A rapid rise in public borrowing was associated with the expansionary fiscal policy of 1971–2, which produced unprecedented output growth in 1973, and rapid growth in bank lending followed the move towards a more liberal regulatory regime in 1971 (the 'competition and credit control' measures). The broad money measure in use at the time (M3) grew by 27·8% in 1972 and 27·6% in 1973, the inflationary explosion following with a two-year lag and an intervening temporary output

rise, exactly as in the basic Friedmanite analysis, although the oil-price rise exaggerated the magnitude of the inflationary response. Foreign prices had a role in the models, as indicated by Klein (1974), quoted above, but no distinction was drawn between oil and other imports, nor between oil and other commodities, whose prices were also rising exceptionally quickly in 1973. The inclusion of foreign prices was by no means enough to prevent the 1975 forecasting errors shown in Figs. 1 and 2; similarly large errors were observed in other variables such as consumption and employment. As elsewhere, the depth of the recession was not correctly anticipated. Again as elsewhere, the response was a pragmatic one, with forecasters attempting to learn from the new information and new experience in order to improve future performance, rather than jettisoning their whole approach.

In the light of the events described above, and of the 1973 move to a system of floating exchange rates, attention concentrated on the modelling of the financial sector, exchange rate equations, and the influence of monetary policy. A new financial sector for the Treasury model became operational in 1978 (Spencer and Mowl, 1978; Spencer, 1986), and a simple financial system was introduced into the NIESR model (Ormerod, 1979b). The fullest account of their learning and respecification process or, in Flemming's (1978) phrase, their 'intellectual odyssey', is provided by the LBS group in papers presented at conferences in late 1977 and mid-1978 (Ball and Burns, 1978; Ball et al., 1979). The LBS approach followed the 'international monetarist' tradition (Dornbusch, 1976), that changes in relative money supplies affect prices via changes in the exchange rate, given long-run purchasing power parity. More generally, the LBS group undertook a 'fundamental reappraisal of the properties of the system as a whole'. In the personal sector, for example, failure to anticipate the sharp fall in consumers' expenditure in 1975 led to a reconsideration of the impact of inflation on real spending, in which consumers were assumed to save more as inflation increased in order to preserve the real value of their financial assets.

Given the central place of the consumption function in neoclassical synthesis econometrics, the rise in the savings ratio that occurred in the early to mid-1970s in association with a rising rate of inflation prompted much empirical research by forecasters and others, which continues to this day. Notable examples in this JOURNAL are the papers by Davidson et al. (1978) and Pesaran and Evans (1984), and the current forecasting models incorporate the effect of inflation on consumers' expenditure, occurring either directly, or indirectly via an adjustment to real income for the inflation loss on liquid assets and/or wealth. The detection of this effect provides a good example of the model development opportunities offered by turbulent periods of history, alluded to in the quotation by Eckstein, above.

The technical point is that forecasts based on a mis-specified model, for example, one that omits a relevant explanatory variable, will continue to perform as well as expected from past experience and hence arouse no suspicions as long as the omitted variable continues to behave in the same manner as before. Once its behaviour changes, however, then so does that of

variables related to it, and forecasts of these variables based on the mis-specified model exhibit unexpected errors, so drawing attention to the model's inadequacy. Over the time span of this paper the changes have been from relatively smooth to relatively turbulent behaviour, and it may be that only after the perturbation of the data, and a period in which events appear inexplicable, is it possible to identify and estimate an appropriately extended model. On the other hand, there may have been enough information in the 'smooth' data to permit the satisfactory estimation of a revised model, once the inadequacies of the original model are revealed. In the case of the consumption function, Davidson *et al.* (1978) find that there is enough variation in the 1960s data to allow the estimation of a new equation, incorporating inflation effects, that does not suffer from the mid-1970s predictive failure of the old equation. The forecast errors in consumers' expenditure could have been avoided if the inflation effects had been looked for and, in an *ex ante* forecast context, if inflation itself had been well predicted.

In the case of the exchange rate, other UK forecasters, like the LBS group, introduced exchange rate equations and made other related changes post-1975. None of the exchange rate systems performed well when used for forecasting, however, and the exchange rate equations often had to be overridden or subject to heavy residual adjustments. In the first instance the modelling difficulties were not surprising, given the short series of data available for the 'floating' regime and the completely uninformative nature of the 'fixed' regime. These have persisted, however, Isard (1988) concluding in his international survey that 'empirical modeling of exchange rates over the past decade has been largely a failure', which is evident 'from documentation of the poor post sample forecasting accuracy of the models, from data that appear to reject important building blocks for the monetary models (in particular, the assumption of uncovered interest rate parity), and from the lack of statistically significant in-sample support for existing portfolio-balance models of the exchange risk premium'. Forecasting the exchange rate remains arguably the greatest single problem facing UK forecasters.

The experience of the 1974–5 recession was a major blow to the growing confidence in forecasting of the 1960s and early 1970s. It represented a valuable learning experience, however, and caused a sharp spurt in the continuing evolution of forecasting models, as existing systems were amended, not abandoned. In any event the advice that they should be abandoned was by the end of the decade itself amended, Lucas (1980) observing 'To what extent this forecast error should be interpreted as a "fatal" error in models based on the neoclassical synthesis or simply as one suggesting some modifications is not so easy to determine.'

III. THE 1979–81 RECESSION

The second oil-price shock occurred in several stages in 1979 and initiated a second recession in the OECD economies. 'Inflation re-accelerated, current account deficits increased, public sector deficits swelled, and unemployment rose yet further. Taken together, these presented a greater problem for policy

in nearly all economies than in any previous post-war cycle' (Llewellyn *et al.*, 1985, p. 39), although unlike the first oil price shock, there was now no substantial disagreement among the OECD economies about the appropriate policy response. The UK experience differed from that of the rest of the OECD, however. Despite its near self-sufficiency in oil by 1979, the United Kingdom suffered a recession which started earlier than elsewhere and was of much greater severity. Many authors have compared this recession with the Great Depression of the 1930s, output falling by 5 % in both 1929–31 and 1979–81. Unemployment rose from 1·2 million in mid-1979 to reach 2 million in late 1980 and 3 million by autumn 1982.

May 1979 had seen the election of a new Conservative government in the United Kingdom, committed to the reduction of inflation and the creation of conditions favourable to sustained economic growth, but no longer accepting responsibility for high employment. Inflation was to be controlled by restrictive monetary policy, not incomes policy, and the supply side was to be strengthened through fiscal measures such as a shift from direct to indirect taxation and a reduction in public expenditure, and through the liberalising of financial markets and the labour market. Thus in the June 1979 Budget monetary policy was tightened, public expenditure cuts were announced, income tax was reduced and the rate of indirect taxation (value added tax) was increased. In October 1979 exchange controls were abolished. The March 1980 Budget introduced the medium-term financial strategy (MTFS) comprising a four-year declining target path for the growth of a broad monetary aggregate (sterling M3) and an accompanying path for the public sector borrowing requirement (PSBR), which implied a declining PSBR/GDP ratio. The immediate result of these policies, in conjunction with external developments, was a major loss of competitiveness accompanied by a large fall in production and the rise in unemployment noted above. Company finances were adversely affected, real profitability falling to its lowest recorded level in the second half of 1980, and the reaction was a massive reduction in inventories. Arithmetically, the reduction in real GDP between 1979 and 1980 was more than accounted for by the reversal in stockbuilding; substantial destocking continued in 1981. Inflation increased to 18 % in 1980, but then began to fall, reaching 5·4 % by the end of 1982.

This brief description serves to set the scene for the discussion of forecast performance over this period, of which a good account is given by Barker (1985), and of subsequent developments. The 'Thatcher experiment', or what is now the 'Thatcher experience', has a considerable literature of its own, two notable early contributions being the Brookings papers of Buiter and Miller (1981, 1983), and that literature falls outside the scope of the present survey.

The recession was not well forecast, being unanticipated until mid-1979, although the errors were smaller than in 1974–5. As Barker (1985) points out, the difficulties are demonstrated by an exercise presented to the Bank of England Academic Panel (Worswick and Budd, 1981), in which the LBS and NIESR models were used to explain the shortfall of output over 1978–80

relative to its trend. (Budd's contribution also includes a post-mortem analysis of LBS forecasts over this period; Surrey (1982) likewise examines the NIESR track record.) The shortfall in output was estimated at 4%, the main contributions being

(i) policy changes in the June 1979 Budget, in particular the VAT increase and the planned cuts in public expenditure (1·6–1·7%);

(ii) the deterioration in competitiveness (1·1–1·3%);

(iii) the oil-price rise (0·8–1·3%).

As all these factors appeared in the models as relevant exogenous variables, three possible explanations of the forecast errors remain. First, the movements in these variables may not have been correctly anticipated. Secondly, their effects on the economy may not have been accurately modelled. Thirdly, the model forecasts may have been adjusted to show less dramatic changes, although no evidence on this is available (an analysis of the impact of adjustments on later forecasts is presented in Section IV.5 below). After the onset of the oil-price increases and the June 1979 Budget growth forecasts were revised downwards, with the LBS correctly forecasting that output would fall in 1980, although the extent of the fall, and in particular the collapse of GDP in the fourth quarter of 1980, was not anticipated. The NIESR forecast of growth in 1980 was revised to a negative number only in February 1980, and again the fall in output was considerably underestimated. By this time the full extent of the oil-price rise was appreciated, but the extent of the financial squeeze on the company sector was not. Cuts in public expenditure continued to be forecast into 1980, but these never materialised, and these errors served to offset the failure to capture the large volume of destocking. As Budd notes, to predict the speed of adjustment to major shocks, particularly financial shocks, is especially difficult, and the previous recession offered little guidance. Whereas companies allowed exceptionally high deficits to accrue in 1974 and did not adjust inventories and employment until 1975, in 1980 the adjustment occurred almost instantaneously. Thus the rise in unemployment was also underpredicted. Finally the exchange rate was underpredicted, and this would have remained the case even had the oil-price increase been fully anticipated.

Once again the experience of forecasting through a relatively turbulent period led to a reappraisal of the forecasting models. While this resulted in further small steps in the continuing process of model evolution, such as the introduction of financial considerations into the determination of investment and stockbuilding, two major items appeared on the agenda, namely expectations and the supply side.

Expectations of future developments have long been recognised as important determinants of current behaviour and, as Klein (1987, p. 420) notes, forecasters have long endeavoured to use survey data on expectations and anticipations wherever possible. Reliable quantitative data on expectations are relatively rare, however, and auxiliary hypotheses about the way in which expectations are formed are commonly used in their place. In the forecasting models of the 1960s and 1970s expectations were assumed to be formed by extrapolating from past experience, usually in respect of the variable of interest

alone. The simplest example of this approach is the adaptive expectations hypothesis, which was widely used, and in this and more general cases the resulting model can be described as backward-looking. The need to accommodate explicit forward-looking behaviour was increasingly recognised, at the theoretical level through the attention given to the rational expectations hypothesis and at the practical level through the experience of 1979–81. The role of expectations in the face of a change of policy regime and the associated questions of credibility began to feature in the discussion. Credibility featured without agreement, however. On the one hand Buiter and Miller (1981, p. 362) assert

> The government established the credibility of its restrictive policy stance at the start of its term of office. The perception that current and future monetary policy would be restrictive was reflected promptly in the exchange rate, interest rates, and financial markets generally, but only gradually in domestic costs, especially wages. This led to a major appreciation of the real exchange rate along the lines of the overshooting model, a rise in real interest rates, and a decline in Tobin's q.

On the other hand, Matthews and Minford (1987, p. 62) argue that 'in fact, most people had written off the early actions of the government as unlikely to be followed through,...so if anything people expected a "U-turn" towards much looser policies. It was for this reason, the lack of credibility, that prices and wages in 1979–80 were accelerating towards 20% p.a. growth.'

In March 1980 the first forecasts based on the Liverpool model (Minford *et al.*, 1984) were published. This model represented a break with the existing models of the UK economy, being a new classical equilibrium model, incorporating the hypothesis of rational expectations; it is a monetarist model in the sense that higher monetary growth directly increases inflation, with no role for cost factors. The rational expectations literature is usually assumed to start with Muth (1961), although following much older discussions of the influence of forecasts on outcomes, Grunberg and Modigliani (1954) had already shown that where agents react to forecasts and thereby alter the course of events, this reaction can be taken into account to yield a correct, self-fulfilling forecast. This same kind of internal consistency is imposed by the rational expectations algorithms used to calculate a sequence of forecasts based on a model containing explicit forward-looking expectations variables, in that each period's future expectations coincide with the model's forecasts for the future period. The approach is more appropriately and perhaps less controversially termed 'model-consistent' expectations. In policy evaluation exercises, the use of consistent expectations allows policies to be tested under conditions in which their effects are understood. As Currie (1985) argues, good performance in these conditions is a necessary condition for a satisfactory policy: 'a policy that performs badly when its effects are understood must be unsatisfactory'. In both forecasting and policy analysis, the explicit treatment of forward expectations allows such issues as the announcement effects of future policy changes, their credibility and, indeed, the consequences of false expectations to be dealt with.

Fischer (1988) recalls that his 1976 survey with Barro drew 'a clear distinction between the rational expectations hypothesis as a theory of expectations, and the type of equilibrium model in which the hypothesis was typically embodied at the time'. This distinction was blurred by leading US forecasters, however, thus in discussing the 'rational expectations school', both Eckstein (1983) and Klein (1986) associate the rational expectations hypothesis with the policy ineffectiveness proposition, and Klein *et al.* (1983) describe the 'many ways in which the thinking of monetarists and proponents of rational expectations are congruent'. Given the opposition of the 'mainstream models' to the positions of the 'rational expectations school' and the extreme monetarists indicated in these writings, it is perhaps not surprising that consistent expectations were not widely embraced in the US models. In the United Kingdom, however, despite the first appearance of rational expectations being in the new classical Liverpool model, the distinction appears to have been more clearly appreciated, and the incorporation of explicit forward-looking expectations handled in a model-consistent manner was part of the post-1981 revision process of other, more mainstream 'sticky price' and quantity adjustment models.

In the case of the LBS model, the introduction of consistent expectations was associated with the introduction of a detailed financial sector, in which asset demands are determined in a general portfolio choice model featuring expected future prices of gilts, equities and overseas assets, and in which the exchange rate is determined as a market-clearing price. Various questions, now of increased importance, such as the conduct of monetary and exchange rate policy, the use of open-market operations and debt management, and the finance of fiscal deficits could then be addressed.

The introduction of forward-looking behaviour into the NIESR model (Hall and Henry, 1985) was motivated by its forecasting performance: backward-looking equations for employment, investment and stockbuilding, depending mainly on lagged output, missed the turning point in 1979, as noted above, and did not capture the speed and depth of the recession. Accordingly, forward-looking behaviour was introduced into these equations, together with wage equations, money demand and exchange rate equations. Not only do expectations of various prices appear, as in the LBS and Liverpool models, but also expectations of future output and personal income, so retaining the model's neo-Keynesian quantity adjustment approach.

The influence of the supply side on macroeconomic phenomena received increasing attention as a result of the major supply shocks of the 1970s and the criticisms of the new classical macroeconomists. Among US modellers and forecasters, for example, Klein (1978) in his AEA Presidential Address and Eckstein (1983), cited above, advocate a consensus approach in which mainstream models, now recognised to be over-emphasising effective demand, are extended by incorporating a full supply-side analysis into an appropriately elaborated IS–LM system. The over-emphasis on the demand side does not imply that supply theory had been neglected in the academic literature in preceding decades, as Eckstein (1983, p. 56) notes, but that body of work had

had little impact on the macroeconomics used for policy. In the United Kingdom the Conservative government elected in May 1979 paid increasing attention to supply-side policies, and the Liverpool model contained powerful supply effects in its representation of the labour market. As Matthews and Minford (1987) note, the size of the effects attributed to variations in unemployment benefits, direct taxes and trade union membership (proxying the power of unions) remain controversial, and their claim that these are well-determined empirically has been rejected by other researchers (for example, Nickell, 1987; Wallis et al., 1986, ch. 5). A further model, specifically designed as a supply-side model, namely that of the City University Business School (Beenstock et al., 1986) was first used for forecasting at the time of the June 1983 general election. The CUBS model abandoned the usual income-expenditure framework and determined the supply of output through a KLEM production function (capital, labour, energy, materials); it also included a formal labour supply schedule, unlike other models. Like the Liverpool model, it is an annual, not quarterly model, and emphasises the medium-term development of the economy. In the quarterly models the consensus has been slower to arrive than in the United States, but the current version of the LBS model (Dinenis et al., 1988) incorporates explicit supply influences into the income–expenditure framework. Output is determined not as the sum of the expenditure components but by both supply and demand factors at the sectoral level, with prices in the long run ensuring that the goods market clears.

IV. EVALUATION OF FORECASTS

The evaluation of past forecasting performance is an important input into the forecasting process and, since it provides a forecast of future performance, it is of interest to the users of forecasts. The forecasters themselves regularly publish accounts of their own performance, and occasionally contrast this with that of other groups. Several authors have undertaken independent studies of the performance of different forecasters with respect to a range of variables and forecast horizons, for example, in the United Kingdom Ash and Smyth (1973), Holden and Peel (1983, 1986) and Wallis et al. (1986, 1987), and in the United States McNees (1982, 1986) and Zarnowitz (1979, 1985) and references therein, including papers cited above; Artis (1988) examines the forecasting record of the IMF World Economic Outlook, and compares it with that of the separate national forecasting agencies. Evaulations range from descriptive accounts of forecasting ability in particular periods, especially turning points, as in the preceding sections, to the statistical analysis of forecast errors over a period of years, considered in this section. Evaluations are typically addressed to one variable at a time, whereas a multivariate assessment may be more relevant, particularly if trade-offs between different variables have a bearing on the specific decision problem.

Given time series of forecasts and outcomes, the first step is usually to calculate summary statistics such as the mean absolute error (MAE) and root mean square error (RMSE), or comparable statistics for the forecasts and

outcomes separately, or even their correlation coefficient. An immediate difficulty is that there is no absolute measure of the forecastability of a series, and so there is no absolute standard against which to compare these summary statistics. For a given definition of optimality, usually linear least squares, statistical prediction theory provides the optimal forecast with respect to a given information set, but economic forecasters may not agree about the relevant information set, which in the widest sense is in any event unknown, *ex ante* and *ex post*, and unmanageably large. In the absence of an absolute standard, various comparative procedures have been developed, discussed in the following sections. The first approach is to test whether the forecast satisfies certain properties of an optimal forecast, other than that of minimum mean square error. The second approach is to limit attention to a particularly restricted information set, namely that comprising past values of the variable of interest alone, and to compare a given forecast with the 'pure time series' forecast based on this 'own-variable' information set. The proper interpretation of such comparisons is considered, together with some recent evidence on the variation over time of the performance of economic forecasts relative to a particularly simple time series forecast. A third possibility is to conduct comparisons across a number of models or forecasts, and the issues that these raise, together with the possibility of combining forecasts, are discussed next. Whereas all these approaches limit attention to the published *ex ante* forecast, some recent systematic appraisals of the major sources of forecast error, *ex post*, are then described. Finally, the assessment of forecast uncertainty is considered.

IV. 1 *Properties of Optimal Forecasts*

An optimal forecast, which is the same thing as a rational expectation in the macroeconomic context, is unbiased and efficient. That is, the forecast error has an expected value of zero and cannot be predicted by any variable in the information set: full use of the given information has already been made in constructing the forecast.

A simple test of unbiasedness is to calculate the sample mean forecast error and compare it to its standard error. Many studies, instead or in addition, estimate the realisation-forecast regression

$$A_t = \alpha + \beta F_t + u_t,$$

where A and F denote actual value and forecast respectively, and test the (joint) hypothesis $\alpha = 0$, $\beta = 1$. While this is often interpreted as a test of unbiasedness, since if $\alpha = 0$ and $\beta = 1$ the forecasts are unbiased, it is in fact a stricter test, and was originally presented as a test of efficiency by Mincer and Zarnowitz (1969, p. 9). Since

$$A_t \equiv F_t + e_t,$$

where e_t is the forecast error, the estimate of β in the above regression only deviates from 1 if F_t and e_t are correlated. Such a correlation indicates an

inefficient forecast, since the correlation could be exploited to help predict the forecast error and so improve the forecast. But a significant deviation of the estimates of α and β from o and 1, respectively, does not necessarily imply significant bias, for it is possible that the sample mean forecast error, $\bar{e}$, is nevertheless close to zero in such circumstances. Since the regression estimates of α and β are in general correlated, their individual t ratios provide inappropriate tests of the efficiency hypothesis, and a *joint* test is required: Artis (1988) presents examples in which the individual and joint tests are in conflict.

Granger and Newbold (1977, p. 284) raise a practical objection to the realisation-forecast regression, namely that the validity of the usual test procedures rests on the non-autocorrelation of u_t, which need not necessarily hold for sub-optimal one-step-ahead forecasts, and does not generally hold for optimal forecasts more than one step ahead: Hansen and Hodrick (1980) provide corrected estimates of the asymptotic covariance matrix to accommodate this possibility. These are used by Holden and Peel (1985), for example, who study NIESR forecasts of six variables one to four quarters ahead over the period 1975–80: the hypothesis $\alpha = 0$, $\beta = 1$ is rejected in only one of the twenty-four cases, namely for forecasts of inflation four quarters ahead. Zarnowitz (1985) studies ASA–NBER business outlook survey respondents and likewise finds that inflation is the difficult variable: across six variables, five horizons and 79 respondents, the hypothesis $\alpha = 0$, $\beta = 1$ is rejected at the 5 % level in 15·4 % of the tests, but nearly half of these rejections refer to the inflation forecasts. In both of these papers the hypothesis $\alpha = 0$, $\beta = 1$ is referred to as the unbiasedness hypothesis.

The NIESR inflation forecasts presented in Fig. 2 are similar to the four-quarter-ahead forecasts considered by Holden and Peel, but cover a longer period. It is clear from inspection that the forecast errors are positive during the period 1964–81 (except 1968, which has a zero error), indicating persistent underprediction of inflation, followed by persistently negative errors during 1982–87. Given this pattern it is not surprising that the realisation-forecast regression has a significant Durbin-Watson statistic, and in a joint test based on the corrected coefficient covariances the hypothesis $\alpha = 0$, $\beta = 1$ is rejected. The mean forecast error overall is 1·0 % p.a. The GDP forecasts, however, are unbiased according to these tests.

It is now recognised that the concept of efficiency underlying the realisation-forecast regression is a relatively weak one. Granger and Newbold (1977, p. 284) argue that $\alpha = 0$, $\beta = 1$ 'constitutes a necessary condition for forecast efficiency, but according to any acceptable interpretation of that word it cannot be regarded as a *sufficient* condition'; in particular, it neglects possible autocorrelation of the forecast error, as their counter-example implicitly suggests. Autocorrelation of one-step-ahead forecast errors indicates that the forecast is not making efficient use at least of the own-variable information set, since knowledge of past forecast errors for the variable in question can then improve current forecasts. Errors in forecasts more than one step ahead cumulate step-by-step and so are autocorrelated, as noted above, but the errors

in an optimal n-step-ahead forecast exhibit autocorrelation of order n-1, not n, so this cannot be exploited to improve the forecast: it is efficient with respect to the own-variable information set. This efficiency property can be tested in a variety of ways. Holden and Peel (1985) regress the forecast error on the four most recent values of the variable known when the forecast was made. Of the twenty-four cases considered, only the inflation forecasts over three and four quarters fail this test, and the remaining NIESR forecasts could not be improved by using own-variable information more efficiently.

Increased use of the rational expectations hypothesis has been associated with a further strengthening of the concept of efficiency, which in its most extreme form, 'full rationality', requires that all available information be used in an optimal manner in constructing a forecast. Efficiency with respect to an information set containing other variables can be tested as in the preceding paragraph, by regressing the forecast error on lagged values of these variables known at the time the forecast was prepared. But the notion of 'all available' information presents practical difficulties, and whereas a rejection of full rationality in such a test is convincing, a failure to reject does not dispel the thought that there might be a relevant variable, untested, lurking around the corner. Brown and Maital (1981) assess one of the best-known surveys of experts' anticipations, namely the Livingston data, over the period 1962–77. Of the nine variables considered, only the wholesale price inflation expectations were found to be inefficient with respect to own-variable information, and the forecast errors in this variable and consumer price inflation could each be partly explained by past values of the other variable, indicating that their interrelationship was not properly appreciated by the forecasters. The more interesting finding, however, was that monetary growth helped to explain the forecast errors in both inflation variables, indicating that had monetary growth been correctly understood and fully incorporated into expectations over this period, the forecasts would have been considerably improved. Neglect of monetary influences does not appear to be the explanation of the problems in the NIESR annual inflation forecasts described above, however.

IV.2 *Comparisons with Time-series Forecasts*

Forecasts based exclusively on the statistical time-series properties of the variable in question have often been used to provide a yardstick against which economic forecasts, whether model-based or not, can be assessed. At a conference in 1972, Granger and Newbold (1975) commented that 'so far the sparring partner [the time series forecast] is consistently out-pointing the potential champion [the econometric model]', but the potential champion quickly reached match-fitness. A typical result is that of McNees (1982), who finds that published model forecasts generally outperform their time series competitors, the margin being greater four quarters ahead than one quarter ahead.

Comparison of a given forecast with a particularly naive alternative is implicit in a widely used 'inequality coefficient', attributed to Theil, defined as the ratio of the RMSE of the forecast to the RMSE of a 'no-change' forecasting

rule. This rule projects forward the last available observed value, and so the error in such a one-step-ahead forecast is simply the first difference of the variable in question. This inequality coefficient appears in Theil's *Applied Economic Forecasting* (1966), also in Ferber and Verdoorn (1962), Zarnowitz (1967), and elsewhere; in *Economic Forecasts and Policy* (1958) Theil had proposed a coefficient with a different denominator, designed to ensure that the coefficient lies between 0 and 1, but this has other disadvantages, as noted by Ferber and Verdoorn (1962, ch. 10.4), Sims (1967), Granger and Newbold (1977, pp. 281–2), and by Theil (1966, p. 28) himself. Sims (1967) and Zarnowitz (1967) also use inequality coefficients in which the denominator is the RMSE of less naive time series forecasts, initiating a trend in forecast comparisons that has continued since that time. After the initial use of no-change or 'same-change' forecasting rules, autoregressive models and the ARIMA models of Box and Jenkins (1970) were employed as benchmarks. The increased complexity of these time series methods itself indicates the progress that has occurred in forecasting models since their early development. Finally, moving to a multivariate context, vector autoregressive (VAR) models as used for forecasting by Litterman (1986) have entered the competition. In this last respect the US picture is somewhat mixed: VAR forecasts are part of the comparison carried out by McNees (1982), summarised above; with Litterman's Bayesian modification they also feature in the comparison of McNees (1986), where they are generally the most accurate or among the most accurate for real GNP, unemployment and investment. Curiously, for four of the seven variables considered by McNees, the RMSE of the VAR forecast declines as the forecast horizon increases from one to eight quarters. In the United Kingdom, however, VAR forecasts have not been found to dominate published model-based forecasts (Wallis *et al.*, 1986, 1987).

Formal comparisons between model-based forecasts and time series forecasts face two difficulties. First, since the data used in empirical specification and estimation of the two forms are the same, their forecasts, forecast errors and resulting summary measures are not statistically independent, hence a formal test based on a direct comparison of the two forecast error variances, such as an F test, cannot be employed. Secondly, at the theoretical level, a univariate ARIMA model can be regarded as an approximation to a solution form of an econometric model (its 'final equation'), and hence again cannot provide an independent check on the econometric model, in terms of forecast or any other comparison. Since the time series models emphasise dynamic and stochastic features of the data, early comparisons in which they outperformed econometric models simply suggest that the latter were deficient in these respects (Prothero and Wallis, 1976). Indeed, the equations of such models often exhibited substantial residual autocorrelation. Subsequently, the dynamic and stochastic specification of the large-scale econometric models has improved, through the application of developments in time-series econometrics. In the UK context, these stem from the classic Colston paper by Sargan (1964); see, for example, Wallis (1972), Hendry (1974) and Hendry and Richard (1983). In particular, comparisons with time series models now can be seen to represent a useful

diagnostic device during the model specification process. The advice offered at that same 1972 conference, 'that a suggested specification should be tested in all possible ways, and only those specifications which survive and correspond to a reasonable economic model should be used' (Sargan, 1975) has been heeded, and a range of diagnostic tests is readily available in user-friendly micro-computer software (for example, Hendry, 1986).

Finally, an inequality coefficient based on a naive time-series forecast is reconsidered, for a different purpose. In his analysis of Treasury forecasts, Burns (1986) calculates five-year moving averages of absolute forecast errors and compares these to an index of variation which gives some impression of the difficulty of forecasting at different points in time. The index is defined as the five-year moving average of the absolute error of a no-change forecast of the growth rate or the inflation rate, that is, the average of the absolute difference between successive years' growth or inflation rates; appropriate extensions are needed for comparison with two-year-ahead forecasts. The resulting ratio modifies the Theil inequality coefficient for forecasts of growth and inflation in two ways: it replaces RMSEs by MAEs, and it replaces the overall sample means by five-year moving averages. Burns' results for Treasury forecasts over one-year and two-year horizons are shown, slightly updated, in Fig. 3, together with equivalent information for LBS and NIESR forecasts. His broad conclusion about Treasury forecasts over this period is that while there has not been any marked improvement in accuracy of the short-term forecasts, the results for the two-year-ahead forecasts are more encouraging, with some evidence of improvement over time. Over the shorter period for which data are available, the LBS and NIESR two-year forecasts show no substantial change in performance in respect of these two variables, and the Treasury forecasts appear to dominate, which may reflect their more effective incorporation of policy measures. For the one-year-ahead forecasts, however, the ranking of the three groups varies over time, and at any point in time is not consistent across the two variables considered. This is illustrative of a general finding in cross-model comparisons, discussed next.

IV.3 *Cross-model Comparisons and Combinations*

Comparisons of the *ex ante* forecasts published by different forecasting groups place all competitors on an equal footing with respect to information about the future. If one forecaster uses a model which treats as exogenous a variable which another treats as endogenous, then the former needs to provide an off-model projected value for that variable. The difference in classification of variables is immaterial, provided that it is recognised that a forecaster-model combination is under scrutiny. Remembering also the process of adjustments described in Section I, it is clear that *ex ante* forecasts do not provide useful evaluations of models alone. Furthermore, since models evolve, as discussed above, comparisons based on summary statistics such as MAE or RMSE over a period of years are in effect evaluating forecasting groups together with whatever model specification was in use at the time. Their personnel changes over time, too, although Burns (1986) attributes some of the improvement in
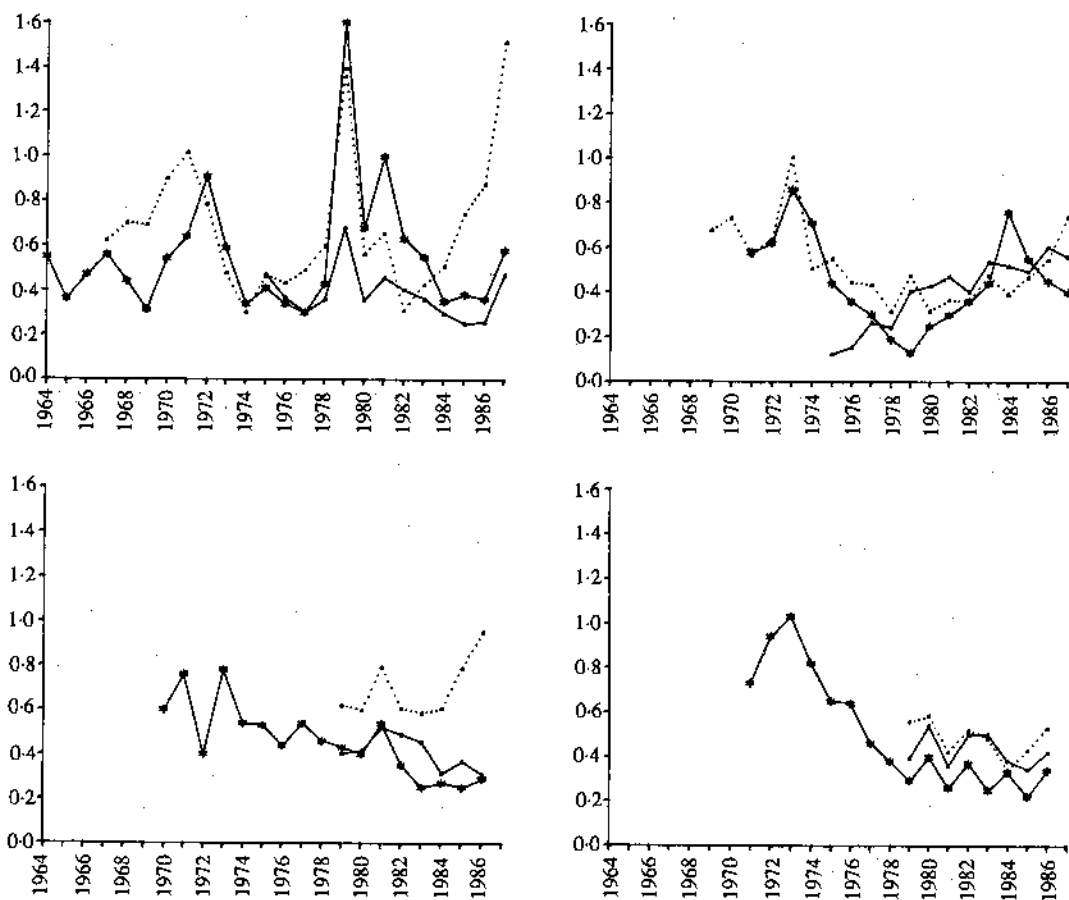
Fig. 3. Ratio of mean absolute error to index of variation (five-year moving averages shown against last year of period). Upper left, GDP growth, one-year-ahead forecast; lower left, GDP growth, two-year-ahead forecast; upper right, Inflation, one-year-ahead forecast; lower right, Inflation, two-year-ahead forecast. Key: —■—, LBS; ··■··, NIESR; —•—, HMT.

Treasury medium-term forecasts to an increase in human capital: 'many members of the forecasting teams were engaged in forecasting for many years'.

One practical use of forecasts is in informing decision-making, and in principle the evaluation of forecasts can be associated with any specified loss function, relevant to the particular objective. Given the dynamic multivariate nature of the forecasts, this would then help to specify the time horizon and the set of variables to be considered. In practice this does not feature in the literature, since decision makers and the users of forecasts seldom discuss such matters, at least not in public. As a result comparative studies usually focus on the small number of key macroeconomic indicators that feature in public discussion, and the possibility that one forecaster may offer information about a further group of variables about which another forecaster is completely silent is usually ignored. Clearly, however, if a user's interest is in that further group

of variables, then there is no difficulty in deciding which forecaster is 'best'. Although the loss function is not made explicit, the common measures used to evaluate forecast accuracy imply particular forms of the loss function, the use of mean square error, for example, implying that it is quadratic. Also the period over which these are calculated implies a particular interest in short-term or long-term forecasts, for example.

Given the rather general nature of the criteria, and the existence of a free market in data and economic ideas, the common finding of comparative studies that there are no unambiguous rankings is perhaps not surprising. For example, Zarnowitz (1979) finds that rankings among six US forecasting groups show appreciable differences with respect to particular variables, subperiods and forecast horizons, but the differences in summary statistics are typically small: 'the main lesson ... is that the similarities greatly outweigh the differences between the forecasters' performance records'. In the United Kingdom, Holden and Peel (1986) compare LBS and NIESR quarterly forecasts of growth and inflation, and find that for both variables NIESR is preferred to LBS in 1975–9, whereas the ranking is reversed in 1980–4 (in each case both are superior to the forecasts published by a London firm of stockbrokers, however). As an example of a multivariate exercise, Wallis *et al.* (1987) consider a single forecast published in autumn 1984 by four groups – LBS, NIESR, LPL and CUBS – and study four variables, namely the level of GDP, its growth rate, the inflation rate and the unemployment rate. For each variable and each group the RMSE over the two years of the forecast period is calculated and then expressed as a ratio of the average for a given variable across all groups, in order to standardise the comparison in respect of the degree of difficulty of forecasting that variable. The relative RMSEs are shown in Fig. 4: a value less than one for a given group and a given variable implies better-than-average performance for that variable. If one tetrahedron lies inside another, then there is an unambiguous ranking not only when the variables are deemed to have equal importance, as in the illustration, but for all values of their relative weights. While this occurs in this particular short-term forecast for one pair, LBS and LPL, in all other pairwise comparisons the tetrahedra intersect and so no rankings that are invariant to the choice of weights on the different variables can be derived.

The statistical significance of differences in summary statistics such as the RMSE cannot be directly tested, as noted above, since the competing forecasts are not independent. The idea of combining forecasts (Bates and Granger, 1969) offers an alternative approach to testing the *ex post* performance of competing forecasts, and also a method of improving *ex ante* forecasting performance. Consider a combination of one forecast with its competitor: if the combined forecast has an error variance that is not significantly smaller than that of the first forecast then the competing forecast appears to offer no useful additional information. On the other hand, different forecasts based on different informations sets, different models, or even different approaches to data analysis, as in econometric vs. time series comparisons, in general may each be expected to contribute usefully to the forecasting problem, and so the
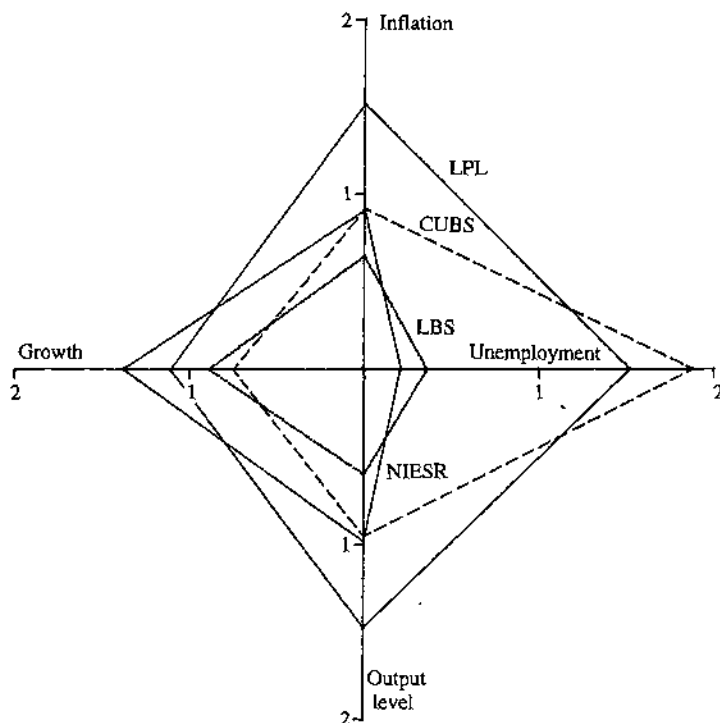
Fig. 4. Relative root mean square errors. The average root mean square errors are: GDP level, 2 % of actual; GDP growth rate, 1·0 % per annum; inflation rate 1·8 % per annum; unemployment rate, 0·7 %.

combined forecast may be more accurate than any of the individual components. Thus an *ex post* comparison of two forecasts, F_{1t} and F_{2t}, may be based on the regression equation

$$A_t = \beta F_{1t} + (1 - \beta) F_{2t} + u_t$$

and a test of the null hypothesis $\beta = 1$, while a combination of *ex ante* forecasts for subsequent use may be based on the estimated coefficient.

Recent discussion (Granger and Ramanathan, 1984; Clemen, 1986; Holden and Peel, 1986) has focused on the general combination of k forecasts through the equation

$$A_t = \alpha + \beta_1 F_{1t} + \beta_2 F_{2t} + \ldots + \beta_k F_{kt} + u_t$$

and the question of whether or not the coefficient restrictions

$$\alpha = 0, \quad \beta_1 + \beta_2 + \ldots + \beta_k = 1$$

should be imposed. The unconstrained regression clearly achieves a smaller residual sum of squares *ex post*, as Granger and Ramanathan indicate, but the practical objective is to improve *ex ante* forecast performance, and Clemen shows that the impositon of the restrictions can improve forecast efficiency. Whereas some forecasts may be found to be biased, *ex post*, over certain periods,

each individual *ex ante* forecast is offered as a best estimate of the future outcome, corrected for any past bias through the best endeavours of the forecaster. Holden and Peel (1986) estimate the above regression using only data available at a given point in the past, and with the estimated coefficients form a combined forecast of a future outcome, which is then compared with the actual outcomes: for quarterly forecasts of growth and inflation, a restricted combination always has a smaller RMSE than an unrestricted combination. The three 'economic' forecasts noted above are considered, along with three time series forecasts, and the winning combination is that based on the three economic forecasts. However, given the need to estimate the regression coefficients, the comparison is restricted to the later period of available data, namely 1980–4: here the LBS provides the best economic forecast, as noted above, and this single forecast also dominates the combined forecast over this period.

IV.4 *Decomposition of Forecast Error*

Ex ante forecasts are of little assistance in the evaluation of models, as noted above, since the forecast is a joint product of model and forecaster. Statistical summaries of errors in published forecasts provide no information about possible causes of forecast error. Errors may arise because the model is inadequate and because exogenous variables behave in a different manner to that on which the forecast is conditioned. Recomputation of the forecast, *ex post*, using realised values of exogenous variables, allows this latter source of error to be assessed. Forecasters make adjustments to the pure model forecasts in an attempt to overcome perceived model inadequacies and to improve the forecast in other ways, and the extent to which this succeeds may be assessed by recomputing the forecast with such adjustments removed. There is little consideration in the literature of these various sources of forecast error, an exception being Osborn and Teal (1979), who analyse two NIESR forecasts. In general, independent researchers have not been able to recompute forecasts under alternative assumptions, using the precise models on which the forecasts were based. An exception, however, is the ESRC Macroeconomic Modelling Bureau, which has taken successive deposits of models and forecasts since autumn 1983, and archived this information. With the passage of time comparative *ex post* forecast error decompositions become feasible, as described in this section, based on Wallis *et al.* (1986, ch. 4; 1987, ch. 4).

Ex post forecast comparisons give an advantage to a model that treats as exogenous a variable that is difficult to forecast, when other models are trying to explain its behaviour. Accordingly such comparisons are often criticised, because different models may have different degrees of exogeneity. Practical experience with the UK models, however, is that the broad classification of variables as endogenous or exogenous does not differ across models. In general, the variables treated as exogenous in models of small open economies fall into three main groups, describing respectively the economic environment in the rest of the world, domestic economic policy, and various natural resource and demographic trends, such as the growth of North Sea oil production and

changes in the population of working age. Within these generally agreed exogenous areas differences of detail may arise, for the same reasons that apply in other areas of the models. For example, different models may measure an identical concept (the world price of oil, the real wage) in different ways; some concepts (world money supply) may appear in some models but not in others; the level of aggregation (distinguishing the price of oil from the general price level) may differ. Also, where the line is drawn between exogenous assumption and endogenous consequence may differ, particularly in respect of domestic economic policy: the more medium-term models tend to take as given a broad policy stance and allow the response of public expenditure and taxation to emerging pressures to be determined by the model, whereas the more short-term models take the detailed policy settings as given. Nevertheless these differences are not sufficient to inhibit cross-model comparisons.

Wallis *et al.* (1986, 1987) compute two variant forecasts after the event. First, the projected values of variables treated as exogenous are replaced by the actual outcomes: a comparison with the published forecast indicates the effect of incorrect anticipations about external developments. Secondly, by setting all residual adjustments to zero an *ex post* forecast, variously described as pure model-based, mechanical, or hands-off, is obtained. The error in this forecast includes the effect of model misspecification and the contribution of random disturbances in the forecast period, together with any effect of data revisions not explicitly accommodated elsewhere. The forecaster's residual adjustments represent an attempt to reduce this error and a comparison with the first variant indicates how successful this was. The error in the published forecast can then be decomposed as illustrated in Fig. 5, which describes the 1983 and 1984 forecasts of the price level, one and two years ahead, from four groups. In each segment of the figure the first block gives the error in the *ex post* hands-off forecast ('model error'), and the second block is the contribution of the forecaster's adjustments. These should be of opposite sign if the adjustments indeed act in an offsetting manner: in the majority of cases this has occurred, although the adjustment is seldom of the correct magnitude and sometimes rather large offsets were called for. The third block gives the contribution of errors in projecting exogenous variables, and these three contributions sum to the fourth block, namely, the error in the published forecast. In all cases these last errors increase as the forecast horizon increases, as expected. Overall the LPL group persistently underestimated inflation, that is, took an overoptimistic view in this respect, whereas the other three groups tended to be unduly pessimistic. It is often argued that a good track record in forecasting is a prerequisite for the use of a model in policy analysis, but it is important to disentangle the role of the model and the role of the forecaster, as in this illustration.

Perhaps the most striking feature of Fig. 5 is the contribution of exogenous uncertainty to the forecasts for 1986. In the first three cases the fortuitous choice of wrong exogenous assumptions has resulted in a relatively small final error, while in the LPL forecast the exogenous outcome transposes a pessimistic *ex post* adjusted forecast into an optimistic one. The principal cause is the
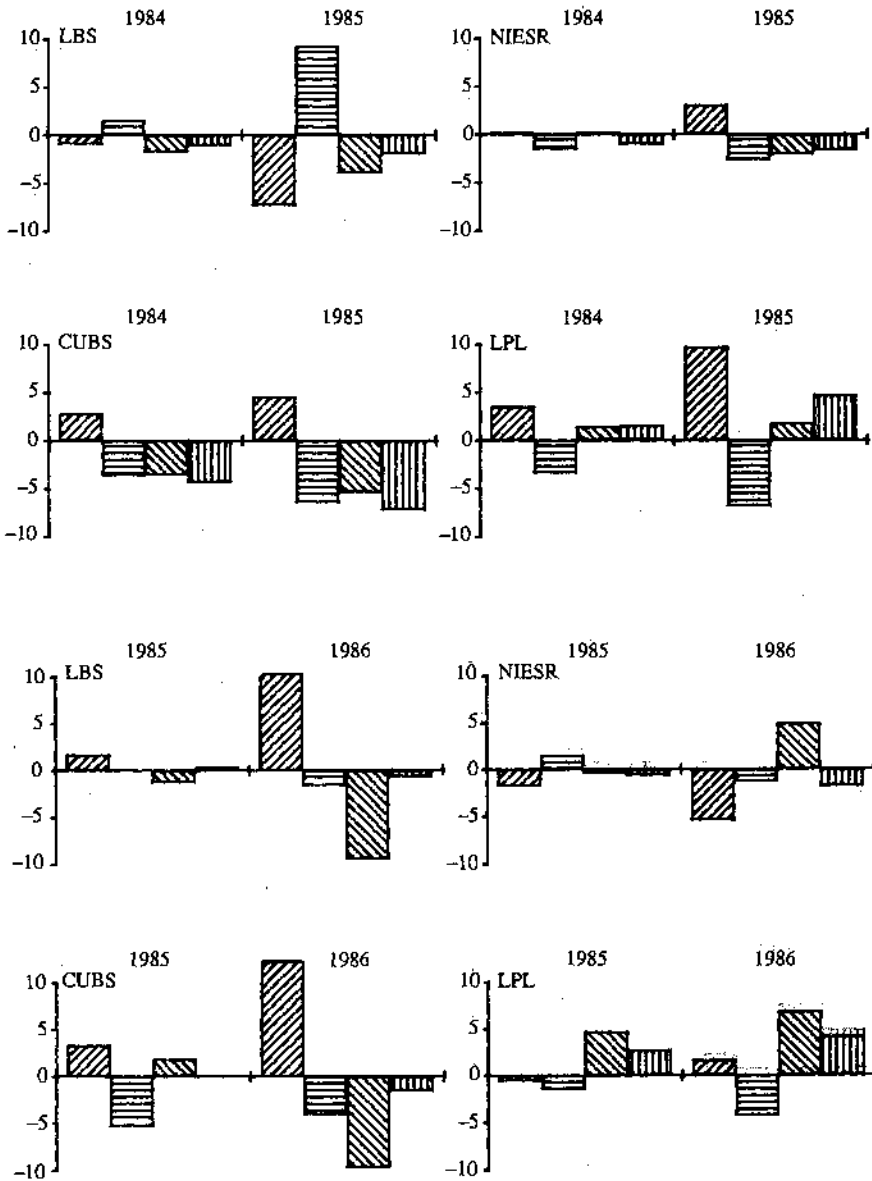
Fig. 5. Contributions to errors in price level forecasts, one and two years ahead (% of actual). Upper panel, autumn 1983 forecasts; lower panel, autumn 1984 forecasts. Key: ▨, model error; ▤, contribution of residuals; ▧, contribution of exogenous uncertainty; ▥, error in published forecast.

unanticipated and unprecedented fall in the world oil price in 1986. This variable appears directly in the LBS, NIESR and CUBS models, which would have produced substantial forecast errors had its movements been anticipated, as the first block of the decomposition indicates. The fall in the world oil price was associated with a level of world demand lower than anticipated, the

depressing impact of lower oil prices on the oil-producing countries being more immediate than the expansionary impact in the oil-importing countries, and this is the principal source of exogenous error in the LPL model, where oil prices do not appear. In the other three cases there is no agreement: LBS and CUBS show a large positive model error and knowledge of the forthcoming oil price fall would have led them to underestimate the domestic price level by a considerable amount, whereas in the NIESR forecast lower oil prices would have increased domestic prices via the lower exchange rate. This is the principal transmission mechanism in this case, and the actual impact on exchange rates of the oil price fall was much smaller than suggested by the then-current models. These were projecting forward from a different kind of sample experience and did not reflect whatever asymmetric effects were in operation. Once again, a change in data characteristics has been informative, and including this episode in the sample period has resulted in newly estimated equations that are much less sensitive to world oil prices.

A principal use of *ex post* forecast assessment is in studying the performance of models, as noted above. As far as the forecasts themselves are concerned, whereas calculations such as those that lie behind Fig. 5 permit discussion of what might have been, they cannot tell what would have been. The production of a forecast is an interactive process, and had the oil price fall been anticipated forecasters might have realised that its effects would not simply be the negative of the effects of the preceding oil price rise. Adjustments to the pure model-based forecast might then have been made, anticipating the model modifications that occurred subsequently. However forecasters seldom reveal the rationale for specific adjustments that they make (revelation of the adjustments themselves is a relatively new phenomenon, as discussed below), and it is not possible to reconstruct the forecasting team and its state of mind at the time in order to answer 'what-if' questions. Nevertheless a model that produces accurate *ex post* forecasts only after substantial adjustment must represent an unreliable vehicle for policy analysis.

IV.5 *Assessing Uncertainty*

Forecasts are subject to errors, and the user of forecasts needs to know the likely margin of error. Theoretical discussion of model-based forecasts usually takes the model as given and then analyses the forecast error in terms of its three sources: the model's random disturbance terms, its coefficient estimation errors, and forecast errors in exogenous variables. In the textbook linear model formal expressions for the variance attributable to each source are available, and in the practical nonlinear model stochastic simulation can be employed to estimate these quantities. This approach is of little help to the practitioner. It neglects the contribution of the forecaster's subjective adjustments, and its last element simply pushes the problem of assessing the margin of forecast error one stage further back, from endogenous to exogenous variables. More fundamentally, the model's specification is uncertain. At any point in time competing models coexist, over time model specifications evolve, and there is no way of assessing this uncertainty. Thus the only practical indication of the likely margin of

future error is provided by the past forecast errors. Published evaluations of forecasts, such as those cited above, are a source of this information, and the forecasters themselves commonly provide an estimate of their past performance. For example, Treasury forecasts published in the annual *Financial Statement and Budget Report* (the 'Red Book') are accompanied by the MAE of the preceding ten years' forecasts.

Estimating the future margin of error is itself a forecasting problem, in that not only the first moment but also the second moment of the conditional probability distribution of future outcomes is now under consideration. As with the first moment, errors in forecasting the second moment are likely to occur at times of rapid change in the underlying situation, when the economy becomes harder or easier to forecast for some reason. Standardising measures of forecast performance with reference to the underlying variation, as in the examples in Fig. 3, accommodates such changes in the past, but in projecting forward the underlying variation is assumed to remain constant, unless this forecast is in turn subjectively adjusted. For example, in discussing the margin of error of Treasury forecasts, the *Financial Statement and Budget Report* that accompanied the June 1979 Budget noted the 'possibility that large changes in policy will affect the economy in ways which are not foreseen'.

The uncertainty of a single forecast, indicated by the spread of the probability distribution of possible outcomes, should be distinguished from disagreement among several (point) forecasts, as Zarnowitz and Lambros (1987) note. Consensus among forecasters need not imply a high degree of confidence about the commonly predicted outcome. To study this question, Zarnowitz and Lambros analyse the ASA-NBER survey responses on GNP growth and inflation, in which individual respondents provide their subjective (joint) probability distribution of outcomes. It is found that the variance of point forecasts tends to understate uncertainty as measured by the variance of the predictive probability distributions. The former varies much more over time than the latter, although these measures of consensus and uncertainty are positively correlated.

V. CONCLUDING COMMENTS

This paper surveys developments in macroeconomic forecasting over the last twenty years, during which forecasting methods have become more formalised and considerable progress has been made. 'Success in forecasting may be occasional and fortuitous or intuitive, but progress in forecasting, to the extent it is possible, can only come from advances of science, not art or chance' (Zarnowitz, 1979). Such advances have occurred gradually, with small improvements building on one another, 'but nothing really has been a complete breakthrough for solving the problems that confront us' (Klein, 1987). Forecasting disappointments led not to complete changes of direction but to constructive reappraisals and eventual improvements, also changes of emphasis. McNees (1988) concludes that 'annual forecasts of real GNP and the inflation rate have improved over time. Summary error measures have

declined slightly in absolute terms but even more relative to either naive standards of comparison or the variability of the actual outcomes.'

The role of forecasts in macroeconomic policy-making, and the policy-making itself, have seen considerable changes during this period. Although the uncertainty in forecasting was clearly appreciated, Cairncross (1969) noting that 'the margins of error... even when small in relation to the magnitudes involved, remain large in relation to policy objectives and to the reserves held against contingencies', during the years of demand management short-term forecasts played an increasing role in fine tuning the level of demand. But by 1976, as Burns (1986) notes, 'there was not only disillusion with demand management; there was also growing frustration with the forecasts' as the increased level of noise in the economic system led to increased margins of error. In the 1980s the focus of the forecasts and policy analysis has shifted to a more medium-term horizon, and it is here that 'many of the developments in our understanding of the way economies work are likely to have their pay-off', again quoting Burns.

Cairncross (1969) concluded his Presidential Address by commenting on the issue of the publication of official forecasts. While acknowledging that there were 'great advantages in making official forecasts as widely available as possible', he also emphasised 'some of the inconveniences from the Ministerial point of view in putting official forecasts in front of an unsophisticated public' and expressed his own view 'that the issue is not nearly as important as it appears'. In any event, the Bray amendment to the 1975 Industry Act required the Treasury to publish forecasts produced with the aid of the model at least twice a year, and the first 'Industry Act forecast' was published in December 1976. As noted above, these are accompanied by indications of forecasting error, also as required by the Act. Five years after the Act, the Treasury could conclude that 'the impact of the Government's publishing forecasts has been, perhaps, more limited than was suggested by some of the strongly-held beliefs – pre-1976 – for and against publication' (HM Treasury, 1981).

Considerable improvements in publication and dissemination of forecasts have occurred in the period under review. Models and forecasts have become better documented and more widely accessible, most recently being disseminated for implementation on micro-computers. In the United Kingdom, forecast analysis such as that described in Section IV. 4 is possible because the models, forecasts, and associated databases are made available to an independent research group (the ESRC Macroeconomic Modelling Bureau) charged with improving accessibility and undertaking comparative research. Before the establishment of the Bureau in 1983, the LBS forecasters had begun to publish the residual adjustments used in each forecast in their *Economic Outlook*. (Treasury forecasts do not enter such comparative analysis, however, since although the model is publicly available, again thanks to the Industry Act, the forecast assumptions and adjustments are not.) This degree of openness is unique, and perhaps could not be achieved in countries where the leading forecasting groups are commercial. Nevertheless in such countries public bodies are often among the main customers of commercial forecasters, and when such

forecasts enter the public debate the same arguments about understanding the assumptions and replicating the results apply. Models and forecasts benefit from public discussion and assessment, and in most countries much more could be done.

University of Warwick

Date of receipt of final typescript: September 1988

REFERENCES

Artis, M. J. (1988). 'How accurate is the World Economic Outlook? A *post mortem* on short-term forecasting at the International Monetary Fund.' In *Staff Studies for the World Economic Outlook*, July 1988, pp. 1–49. Washington, DC: International Monetary Fund.

Ash, J. C. K. and Smyth, D. J. (1973). *Forecasting the United Kingdom Economy*. Farnborough: Saxon House.

Ball, R. J. and Burns, T. (1978). 'Stabilisation policy in Britain 1964–81.' In *Demand Management* (ed. M. V. Posner), pp. 66–100. London: Heinemann.

——, —— and Warburton, P. J. (1979). 'The London Business School model of the UK economy: an exercise in international monetarism.' In *Economic Modelling* (ed. P. A. Ormerod), pp. 86–114. London: Heinemann.

Barker, T. S. (1985). 'Forecasting the economic recession in the U.K. 1979–1982: a comparison of model-based *ex ante* forecasts.' *Journal of Forecasting*, vol. 4, pp. 133–51.

Barro, R. J. and Fischer, S. (1976). 'Recent developments in monetary theory.' *Journal of Monetary Economics*, vol. 2, pp. 133–67.

Bates, J. M. and Granger, C. W. J. (1969). 'The combination of forecasts.' *Operational Research Quarterly*, vol. 20, pp. 451–68.

Beenstock, M., Warburton, P., Lewington, P. and Dalziel, A. (1986). 'A macroeconomic model of aggregate supply and demand for the UK.' *Economic Modelling*, vol. 3, pp. 242–68.

Blanchard, O. J. (1987). 'Neoclassical synthesis.' In *The New Palgrave: A Dictionary of Economics*, vol. 3 (ed. J. Eatwell, M. Milgate and P. Newman), pp. 634–6. London: Macmillan.

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

Brown, B. W. and Maital, S. (1981). 'What do economists know? An empirical study of experts' expectations.' *Econometrica*, vol. 49, pp. 491–504.

Buiter, W. H. and Miller, M. H. (1981). 'The Thatcher experiment: the first two years.' *Brookings Papers on Economic Activity*, no. 2, pp. 315–67.

—— and —— (1983). 'Changing the rules: economic consequences of the Thatcher regime.' *Brookings Papers on Economic Activity*, no. 2, pp. 305–65.

Burns, T. (1986). 'The interpretation and use of economic predictions.' *Proceedings of the Royal Society of London*, A, no. 407, pp. 103–25. Reprinted in *Predictability in Science and Society* (ed. J. Mason, P. Mathias and J. H. Westcott). London: Royal Society and British Academy.

Cairncross, A. (1969). 'Economic forecasting.' ECONOMIC JOURNAL, vol. 79, pp. 797–812.

Chow, G. C. and Corsi, P. (eds.) (1982). *Evaluating the Reliability of Macroeconomic Models*. New York: Wiley.

Clemen, R. T. (1986). 'Linear constraints and the efficiency of combined forecasts.' *Journal of Forecasting*, vol. 5, pp. 31–8.

Currie, D. A. (1985). 'Macroeconomic policy design and control theory – a failed partnership?' ECONOMIC JOURNAL, vol. 95, pp. 285–306.

Davidson, J. E. H., Hendry, D. F., Srba, F. and Yeo, S. (1978). 'Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom.' ECONOMIC JOURNAL, vol. 88, pp. 661–92.

Dinenis, E., Holly, S., Levine, P. and Smith, P. (1988). 'The London Business School econometric model: some recent developments.' *Economic Modelling*, forthcoming.

Dornbusch, R. (1976). 'Expectations and exchange rate dynamics.' *Journal of Political Economy*, vol. 84, pp. 1161–76.

Eckstein, O. (1983). *The DRI Model of the U.S. Economy*. New York: McGraw-Hill.

Fair, R. C. (1984). *Specification, Estimation and Analysis of Macroeconometric Models*. Cambridge, Mass: Harvard University Press.

Ferber, R. and Verdoorn, P. J. (1962). *Research Methods in Economics and Business*. New York: Macmillan.

Fischer, S. (1988). 'Recent developments in macroeconomics.' ECONOMIC JOURNAL, vol. 98, pp. 294–339.

Flemming, J. S. (1978). Review of *Demand Management* (ed. M. V. Posner), Heinemann, 1978. *Times Higher Educational Supplement*, (15 September).

Granger, C. W. J. and Newbold, P. (1975). 'Economic forecasting: the atheist's viewpoint.' In *Modelling the Economy* (ed. G. A. Renton), pp. 131–47. London: Heinemann.

—— and —— (1977). *Forecasting Economic Time Series*. New York: Academic Press. (Second edition, 1986.)

—— and Ramanathan, R. (1984). 'Improved methods of combining forecasts.' *Journal of Forecasting*, vol. 3, pp. 197–204.

Grunberg, E. and Modigliani, F. (1954). 'The predictability of social events.' *Journal of Political Economy*, vol. 62, pp. 465–78.

Hall, S. G. and Henry, S. G. B. (1985). 'Rational expectations in an econometric model: NIESR model 8.' *National Institute Economic Review*, No. 114, pp. 58–68.

Hansen, L. P. and Hodrick, R. J. (1980). 'Forward exchange rates as optimal predictors of future spot rates: an econometric analysis.' *Journal of Political Economy*, vol. 88, pp. 829–53.

Heller, W. W. *et al.* (1968). *Fiscal Policy for a Balanced Economy*. Paris: OECD.

Hendry, D. F. (1974). 'Stochastic specification in an aggregate demand model of the United Kingdom.' *Econometrica*, vol. 42, pp. 559–78.

—— (1986). 'Using PC-GIVE in econometrics teaching.' *Oxford Bulletin of Economics and Statistics*, vol. 48, pp. 87–98.

—— and Richard, J. F. (1983). 'The econometric analysis of economic time series.' *International Statistical Review*, vol. 51, pp. 111–63.

HM Treasury (1981). 'Forecasting in the Treasury.' *Economic Progress Report*, No. 134 (June), pp. 2–6.

Higgins, C. I. (1988). 'Empirical analysis and intergovernmental policy consultation.' In *Empirical Macroeconomics for Interdependent Economies* (ed. R. C. Bryant *et al.*), pp. 285–302. Washington, DC: Brookings Institution.

Hilton, K. and Heathfield, D. F. (eds.) (1970). *The Econometric Study of the United Kingdom*. London: Macmillan.

Holden, K. and Peel, D. A. (1983). 'Forecasts and expectations: some evidence for the UK.' *Journal of Forecasting*, vol. 2, pp. 51–8.

—— and —— (1985). 'An evaluation of quarterly National Institute forecasts.' *Journal of Forecasting*, vol. 4, pp. 227–34.

—— and —— (1986). 'An empirical investigation of combinations of economic forecasts.' *Journal of Forecasting*, vol. 5, pp. 229–42.

Isard, P. (1988). 'Exchange rate modeling: an assessment of alternative approaches.' In *Empirical Macroeconomics for Interdependent Economies* (ed. R. C. Bryant *et al.*), pp. 183–201. Washington, DC: Brookings Institution.

Keating, G. (1985). *The Production and Use of Economic Forecasts*. London: Methuen.

Klein, L. R. (1968). *An Essay on the Theory of Economic Prediction*. Helsinki: Yrjo Jahnsson Foundation.

—— (1974). 'Supply constraints in demand-oriented systems: an interpretation of the oil crisis.' *Zeitschrift für Nationalökonomie*, vol. 34, pp. 45–56.

—— (1978). 'The supply side.' *American Economic Review*, vol. 68, pp. 1–7.

—— (1986). 'Economic policy formation: theory and implementation (applied econometrics in the public sector).' In *Handbook of Econometrics*, vol. 3 (Z. Griliches and M. D. Intriligator, eds.), pp. 2057–93. Amsterdam: North-Holland.

—— (1987). 'The ET interview: Professor L. R. Klein interviewed by Roberto S. Mariano.' *Econometric Theory*, vol. 3, pp. 409–60.

—— Ball, R. J., Hazlewood, A. and Vandome, P. (1961). *An Econometric Model of the United Kingdom*. Oxford: Basil Blackwell.

—— Friedman, E. and Able, S. (1983). 'Money in the Wharton quarterly model.' *Journal of Money, Credit and Banking*, vol. 15, pp. 237–59.

—— and Young, R. M. (1980). *An Introduction to Econometric Forecasting and Forecasting Models*. Lexington, Mass: D. C. Heath.

Kmenta, J. and Ramsey, J. B. (eds.) (1981). *Large-Scale Macroeconometric Models: Theory and Practice*. Amsterdam: North-Holland.

Litterman, R. B. (1986). 'Forecasting with Bayesian vector autoregressions – five years of experience.' *Journal of Business and Economic Statistics*, vol. 4, pp. 25–38.

Llewellyn, J., Potter, S. and Samuelson, L. (1985). *Economic Forecasting and Policy – the International Dimension*. London: Routledge and Kegan Paul.

Lucas, R. E. (1975). 'An equilibrium model of the business cycle.' *Journal of Political Economy*, vol. 83, pp. 1113–44.

—— (1977). 'Understanding business cycles.' In *Stabilization of the Domestic and International Economy* (Carnegie-Rochester Series on Public Policy, vol. 5, ed. K. Brunner and A. H. Meltzer), pp. 7–29. Amsterdam: North-Holland.

—— (1980). 'Methods and problems in business cycle theory.' *Journal of Money, Credit and Banking*, vol. 12, pp. 696–715.

—— and Sargent, T. J. 'After Keynesian macroeconomics.' In *After the Phillips Curve: Persistence of High Inflation and High Unemployment*, pp. 44–72. Federal Reserve Bank of Boston, Conference Series No. 19.

McNees, S. K. (1979a). 'The forecasting record for the 1970s.' *New England Economic Review* (September/October), pp. 33–53.

—— (1979b). 'Lessons from the track record of macroeconomic forecasts in the 1970s.' In *Forecasting* (TIMS Studies in the Management Sciences, vol. 12, ed. S. Makridakis and S. C. Wheelwright), pp. 227–46. Amsterdam: North-Holland.

—— (1982). 'The role of macroeconometric models in forecasting and policy analysis in the United States.' *Journal of Forecasting*, vol. 1, pp. 37–48.

—— (1986). 'Forecasting accuracy of alternative techniques: a comparison of U.S. macroeconomic forecasts (with discussion).' *Journal of Business and Economic Statistics*, vol. 4, pp. 5–23.

—— (1988). 'How accurate are macroeconomic forecasts?' *New England Economic Review* (July/August), pp. 15–36.

Malgrange, P. and Muet, P. A. (eds.) (1984). *Contemporary Macroeconomic Modelling*. Oxford: Blackwell.

Matthews, K. G. P. and Minford, A. P. L. (1987). 'Mrs Thatcher's economic policies 1979–87 (with discussion).' *Economic Policy*, vol. 5, pp. 59–101.

Mincer, J. and Zarnowitz, V. (1969). 'The evaluation of economic forecasts.' In *Economic Forecasts and Expectations*, National Bureau of Economic Research Studies in Business Cycles No. 19 (ed. J. Mincer), pp. 3–46. New York: Columbia University Press.

Minford, A. P. L., Marwaha, S., Matthews, K. and Sprague, A. (1984). 'The Liverpool macroeconomic model of the United Kingdom.' *Economic Modelling*, vol. 1, pp. 24–62.

Muth, J. F. (1961). 'Rational expectations and the theory of price movements.' *Econometrica*, vol. 29, pp. 315–35.

Nickell, S. J. (1987). Discussion (of Matthews and Minford, 1987). *Economic Policy*, vol. 5, pp. 93–5.

Organisation for Economic Co-operation and Development (1965). *Techniques of Economic Forecasting*. Paris: OECD.

Ormerod, P. A. (ed.) (1979a). *Economic Modelling*. London: Heinemann.

—— (1979b). 'The National Institute model of the UK economy: some current problems.' In Ormerod (1979a), pp. 115–40.

Osborn, D. R. (1979). 'National Institute gross output forecasts: a comparison with US performance.' *National Institute Economic Review*, No. 88, pp. 40–9.

—— and Teal, F. (1979). 'An assessment and comparison of two NIESR econometric model forecasts.' *National Institute Economic Review*, No. 88, pp. 50–62.

Pesaran, M. H. and Evans, R, A. (1984). 'Inflation, capital gains and U.K. personal savings: 1953–1981.' ECONOMIC JOURNAL, vol. 94, pp. 237–57.

Pindyck, R. S. and Rubinfeld, D. L. (1976). *Econometric Models and Economic Forecasts*. New York: McGraw-Hill. (Second edition, 1981.)

Prothero, D. L. and Wallis, K. F. (1976). 'Modelling macroeconomic time series (with discussion).' *Journal of the Royal Statistical Society*, Series A, vol. 139, pp. 468–500.

Renton, G. A. (ed.) (1975). *Modelling the Economy*. London: Heinemann.

Roy, A. D. (1970). 'Short-term forecasting for central economic management of the U.K. economy.' In *The Econometric Study of the United Kingdom* (ed. K. Hilton and D. F. Heathfield), pp. 463–73. London: Macmillan.

Samuelson, P. A. (1955). *Economics* (3rd edn). New York: McGraw-Hill.

Sargan, J. D. (1964). 'Wages and prices in the United Kingdom: a study in econometric methodology.' In *Econometric Analysis for National Economic Planning* (ed. P. E. Hart, G. Mills and J. K. Whitaker), pp. 22–54. London: Butterworth. Reprinted in *Econometrics and Quantitative Economics* (ed. D. F. Hendry and K. F. Wallis), pp. 275–314. Oxford: Basil Blackwell.

—— (1975). 'Discussion on misspecification.' In *Modelling the Economy* (ed. G. A. Renton), pp. 321–22. London: Heinemann.

Sims, C. A. (1967). 'Evaluating short-term macro-economic forecasts: the Dutch performance.' *Review of Economics and Statistics*, vol. 49, pp. 225–36.

Spencer, P. D. (1986). *Financial Innovation, Efficiency and Disequilibrium: Problems of Monetary Management in the United Kingdom 1971–1981*. Oxford: Clarendon Press.

—— and Mowl, C. J. (1978). 'A financial sector for the Treasury model.' Government Economic Service Working Paper No. 17.

Suits, D. B. (1962). 'Forecasting and analysis with an econometric model.' *American Economic Review*, vol. 52, pp. 104–32.

Surrey, M. J. C. (1982). 'Was the recession forecast?' *National Institute Economic Review*, No. 100, pp. 24–8.

Theil, H. (1958). *Economic Forecasts and Policy*. Amsterdam: North-Holland. (Second edition, 1961.)

—— (1966). *Applied Economic Forecasting*. Amsterdam: North-Holland.

Wallis, K. F. (1972). 'Testing for fourth order autocorrelation in quarterly regression equations.' *Econometrica*, vol. 40, pp. 617–36.

—— (1986). 'Forecasting with an econometric model: the "ragged edge" problem.' *Journal of Forecasting*, vol. 5, pp. 1–13.

—— (ed.), Andrews, M. J., Bell, D. N. F., Fisher, P. G. and Whitley, J. D. (1984). *Models of the UK Economy: A Review by the ESRC Macroeconomic Modelling Bureau*. Oxford: Oxford University Press.

—— (ed.), ——, ——, —— and —— (1985). *Models of the UK Economy: A Second Review by the ESRC Macroeconomic Modelling Bureau*. Oxford: Oxford University Press.

—— (ed.), —— Fisher, P. G., Longbottom, J. A. and Whitley, J. D. (1986). *Models of the UK Economy: A Third Review by the ESRC Macroeconomic Modelling Bureau*. Oxford: Oxford University Press.

—— (ed.), Fisher, P. G., Longbottom, J. A., Turner, D. S. and Whitley, J. D. (1987). *Models of the UK Economy: A Fourth Review by the ESRC Macroeconomic Modelling Bureau*. Oxford: Oxford University Press.

Worswick, G. D. N. and Budd, A. P. (1981). 'Factors underlying the recent recession.' Paper presented to the Panel of Academic Consultants, no. 15, Bank of England.

Zarnowitz, V. (1967). *An Appraisal of Short-Term Economic Forecasts*. National Bureau of Economic Research Occasional Paper, no. 104. New York: Columbia University Press.

—— (1979). 'An analysis of annual and multiperiod quarterly forecasts of aggregate income, output and the price level.' *Journal of Business*, vol. 52, pp. 1–33.

—— (1985). 'Rational expectations and macroeconomic forecasts.' *Journal of Business and Economic Statistics*, vol. 3, pp. 293–311.

—— and Lambros, L. A. (1987). 'Consensus and uncertainty in economic prediction.' *Journal of Political Economy*, vol. 95, pp. 591–621.