SCOPE AND LIMITATIONS OF STATISTICAL METHODS
FOR MACROECONOMIC FORECASTING

F.A.G. den Butter*

Summary

This essay confronts the science of forecasting with the art of forecasting and uses the crystal ball as a metaphor. It shows how the macroeconomic forecasts used in policy analysis originate as a mixture of scientific knowledge and judgement: economic methodology may provide the macroeconomic crystal gazer with a valuable crystal ball, but the final forecast relies heavily on the judgement and intuition of the crystal gazer. A comparison of the predictive performance of models and modellers is only sensible when it aims at investigating the different information contents of the forecasts and at answering the question why the forecasts were wrong.

Key Words & Phrases: economic forecasting, macroeconomic modelling, organisation of forecasting procedure.

1. Introduction

The crystal ball is not only an indispensable instrument for the initiated crystal gazer to reveal esoteric knowledge about the future. It is also a neat symbol for macroeconomic forecasting. In this metaphor the crystal ball represents the model or, more in general, the prediction method that the economic forecaster uses for shaping his (or her) view on the economic future. Up to date econometric methodology and/or economic theory may be applied for constructing the prediction method which is most suitable for the forecasting purposes.

However, the crystal ball only mirrors and colours reality according to its special shape and polishing, whereas it is the crystal gazer who has to interpret these mirror images so as to verbalize them to events which are likely to happen. In the same way the economic forecaster has to interpret the results that come out of the prediction method ("the

* Professor of Economics and head of the Applied Labour Economics Research Team (ALERT), Free University, PO Box 7161, 1007 MC Amsterdam, The Netherlands, tel. 020 5484609.

computer"), and should add his own judgement before making the forecasts public. Interpretation and judgement mark the difference between the art and the science of economic forecasting.

This essay discusses some aspects of forecasting for macro-economic policy purposes. It gives a personal view on the appropriate mixture of formal statistical forecasting methodology, economic theory and intuitive economic knowledge in order to get the best possible forecast of major macro-economic developments. The next section is on the science of forecasting. It shortly surveys the scientific methods available to the forecaster and considers their merits. Section 3 is on the art of forecasting. Whereas section 2 tells what toolbox the forecaster has for the construction of his crystal ball, section 3 looks at the required skill of the forecaster to use this forecasting device. Economic forecasts will never become (completely) true. Therefore, section 4 shows how we can evaluate the performance of various forecasters and their models, and how we can learn from the prediction errors. Finally section 5 concludes with a number of assertions on macroeconomic forecasting which may seem somewhat alluring for those not familiar with the art of forecasting.

2. Science of forecasting

The macroeconomic forecaster can, for the trade mark of his crystal ball, choose between *subjective* and *objective* methods. Examples of subjective methods are surveys and the Delphi-method. In surveys selected experts or randomly chosen individuals are asked about their (well documented or intuitive) opinion about future economic developments. Accordingly a survey prediction can be calculated as the (weighted) average of expert or individual opinions. An example is the survey conducted by national statistical agencies on consumer sentiments. These survey data are, for instance, used as leading cyclical indicators.

According to the Delphi-method views of experts on future economic developments are confronted with each other, so that after a first round of individual statements the experts can react on each other's opinion. However, this procedure, developed at the RAND Corporation is mainly used for technological forecasting (e.g. on the timing of a technological breakthrough, see Granger, 1980). Yet, in a less formal way this procedure is implicit in the organisation of forecasting rounds (see next section), where experts discuss preliminary forecasts during intradepartmental or intra-organisational meetings.

The scientific research of forecasting methodology focuses on the objective methods. Here we distinguish between *mechanical* methods and *causal* methods. Nowadays a large number

of mechanical methods for describing the data generating processes of macroeconomic time series is available, ranging from simple smoothing and naive extrapolation methods to sophisticated time series modelling. As examples of the latter we mention the vector ARMA-models, for which Box and Jenkins (1970; see also Jenkins and Alavi, 1981) have provided a unified specification, estimation and diagnostic checking method, the vector autoregressive (VAR) models which are a special branch of vector ARMA-models, and the structural time series models. While the specification of vector ARMA-models is restricted to one source of stochastic variation for each time series to be modelled, the structural time series models allow for more than one source of stochastic variation (for instance a stochastic trend combined with random shocks; see Harvey, 1989). Because the various sources of the stochastics of the structural time series models cannot be observed separately, we need Bayesian methods to estimate the parameters of these models. ARMA-models are sometimes regarded as reduced forms of these structural time series models. The common feature of all these mechanical methods is that they yield an algorithm for mechanical extrapolation of the time series. It does, however, not imply that specifying these models is completely mechanized. For instance, the specification method for Box and Jenkins-models contains some judgemental elements and, although it has been tried, cannot be mechanized completely.

The causal methods, as opposed to the mechanical methods, relate to the use of macroeconomic forecasting models, in which technical or behaviourial equations describe causal relationships which are derived from economic theory. This modelling of causal chains should be clearly distinguished from the casual causality relations as defined by Granger (1969) implicit in dynamic mechanical time series models. In the latter case causality is completely determined by the data and has very little to do with causal relationships imposed by economic theory (which, by the way, Sims, 1980, calls "incredible restrictions"). Causal economic models which are used for forecasting, range from very small one-equation models to large systems of equations with thousands of endogenous variables.

Macroeconomic development is, to a certain extent, unpredictable. An essential feature of objective economic forecasting methods is that they are designed to deal with limited predictive ability and that they try to separate as good as possible the signal from the noise in the information contained in economic time series. Moreover, in economic life the near future is usually more easily predictable than the far away future. Technically speaking, the signal to noise ratio becomes smaller the more periods are to be predicted ahead. Experience tells that in mechanical dynamic time series models the signal to noise ratio rapidly goes to zero when the number of periods to be predicted ahead increases. Hence, for long term forecasting these models boil down to very naive extrapolation models, such as no

change. In that case the variance of the prediction error is as large as the variance of the series itself.

A difference between mechanical forecasting methods and causal economic models is that the causal models require future values of exogenous variables as an input. Because generating input values can be mechanized by use of extrapolative formulas for the exogenous variables, the causal models can be used as a mechanical method as well. There is no essential difference between both methods in this respect.

The crucial difference, however, is that causal methods are very helpful for the art of forecasting, while mechanical methods are not, or only very little. In this respect I believe that, whenever there is room for improvement in macroeconomic forecasting, it will come through improving the art of forecasting. Therefore the development of the science of forecasting should be such that it supports the art of forecasting in the best possible way. That's why the scope for improving our forecasting performance in macroeconomics is rather limited when we concentrate on the development of more sophisticated mechanical methods only. Especially the marginal productivity of new and sophisticated estimation methods with respect to predictive performance of time series models seems very low. The improvement of model selection methods and specification tests will not add much to the predictive performance of these models either. A better economic theory can be instrumental to enlarge our understanding of the future and is essential for the art of forecasting. However, when better model specifications due to sophisticated economic theory are merely used for the production of mechanical forecasts, the gain in predictive accuracy will again be small.

A proper disaggregation of time series models can be useful in case the models for the disaggregate series are more stable and have a higher signal to noise ratio than the model for the aggregated series. Yet, I do not know many practical examples that show a considerable gain in predictive performance due to disaggregation. In my view the most promising approach for scientific help to the art of macroeconomic forecasting is the methodology of combining forecasts. This procedure defies the notion of many old-fashioned technocrats that there exists one superior forecasting model with accompanying estimation method. The combination methodology allows different forecasting models to coexist as these models may contain different information useful to the forecast. Bates and Granger (1969) prove that, when two forecasts contain different information, and hence when one forecasting method does not encompass the other one, a combination of both forecasts will always yield better predictions than one of the two forecasts separately. Combining forecasts enables to make optimal use of this different information. The methodology can be applied to combine various mechanical forecasting models, for exam-

ple based on both aggregated and disaggregated data (see Den Butter and Van de Gevel, 1989, where the gain from disaggregation appeared to be marginal indeed), but it can also be applied to combine mechanical forecasts with forecasts which are partly or fully based on judgement. In the latter case the combination method can be labelled as an objective method to combine subjective forecasts. This part of the science of forecasting may provide a good tool to be used in the art of forecasting, when it is applied to investigate differences of informational content of various forecasts (e.g. as a part of a post-mortem analysis, see section 4). However, no much gain in predictive performance will be obtained when utilizing the combination methods in a sheer mechanical way only.

3. Art of forecasting

The first lines of this essay argue that the crystal ball itself is not a good forecaster, but that we need the crystal gazer to interpret the reflections of the crystal ball. Similarly mechanical methods on their own are not very useful for forecasting: no professional macroeconomic forecaster will publish a forecast just as it first comes out of the computer. He definitely wants to comment on the forecast and tell the story to go with it. As a matter of fact macroeconomic forecasts are never published just on their own, but are always accompanied by a text which explains how they are obtained and under what conditions they may come true (or be defied). That's why mechanically generated forecasts alone are not very useful to the art of macroeconomic forecasting: they do not tell a story. They just come from a black box where - unlike in the crystal ball, I suppose - there is nothing to see.

On the other hand, causal forecasting models make macroeconomic story-telling possible and enable to include judgement into the forecast. As a matter of fact, building the causal policy model already requires artistic skills and a sound economic judgment. Apart from the design of the model and the specification of the behaviourial equations, judgement is introduced technically into the models by so called con(stant) adjustments or add-factors (information factors) in the behaviourial equations of the model. By means of the add-factor the forecaster judges how much the future course of a specific macroeconomic variable will deviate from normal behaviour. In this way informal and intuitive knowledge, or formal knowledge which is not contained in the model, can be consistently combined with the model's description of past behaviour. Such add-factors may, for instance, relate to the additional demand for television-sets during the Olympic Games or to the intuitive feeling that the negative wealth effects of a stock market crash will not lead to the consequent slowdown of macroeconomic activity. According to Klein (1979) these constant adjustments or add-factors were already used in 1953 for preparing the forecast of the

Klein-Goldberger model. Haitovsky and Treyz (1972) consider the possibility of incorporating judgemental information in a systematical way in the forecast by means of add-factors as the main advantage of causal macroeconomic models over mechanical forecasting methods.

The use of add-factors and hence of a macroeconomic model, forces the forecaster to make his beliefs about atypical future developments explicit in a quantitative way. Note that, for obvious reasons, add-factors are not allowed in definition equations. If, for instance, a forecaster believes that his model yields too low a prediction of national income as the sum of a number of demand categories, he is obliged to indicate what behaviourial equation he wants to add-factor in order to increase his forecast of national income. Moreover he should be aware of the fact that in a simultaneous equation structural economic model add-factoring of one equation will affect the outcomes of all other endogenous variables.

This brings us to the main rationale for the use of a forecasting model: it makes the forecasts consistent and it makes the assumptions in the forecasting-process explicit. The size of the add-factors should be plausible. In this respect it is helpful to compute the add-factors that equate the dynamic model-simulations over the recent past to their realizations. These endogenously determined add-factors indicate how big the add-factors should have been in the recent past in order to get a perfect prediction.

If skilfully applied the mechanical forecasting methods, of course, have their own role in this art of macroeconomic forecasting. For instance, short term mechanical predictions may provide benchmarks for add-factors (see Corrado and Greene, 1988). Mechanical methods are also useful for extrapolating those exogenous variables which, because of their sheer number or for other reasons, cannot be determined by judgement. Yet, the use of mechanical forecasting methods on their own, without interference of the forecaster, has its main scope outside macroeconomic forecasting. For instance, its scope in economics is in marketing research or in financial analysis where, on the very short term, developments on the stock market and the foreign exchange market leave no room for a formal procedure to combine the mechanical forecasts with judgment.

Some international institutions (such as OECD, IMF, World Bank, EC) and commercial companies (DRI, Wharton) are specialized in making a consistent forecast of the world economy. It is obvious that the forecasting of so many interdependent macroeconomic variables is an iterative procedure and involves a number of people who are well trained in the art of forecasting. It also requires a good organisation of the forecasting process. As an example of such a professional forecasting exercise we consider the OECD-procedure for the forecasts published in the *Economic Outlook*. This example is selected because it is well

documented (see Llewellyn, Samuelson and Porter, 1985, chapter 10). The procedure can be summarized as follows:

- The Economic Outlook is a semi-annual publication and contains forecasts for the world over the period of the next one and a half year
- The forecasting procedure is centred around the Secretariat's world model INTER-LINK
- Apart from a small central group responsible for maintaining the model and for organising the regular forecasting round, OECD forecasts are made by a number of country and subject specialists which each devote only a part of their time to that activity
- Therefore the forecasting round should be well organised and has a fixed time table with the following stages:

Stage 1. First climate run

- review of recent data
- updated set of add-factor adjustments
- new exogenous variables, mainly newly announced policy measures
- this first climate run is merely a mechanical update of the previous set of forecasts
- the results of the first climate run are circulated only to those actively participating in the construction of the projections

Stage 2. Second climate run

- inputs are prepared by country desk officers and subject specialists active in the forecasting round
- revised historical data are entered, together with any new information on policy and on the exogenous variables
- the exercise of judgement is limited at this stage
- re-based trade data by the trade specialists
- detailed forecasts are entered for the United States and Germany
- the climate meeting is held two weeks after the beginning of the round in order to inform each country desk officer what sort of international climate his or her country seems likely to be facing

Stage 3. First set of detailed forecasts

- input of forecasts by all country specialists

- discussions on consistency of forecasts; for instance the country specialists may be struck by a collective pessimism or optimism on the balance of payment position of the countries they look at; in that case the balance of payment specialists will see to it that the overall balance of payment in the world is in equilibrium; this consistency urges the country specialists to revise their add-factors of import and export equations.

- input of judgement
- review of various features of the forecasts at intradepartmental meetings (about 4 weeks into the round)

Stage 4. Submission of forecasts to experts from OECD member governments

- revised forecasts after intradepartmental meeting are submitted to experts from OECD governments
- meeting of national experts
- meeting of the Economic Policy Comity (top officials from Economics and Finance Ministries and central banks)
- further modification of the forecasts in the light of comments made either by the experts or in the course of the Economic Policy Comity

Stage 5. Publication of OECD Economic Outlook

- a final set of internationally consistent forecasts for all OECD member countries

Stage 5 concludes the forecasting round at OECD.

In the Netherlands model-based macroeconomic policy analysis and forecasting has a long and rich history. This is due to Tinbergen's pioneering work, and to the legal responsibilities of the Central Planning Bureau (CPB) which find their basis in the Law of the 21st of April 1947 on the Preparation of the Central Economic Plan. According to this law the Central Economic Plan "should contain data on the future size and development of the price level, on national income and its components, on national spending, and on all other macroeconomic quantities which a good co-ordination of the economic social and financial policy demands". This law explicitly charges the Central Planning Bureau to make forecasts. (See Fase, 1986, and Van den Beld, 1979). In spite of some recent proliferation of model-based macroeconomic forecasting in the Netherlands, the CPB has remained by far the most important and prestigious producer of model-based forecasts in the Netherlands. Three annual publications are important in this respect. Each year in September, at the presentation of the new government budget, the CPB publishes its "Macro-Economische

Verkenning" (MEV, Macro-economic Outlook), which contains forecasts for the following year. Then, at the end of April of the following year, the forecasts for that year are updated in the "Centraal Economisch Plan" (CEP, Central Economic Plan). Moreover, in addition to the Central Economic Plan, the Central Planning Bureau has started in 1989 to publish "Het Economisch Beeld" (EB, Economic Prospects) which contains a first analysis of the economic development and forecasts for the next year. The forecasts are again updated in December and June in the so called "Halfjaarlijkse Tussenrapportage over de Nederlandse Economie" (semi-annual intermediate report on the Dutch economy).

The foregoing surveys the CPB's calender of publication of short-term forecasts. Additionally the CPB publishes medium-term forecasts, for the next five years. The CPB used to publish these medium-term forecasts each five years but recently these forecasts are updated more frequently (see Van der Lem and Zalm, 1989). The forecasting rounds for the CPB's short term and medium term macroeconomic forecasts proceed along similar lines as those of the OECD. Don and Van den Berg (1990) provide an interesting survey of the forecasting procedure and the calender of activities of the CPB during the period November 1988 - October 1989. They report that the decision to publish the April forecasts for the following year in a separate publication (EB, Economic Prospects) and not in the Central Economic Plan had a legal and political background. When published in the CEP the Government would be legally bound to the CPB's estimates and views, which was regarded as premature at that stage.

This summary of the forecasting experience indicates that progress in the art of forecasting is mainly made at organisations such as the OECD and CPB, and not so much at the Academia.

The tension between the science of forecasting and the art of forecasting is illustrated in the following quotation of Samuelson (1975):

"When Robert Adams wrote a MIT-thesis on the accuracy of different forecasting methods, he found that 'being Sumner Slichter' was apparently one of the best methods known at that time. This was a scientific fact, but a *sad* scientific fact. For Slichter could not and did not pass on his art to an assistant or to a new generation of economists. It died with him, if indeed it did not slightly predecease him. What he hoped to get by scientific breakthrough is a way of substituting for men of genius, men of talent and even just run-of-the-mill men. That is the sense in which science is public, reproducible knowledge."

Hence Samuelson's main concern with forecasting is that a forecasting artist may outperform a forecasting scientist, while the art of forecasting is non-reproducible.

Macroeconomic forecasting is certainly not the only and probably not even the main reason for building macroeconomic models. Policy analysis also uses these models for *scenario analysis* and for *policy simulations*. Scenario analysis is very closely connected with macroeconomic forecasting. A scenario is a model-based projection of the future under specific assumptions on exogenous variables, especially policy variables. The main rationale for this type of analysis is to show what will happen in case specific policy measures are taken. Or, more dramatically, "pessimistic" scenarios may show, for instance in the case of environmental policy, what will happen if policy measures are not taken.

As a matter of fact a forecast can be viewed as the scenario with the most probable assumptions on exogenous variables. If there is great uncertainty on future policy or on exogenous developments (world trade, oil prices, threat of war etc.) the forecasting agency is to publish a number of scenarios instead of one central forecast. The number of published scenarios should be even since otherwise the public may choose the middle scenario as the actual forecast. Moreover, forecasters should also label their forecasts as scenarios when the forecast shows unwarranted future developments. In that case the forecaster wants his forecast to be self-defying. We can distinguish between three goals for the scenario analysis, namely:
- to advertise a desired development
- to prevent an undesired development
- to cure an undesired development.
For symmetry reasons a fourth goal of scenario analysis would be to advocate the continuation of a desired development, but in practice such scenario's never emerge.

There are two different construction methods in scenario analysis (see Jungermann, 1985). The *exploratory scenario* is based on forward inference and looks at what happens if a specific exogenous event occurs, or policy measure is taken (or not). The *anticipatory scenario* is based on backward inference and indicates how a fixed policy goal can be reached.

Policy simulation is the most common and oldest purpose of use of macroeconomic models. Such simulation gives the model's response to an autonomous shock in a policy (or other exogenous) variable. The result of policy simulations are usually presented in the format of impulse-response tables, which show the difference between the central or baseline projection and the alternative or impulse projection. When Tinbergen presented his macroeconomic model for the Netherlands in 1936, which was also the first macroeconomic model of this type in the world, he already used the model for policy simulations and calculated such impulse-response effects, which in Dutch are known under the label of "spoorboekjes" (railways timetables).

4. Evaluation of predictive performance

The weight that macroeconomic forecasts will carry in the decision-making process of economic agents and in the design of macroeconomic policy, depends much on their reputation. Unlike the qualitative and often enigmatic forecasts made by crystal gazers, the performance of quantitative macroeconomic forecasts can be evaluated by comparing the predictions with the realizations. However, such confrontation of forecasts with realizations is more complicated than it looks at first instance. One reason is that the forecasts may relate to different definitions of variables, because the forecasts were made at different moments of time. Another reason is that even realizations are not always unambiguous: definitive National Accounts' data are published with a lag of some years, and the Netherlands has witnessed recently a remarkable disagreement between the Central Bureau of Statistics and the Chambers of Commerce on the national investment data. The disagreement was not only on the size of the annual growth of investment in the past year, but even on the sign of it.

In spite of these obstacles, there is ample literature on the comparison of the predictive performance of macroeconomic models using a simple root mean square prediction error, Theil's inequality coefficient, or a more sophisticated measure as a criterion to pick the winner. One of the early studies in this field is Nelson's (1972) demonstration that simple mechanical time series models could produce more accurate one quarter ahead forecasts than a large scale macroeconometric model. This superiority of time series models over causal models is somewhat confirmed by, for instance, our own forecasting exercise with money demand functions (see Den Butter and Fase, 1980). However, more recent and much better documented evidence by McNees, who is nowadays commonly regarded in the United States as the extreme judge on the predictive performance of macroeconomic models, challenges these conclusions (see, for instance, McNees, 1990).

A related question is whether forecasters have improved. To this end Westerhout (1990) recently investigated the predictive performance of the CPB and concludes that its performance has improved slightly indeed. A surprising result is that, according to Westerhout, the medium-term predictions of the CPB appear to be more reliable than the short-term predictions (see Ketellapper and Scholtens, 1984; Van der Leeuw, 1984; and Van Schaaijk, 1986, for other studies of the CPB's predictive performance). Incidentally, the finding that the CPB's predictive performance has improved does not necessarily imply that its predictive ability did improve to the same extent. It can also be true that macroeconomic forecasting has become easier, so that the same predictive ability yields better forecasts. However, probably macroeconomic forecasting has become more difficult, but as the CPB has improved even more, its forecasts slightly gained accuracy.

Yet, for a number of reasons this game of picking the winners, and other comparative studies of predictive performance on the basis of average prediction errors, does not make

much sense. Firstly it is very difficult to monitor that all predictions are made at the same moment, and that it is not informational advantage which has resulted in a better forecast. Secondly, the prediction errors relate to various macroeconomic variables and to a different number of periods to be forecasted ahead, so that a specific model or forecasting method never finishes as winner on all points. Thirdly, the differences between the forecast errors are usually small as compared to the forecasting intervals, so that a better forecast in terms of a smaller prediction error may just be good luck and not a systematic result. A fourth reason is that the forecasts may have had the character of self-defying or self-fulfilling prophecy. An illuminating example for this is the conflict that Prime Minister Lubbers had with the former CPB director on the forecasts for 1988, which Lubbers thought were too pessimistic. However, the CPB was very reluctant to adjust the forecast according to the Prime Minister's wishes and suggested that Lubbers was hoping for self-fulfilling prophecies, which would enhance the status of the cabinet[1].

But finally and most importantly, these comparisons do not learn us *why* the prediction errors are made and *why* some models or methods are better than others. Therefore, they are not very helpful in improving the forecasting performance. For that reason, it is much more useful to compare the performance of models in such a way that we can learn about strong and weak points of the forecasting procedure. In this vein Fair and Shiller (1989, 1990) compare the informational content in forecasts from economic models and in forecasts combining economic models with judgement. Such investigation does not only indicate to what extent forecasters use different information, or all relevant information available, but it also gives a clue for combining the forecasts in order to improve them. In the same vein McNees (1990) investigates the extent to which judgement is helpful to improve mechanically generated forecasts. He concludes that the historical records suggest that judgemental adjustment improves the forecasts, despite instances of success of mechanically generated forecasts. Moreover he looks at whether forecasters who combine their forecasts with judgement overadjust or underadjust. In other words, whether they put to much or to little trust in the mechanically generated forecast from their models. McNees finds a slight overadjustment. The message therefore is that forecasters should adjust their models using judgement, but that they should be very careful about it. It is, according to McNees, a mistake to accept adjustments that are made at face value, especially when the adjustments appear without any explanation of the reasoning behind them.

For that reason an extensive so called *post-mortem* analysis of prediction errors should be an essential part of the forecasting procedure. Prediction errors may have various causes,

---

[1] Forecasts of official forecasting agencies, like the CPB, are bound to be based on announced policy measures, even when it is highly improbable that these policy measures will be effectuated. For that reason the predictive performance of 'unofficial' forecasters may be better.

and in order to improve his forecast the forecaster should know which cause has led to his failure in the past. As possible causes we mention

- model specification error
- wrong coefficient values in the behaviourial equations, for instance because of an inadequate estimation technique or because of lack of data
- prediction errors of exogenous variables, for instance of the policy stance
- errors in add-factors, and hence a wrong judgemental adjustment
- the use of preliminary data on lagged endogenous variables, and on exogenous variables, which are revised afterwards[1].

However, in practice it may be very difficult to discriminate in a post-mortem analysis between these various sources of forecasting errors. Even with the benefit of hindsight it is difficult to determine formally whether a prediction error has been caused by the wrong specification of economic behaviour or by a shift in preferences which should have been met by an add-factor. Moreover, even a perfect forecast does not give us good comfort because it can be the result of two large forecasting errors that incidentally compensate each other completely.

A post-mortem analysis is not only useful to teach us the reason of our failures but may also enhance our understanding of economic developments in the recent past. By running a dynamic simulation with the model over the past with realized values for the exogenous and initial lagged endogenous variables, the ex post-prediction errors provide an indication of what developments in the recent past can be considered as atypical. Although Wallis (1989) reports about an operational method for decomposing forecasting errors, it is remarkable that up to now in model-based policy analysis no unified methodology exists for the classification of the different causes of atypical developments. The design of such methodology is desirable. Science can be supportive to the art of forecasting if efforts would be spent to the scientific foundation of the post-mortem analysis. Yet, problems of interpretation and classification will remain, so that judgement will always play a role in a post-mortem analysis as well.

5. Conclusions

This essay confronts the science of forecasting with the art of forecasting. It focuses on macroeconomic forecasts and shows how these forecasts used in policy analysis originate as a mixture of scientific knowledge and judgement. The following statements summarize the views expressed in this essay:

---

[1] See Gallo and Don (1991), and Van Vlimmeren, Don and Okker (1991) for the relationship between forecast errors and data uncertainty due to revisions.

1. When fully mechanized forecasting methods are used as such, they are not very serviceable to the professional macroeconomic forecaster: it resembles a crystal ball without the crystal gazer to interpret the mirror image of reality. However, mechanical methods are useful as an auxiliary tool for the macroeconomic forecaster, like the crystal ball is for crystallomancy.

2. Macroeconomic forecasts made by professionals always contain judgement and are never solely based on the outcomes of the models.

3. Further sophistication of econometric methods will not contribute much to the improvement of the forecasts. If skilfully applied, today's econometric toolbox suffices to avert avoidable specification errors.

4. The scope for economic theory in improving the forecasts lies not so much in providing better specifications for the models, but in gaining a better understanding of the working of the economy so that founded judgement can be incorporated into the forecasts.

5. A substantial improvement of the macroeconomic forecasts is not to be expected as the predictive accuracy of the economic future seems bounded by an upper limit; better data from statistical sources or more sophisticated scientific techniques will not be raise this limit very much; Samuelson (1975) feels as if there is a Heisenberg indeterminacy principle dogging us, which will limit the asymptotic accuracy of forecasting we shall be able to attain.

6. Model-based economic policy analysis is a much broader profession than just making forecasts; scenario analysis and policy simulations can be as fruitful for policy design than the provision of good forecasts.

7. Competitions of forecasting performance between models or modellers, based on a comparison of prediction errors only, does not provide much relevant information for improvement of the forecasts; scarce research resources are much better allocated by analyzing and comparing the information content of the various forecasts.

8. Model-based policy analysis has always played an important part in the design of macroeconomic policy in the Netherlands; the recent proliferation of models which are used for forecasting, brings about the need for comparative studies on the informational content of these models and consequently on the optimal combination of these forecasts.

9. A post-mortem analysis, which shows why forecasts deviate from their realizations, is essential for improving the forecasts (if possible); a better methodology and more data on the forecasts are needed, especially in order to investigate whether the judgemental part or the mechanical part of the forecast is most responsible for the forecast error.

10. The main aim of model-based policy analysis is to indicate what policy measures can improve welfare, or prevent disaster; the provision of accurate forecasts should be auxiliary to this aim.

References

Bates, J.M. and C.W.J. Granger, 1969, The combination of forecasts, Operational Research Quarterly, 20, pp. 451-468.

Beld, C.A. van den, 1979, Het Centraal Planbureau, zijn invloed, zijn macht en zijn onmacht, in: W.M. van den Goorbergh, Th.C.M. van de Klundert and A.H.J. Kolnaar (eds.), Over Macht en Wet in het Economisch Gebeuren (Stenfert Kroese, Leiden), pp. 49-75.

Berg, P.M.C.J. van den, 1986, De betrouwbaarheid van macro-economische voorspellingen - een bijdrage van het CPB, Economisch Statistische Berichten, 71, pp. 1153-1157.

Box, G.E.P. and G.M. Jenkins, 1970, Time Series Analysis, (Holden Day, San Francisco).

Butter, F.A.G. den, and M.M.G. Fase, 1980, Forecasting the money stock in 6 EEC countries, De Nederlandsche Bank N.V., Research Report nr. 8004.

Butter, F.A.G. den, and F.J.J.S. van de Gevel, 1989, Prediction of the Netherlands' money stock, De Economist, 137, pp. 173-201.

Corrado, C. and M. Greene, 1988, Reducing uncertainty in short-term projection: linkage of monthly and quarterly models, Journal of Forecasting, 7, pp. 77-102.

Don, F.J.H. and P.J.C.M. van den Berg, 1990, The Central Planning Bureau of the Netherlands; Its role in the preparation of economic policy, mimeo, November 1990 (CPB, The Hague).

Fair, R.C. and R.J. Shiller, 1989, The informational content of ex ante forecasts, Review of Economics and Statistics, 71, pp. 325-331.

Fair, R.C. and R.J. Shiller, 1990, Comparing information in forecasts from econometric models, American Economic Review, 80, pp. 375-389.

Fase, M.M.G., 1986, De betrouwbaarheid van macro-economische voorspellingen: een bijdrage van de Nederlandsche Bank, Economisch Statistische Berichten, 71, pp. 1148-1152.

Gallo, G.M. and F.J.H. Don, 1991, Forecast uncertainty due to unreliable data, Economic and Financial Computing, 1, pp. 49-69.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross spectral models, Econometrica, 37, pp. 424-438.

20

Granger, C.W.J., Forecasting in Business and Economics (Academic Press, Orlando).

Haitovsky Y. and G. Treyz, 1972, Forecasts with quarterly macroeconomic models, equation adjustments and benchmark predictions: the US experience, Review of Economics and Statistics, 54, pp. 317-325.

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

Jenkins, G.M. and A.S. Alavi, 1981, Some aspects of modelling and forecasting multivariate time series, Journal of Time Series Analysis, 2, pp. 1-47.

Jungermann, H., 1985, Inferential processes in the construction of scenarios, Journal of Forecasting, 4, pp. 321-327.

Ketellapper, R.H. and H.W. Scholtens, 1984, De betrouwbaarheid van macro-economische prognoses op middellange termijn, Maandschrift Economie, 48, pp. 114-122.

Klein, L.R., 1979, Use of econometric models in the policy process, in P. Ormerod (ed.), Economic Modelling (Heinemann, London), pp. 309-329.

Leeuw, J.L. van der, 1984, De toekomst in retrospectief - over de voorspelkwaliteit van de Macro-Economische Verkenning, Economisch Statistische Berichten, 69, pp. 160-164.

Lem, J.T.N. van der, and G. Zalm, 1989, De Nederlandse economie in de periode 1989-1994, Maandschrift Economie, 53, pp. 411-429.

Llewellyn, G.E.J., S.J. Porter and L.W. Samuelson, 1985, Economic Forecasting and Policy: the International Dimension (Routledge and Kegan Paul, Londen).

McNees, S.K., 1990, Man vs. model? The role of judgement in forecasting, New England Economic Review, July/August 1990, 41-52.

Nelson, Ch.R., 1972, The prediction performance of the FRB-MIT-PENN model of the US economy, American Economic Review, 62, pp. 902-917.

Samuelson, P.A., 1975, The art and science of macromodels over 50 years, in G. Fromm and L.R. Klein (eds.), The Brookings Model, Perspective and Recent Developments (North-Holland, Amsterdam), pp. 3-10.

Schaaijk, M. van, 1986, Raming en realisatie van CPB-prognoses, Economisch Statistische Berichten, 71, pp. 727-729.

Sims, C.A., 1980, Macroeconomics and reality, Econometrica, pp. 1-48.

Vlimmeren, J.C.G. van, F. J. H. Don and V.R. Okker, 1991, Modelling data uncertainty due to revisions, CPB Research Memorandum, No. 75 (Central Planning Bureau, The Hague).

Wallis, K.F., 1989, Macroeconomic forecasting: a survey, Economic Journal, 99, pp. 28-61.

Westerhout, E.W.M.T., 1990, De kwaliteit van CPB-voorspellingen, Economisch Statistische Berichten, 75, pp. 840-843.