

Research Memorandum

No 162

**Forecasting in macroeconomics:
a practitioner's view**

F.J.H. Don

CPB Netherlands Bureau for Economic Policy Analysis, The Hague,
February 2000

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, The Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50

ISBN 90 5833 033 8

The responsibility for the contents of this Research Memorandum remains with
the author(s)

CONTENTS	page
1 Introduction	5
2 Why do we forecast?	6
3 What requirements for a good forecast?	9
4 How to forecast?	12
5 Communicating with the client	18
6 Some final remarks	26
References	27
Abstract	30

1. Introduction*

This paper offers some thoughts on forecasting in macroeconomics.

Section 2 starts by asking why we make forecasts. Forecasting is explicitly linked to decision making under uncertainty. We warn against confusing conditional and unconditional forecasts. As loss functions are usually unknown, and certainty equivalence is unlikely to prevail, a forecaster must communicate not just his forecast but also enable the client to form his opinion on the associated uncertainty.

Section 3 attacks the relevance of common statistical criteria for forecast quality, and promotes three non-statistical criteria: logical coherence, economic coherence, and stability.

Against this background, alternative forecasting techniques are discussed in section 4. This section also touches on the value of market expectations and on the matter of self-fulfilling or self-denying prophecies.

Returning to the purpose of forecasting, section 5 discusses communication with the client. Communicating statistical properties tends to be very difficult, because of implicit or explicit conditioning on other variables, on model parameters or specification, and sometimes on subjective probability assessments. For communication with the client, uncertainty variants and alternative scenarios appear adequate. Still policy makers have a strong tendency to choose a single scenario and use it on the assumption of certainty equivalence, rather than use different scenarios as a test on the robustness and flexibility of policy choices.

Section 6 offers some final remarks.

The views given in this paper are based on personal experience at the CPB Netherlands Bureau for Economic Policy Analysis, where the author has worked in various positions for some 20 years; since 1994 as director.

*Earlier versions of this paper were presented at the (EC)² conference on "Forecasting in Econometrics" in Stockholm, December 1998, and at the Seminar in honour of professor Kenneth F. Wallis, University of Groningen, November 1999.

The author is grateful to Frank den Butter, Martin Fase, David Hendry, Ken Wallis and several colleagues at CPB for their comments on an earlier version, and to Henk Kranendonk for research assistance. The usual disclaimer applies.

2. Why do we forecast?

"If one knew everything in advance, then one could travel the world on twopence"
(Dutch proverb, my translation)

Why do we make economic forecasts? Different forecasters may have different purposes. Indeed, forecasts may differ just because they serve a different purpose. In some way or other, however, the purpose of any forecast is to provide guidance in decision making. The type of decision may vary from a simple bet on some future outcome¹, to a complex government or business policy choice.

The criteria for good forecasting should be derived from the purpose the forecast is intended to serve. The natural analytical framework² is that of decision theory, which starts from a loss function $L(d, s)$ specifying what damage is done in future state s if we now choose to take decision d . The decision maker faces uncertainty over s , and this is where the forecaster is supposed to help.

If the forecaster had perfect foresight, then the decision maker would just have to find the decision d that minimizes L for the forecasted value of s . However, if perfect foresight is not available, then any forecast for s necessarily carries uncertainty and the decision maker will have to apply a decision rule to determine d . A common rule is to minimize expected loss, but the decision maker may also want to minimize maximum possible loss, or make some other choice.

Most decision rules will require a distribution for s , not a point forecast. However, there is a strong result known as certainty equivalence: for a quadratic loss function³, minimum expected loss is obtained for the value of d that minimizes $L(d, E_s)$, irrespective of the distribution of s . And for scalar s the forecast with the smallest mean squared forecast error will provide the best result (*i.e.* the smallest value of expected

¹ The decision on what forecast to put out may itself be the relevant decision, when the forecasting institute is primarily interested in publicity. See Laster et al. (1999) for an analysis and empirical evidence.

² In what follows d is a vector in decision space D , s and f are vectors in state space S , and L is a scalar function on $D \times S$.

³ What we need here is that L is quadratic on S , not necessarily on D . It suffices that L admits a representation $L(d, s) = s'As + s'b(d) + c(d)$, where A is a constant matrix and $b(d)$, $c(d)$ are vector functions on D . See Theil (1958, section 8.2).

loss). This makes E_s , the expected value of s , a natural candidate for "the" forecast of s . It also suggests to use the mean squared forecast error as the main criterion for forecast evaluation.

Strong as the certainty equivalence result may be, in many cases the loss function is unlikely to be quadratic (or even symmetric). Often it is very hard to specify a realistic loss function. Indeed, many economic forecasters do not cater to a single well defined decision maker, but rather make their forecasts available to anyone who wants to use them. From a decision theory point of view, they would do best providing distributions of s rather than point forecasts. In section 5 we will argue that this is not a realistic requirement, and suggest some ways to handle the communication problem.⁴

In some cases, including most government economic policy problems, the decision that is taken will also affect the future outcome. Indeed, often the whole point of the exercise is to generate an economic policy decision d which will lead to a preferred result for the future state s . If the decision maker cannot affect the future state, he is best served with an unconditional⁵ forecast. But if his decision can affect s , he can only be served by conditional forecasts, indicating the likely outcome conditional on the policy decision that will be implemented. An unconditional forecast would necessarily imply an assumption on the behaviour of the decision maker, and thus obscure the decision problem.

If s is a function of d , we can still apply the certainty equivalence principle provided the random element in s enters additively, i.e. $s(d) = E_s(d) + u$, with u independent of d . Then, under quadratic loss, the minimum expected loss decision is the same as the minimum loss decision obtained after substituting of $E_s(d)$ for s . See Theil (1958) for derivations and Chow (1975) for multiperiod certainty equivalence in dynamic systems.

The difference between conditional and unconditional forecasts may appear obvious, yet it is often ignored. In fact, most forecasts are conditional, even if they should not be. And in forecast evaluation one usually assumes that the forecasts were unconditional, even if they should not have been. Some examples will help to make this clear.

First, consider the forecasts made in the Summer of 1998 by most macroeconomic forecasters (including CPB). They tended to assume (rather than forecast) that the Asian

⁴ For an early discussion of this communication problem, see Hildreth (1963).

⁵ Of course it should be conditional on all available information, but not on information which is not available, like the decision that will be implemented by some other agent in the system.

crisis would not further deepen or spread to so far unaffected countries. Based on this assumption, they produced conditional forecasts. Yet they hardly gave a clue, if any, how their clients could ensure that the stated condition would indeed prevail. To their credit, some did include alternative scenarios, thus using conditional forecasts as a way of communicating major uncertainties.

Similarly, many macroeconomic forecasters are used to assuming constant exchange rates as of some particular cut-off date, although they would not necessarily consider that assumption to be the best in terms of economic coherence with the rest of the forecast. By refraining from giving an exchange rate forecast, they make their entire forecast conditional (on the exchange rate assumption), while most clients will take it to be unconditional.

As a final example, consider the Dutch stock market analyst who was asked to account for his sizable forecast error over the third quarter of 1998 and explained that "the situation had changed".

In forecast evaluation, the usual procedure is to compare published forecasts with realizations. Such a simple comparison is fair only if the forecasts were unconditional. For a proper evaluation of a conditional forecast, one should check whether the condition has indeed materialized. If it has not, the forecast error should be corrected for the effect of the difference between conditional value and true outcome. Such a correction may be based on the same econometric model, if any, that was used to produce the forecast. See *e.g.* Van den Berg (1986) for an example. But that is a tedious job, which is often skipped in forecast evaluation.

However, if a forecaster had no good reason to condition his forecast, and effectively used the conditions to hide behind when things turned out differently, then it may be fair enough to evaluate his forecasts as if they had been unconditional.

Forecasts produced for the government will normally be conditioned on explicit policy assumptions that reflect the current state of government decision making. Any other policy assumption in the baseline forecast would only confuse the decision making process, which will take the baseline as its starting point. However, a forecasting institute catering to clients outside the government may well draw up an unconditional forecast which, explicitly or implicitly, includes a forecast for policy variables different from the values implied by the current state of decision making. On a simple (unconditional) evaluation of forecasts, one would expect the unconditional forecast to be systematically better than the conditional one.⁶

⁶ I am not aware of any such systematic difference in forecast evaluation studies.

3. What requirements for a good forecast?

*"Many critics, no defenders
Weathermen have two regrets
When they hit, no one remembers
When they miss, no one forgets"*
(anonymous)

Because forecasters cater to many clients simultaneously, and loss functions are generally unknown, we need some guidance as to what we should require from a good forecast.

A common statistical requirement for a forecast of a scalar variable is that the forecast error⁷ has zero mean and minimal variance. Alternatively, one might want it to have zero mode or zero median in stead of zero mean. If small sample properties are hard to assess, one may resort to asymptotic equivalents (like statistical consistency).

Statistical criteria require some knowledge of the true distribution of s . As a rule, this distribution is known only subject to a number of untested (model) assumptions relating to the economic system. Even in a Bayesian approach, some structure is imposed before distributional statements can be made. Hence, any statistical requirement can be valid only in the context of a set of untested assumptions which are usually unknown to and certainly difficult to judge for the decision maker.

Two requirements are often met in practice, although they tend to conflict with the usual statistical criteria just mentioned: coherence and stability. Both relate to the credibility of the forecast: how is the client to understand the forecast in terms of (his perception of) the underlying (economic) system. Of course, some clients may be quite happy with accepting forecasts as coming from some black box, *e.g.* because it has worked so beautifully in the past or because it comes from a reputable institute. But the more sophisticated clients, like those we are used to serving in macroeconomics, will require some credible story or explanation of why the forecast is as it is. They may use the story to develop their own judgement on the associated uncertainty.

⁷ If f serves as the forecast for s , the forecast error is defined as $s-f$.

Logical coherence⁸ requires that a forecast satisfies the (accounting) identities that any possible realization is bound to satisfy. Many such identities are linear, like the Keynesian income-expenditure identity, or the definition of the budget deficit as the difference between government expenditures and government receipts. But many others are nonlinear, like the multiplicative relation between growth rates and levels, or that between real values, corresponding price indexes and nominal values. While linear identities are compatible with statistically unbiased forecasts, nonlinear identities immediately pose a problem: with the exception of the one period ahead forecast, unbiased forecasts for levels must generally differ from levels computed from unbiased forecasts for growth rates. Note that logical coherence is preserved for all monotonous transformations if medians are used rather than means.

Economic coherence goes beyond identities: it also requires the forecast to make economic sense and obey observed regularities in the data. Of course, econometric models try to capture both the relevant economic theory and past data patterns, and they provide the major tool for ensuring both logical and economic coherence. Indeed, an econometric model may well be used as the definition of what is considered coherent, and provides the natural reference for the proper stories and explanations to communicate with the client. Unlike means or medians, the model forecast obtained by putting all error terms at zero will always satisfy both coherence criteria. This probably helps explain why the latter is still the dominant mode for preparing model forecasts, although means and medians can nowadays be obtained without much effort from Monte Carlo runs.

Are logical and economic coherence sensible criteria, or should we ignore them in favour of the usual statistical criteria? Let me repeat that any statistical criterion can be valid only in the context of a set of untested assumptions which are usually unknown to and certainly difficult to judge for the decision maker. In contrast, logical and economic coherence are easy to understand, and very helpful in communicating the relevant information to the decision maker. Indeed, it is often considered to be the major advantage of an econometric model that it provides a coherent framework for bringing all relevant information together, see *e.g.* Smith (1998). Therefore I think we should put logical and economic coherence above satisfying some statistical criteria in a constructed probability space.

⁸ Some authors use the term admissibility to refer to this property.

The stability criterion is even less natural from a statistical point of view. Clients will tend to compare a new forecast with previous forecasts, and the forecaster will have to explain why his forecast has changed. Thus, stability can be viewed as a form of economic coherence between successive forecasts. Britton et al. (1998) observe that the UK Monetary Policy Committee "needs to be able to explain exactly why the chart looks as it does and why it changes between Reports. This is vitally important both for the consistency of policy-making and for the presentation of the analysis" (*op. cit.*, p. 31).

Policy preparation and decision making can be a lengthy and inflexible process, with considerable (and asymmetric) adjustment costs. In his study on forecasting student enrolment, Kuhry (1998) states that successive forecasts should not be strongly different when they are to be used for revolving financial planning: "It is hard to believe and undesirable that estimated budget allocations vary much from one year to the next" (Kuhry 1998, p.101; my translation). Policy adjustment costs are part of the policy maker's loss function, and hence stability can be a legitimate criterion. Nordhaus (1987) offers both empirical and anecdotal evidence for smoothing behaviour of professional forecasters.⁹ In a multi-client setting it is far from clear what weight the stability criterion should receive. Smoothing may also be motivated by reputation and credibility of the forecasting institute. Whatever strategy the forecaster follows, he should inform his clients on the balance of risks that are attached to the forecast he puts out.

Again, coherence and stability reflect that the sophisticated client requires some credible story or explanation why the forecast is as it is. If the forecast refers only to a single scalar variable, then the requirement of coherence is void, but stability can still be applied.

⁹ On this issue, see also Clements (1995, 1997).

4. How to forecast?

*"If you're so smart, why aren't you rich?"
('the American question' discussed in McCloskey, 1994)*

Let us distinguish four basic methods to prepare a forecast:

- (1) use a structural model which features the state variables s as endogenous variables;
- (2) use a reduced form model for the state variables s ;
- (3) use a (vector) ARIMA model which includes the state variables s ; and
- (4) use current market expectations to derive a forecast for s .

In practice, one encounters intermediate forms and mixtures of forecasting methods. For instance, sometimes the exogenous variables required for the structural model themselves are forecasts from an ARIMA model. Or current market expectations may be used to feed a reduced form model to derive forecasts for variables on which no current expectations are available. Also, the ideal structural model (founded on production technology and consumer preferences) does not exist, and any real world model will contain some equations of the reduced form type.

In each case, statistical properties of the forecast can only be derived on more or less strong (and largely untested) assumptions about the match between model and real world: for the first three methods those assumptions amount to assuming that the model is correctly specified and holds true for the forecast horizon; for method (4), the assumption is that the relevant markets are efficient and have access to all relevant information.

Why do I say that those assumptions are largely untested? Have not econometricians devised clever procedures to test almost any model assumption that one can come up with? Yes, they have, but any such test is only valid under a host of other assumptions. A simultaneous test of all model assumptions is impossible, simply because every test requires some structure which is taken for granted to make sure the test statistics have the required properties.

In this respect, method (4) appears to be most robust, in that not many assumptions are required to establish basic statistical properties of market expectations. At the other end, structural models tend to be built on a host of untested assumptions (including the choice of functional forms). All these assumptions are required for a standard derivation of statistical properties of the forecast made with the model. In other words, the statistical properties derived with the model are necessarily conditional on

the model assumptions; or, as stressed by Leamer (1978), all of classical statistics is built on the model axiom. But even in a Bayesian setup it is practically impossible to come up with unconditional statistical properties, which take into account all the underlying model uncertainty.

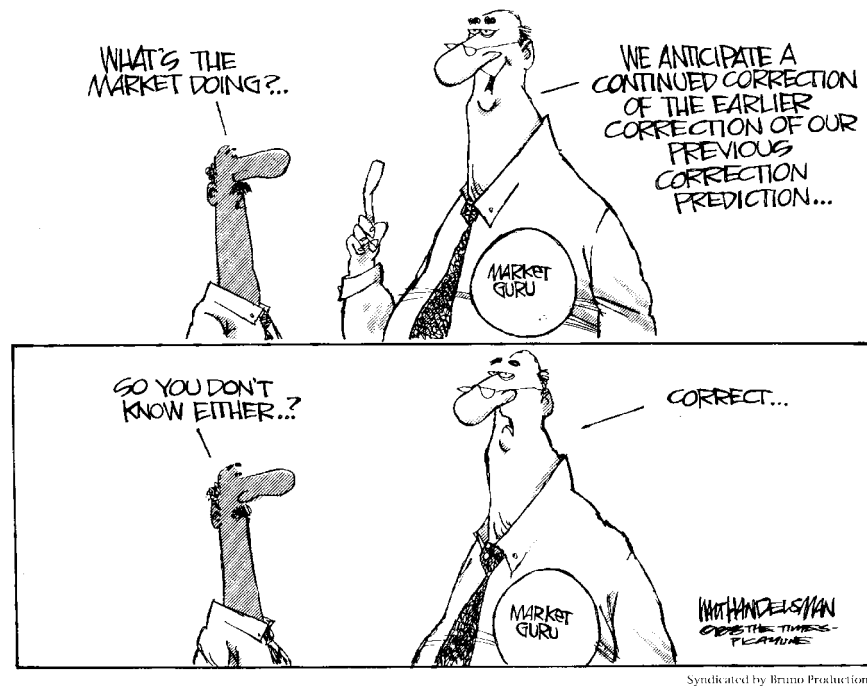
Not only is it uncertain whether the model is correctly specified, there is also the question whether the model will remain valid over the forecast horizon. In his famous critique of econometric policy analysis, Lucas (1976) stresses that often some model parameters are unlikely to be invariant under policy change, and hence the model cannot be used to assess the effects of a policy change. Of course, this would also jeopardize the model forecast when a policy change is anticipated. And the argument is readily extended to include some other (unprecedented) change in environment, making most model forecasts useless because there is almost always some relevant change in policies or environment.

How does this critique affect our different forecast methods? Here methods (1) and (4) appear least vulnerable, for different reasons. A good structural model includes the relevant mechanisms through which policy and environment changes affect the behaviour of economic agents, and should be immune for the Lucas critique. At the other extreme, an anticipated change in policy (or environment) is taken into account by rational market expectations. The critique is most serious for reduced forms and ARIMA models. They cannot convincingly claim to remain valid in the face of new exogenous shocks.

What about coherence and stability? Which forecasts can be supported by a credible story? A creative economist can come up with a story for any forecast. But clearly method (1) is superior to the other three in providing an explanation which is coherent both in terms of the underlying economic system and in terms of changes since the previous forecast. Moreover, this explanation can be checked and be supplemented by a sensitivity analysis. In my experience, these properties of a forecast are indispensable in discussing it with clients, both policy makers and fellow analysts.

To some extent, reduced form models offer the same possibilities, but the explanations that they can provide lack a clear relation to economic theory and structural parameters. This is even more the case for ARIMA models. That a creative economist can come up with one or two stories for any forecast, is illustrated every day by newspaper reports about financial markets. Figure 1 illustrates the economic coherence of such stories.

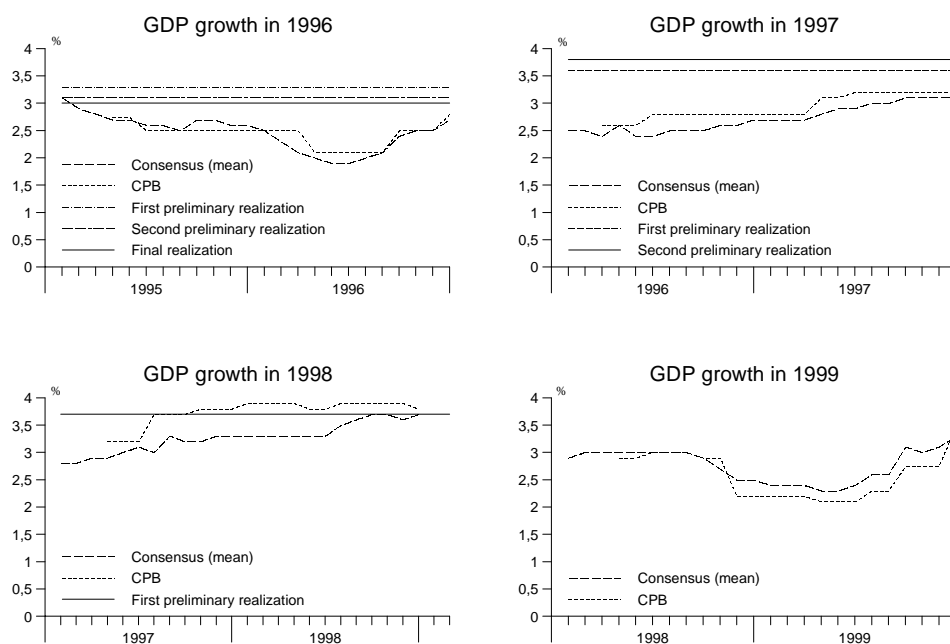
Figure 1 Explaining market expectations



Source: Het Financieele Dagblad, 4 december 1998. Syndicated by Bruno Productions B.V.

As for stability, figures 2 and 3 give some impression of the performance of methods (1) and (4). Both show the evolution in time of the forecasts for Dutch GDP growth rates. Figure 2 contrasts the CPB forecasts, which largely rely on method (1), with the consensus mean as published monthly in Consensus Forecasts (this consensus mean is computed by Consensus Economics from forecasts as put out by a dozen or so commercial and investment banks in the Netherlands). Figure 3 does the same for the Nyfer forecasts, which as far as we know largely rely on method (4). As one might expect, the Nyfer forecasts are less stable than the CPB forecasts, and also less stable than the consensus mean. This observation does not apply to 1999, when Nyfer closely followed the consensus mean. Perhaps they have changed their forecasting method?

Figure 2 CPB and consensus forecasts for Dutch GDP growth, 1996-1999

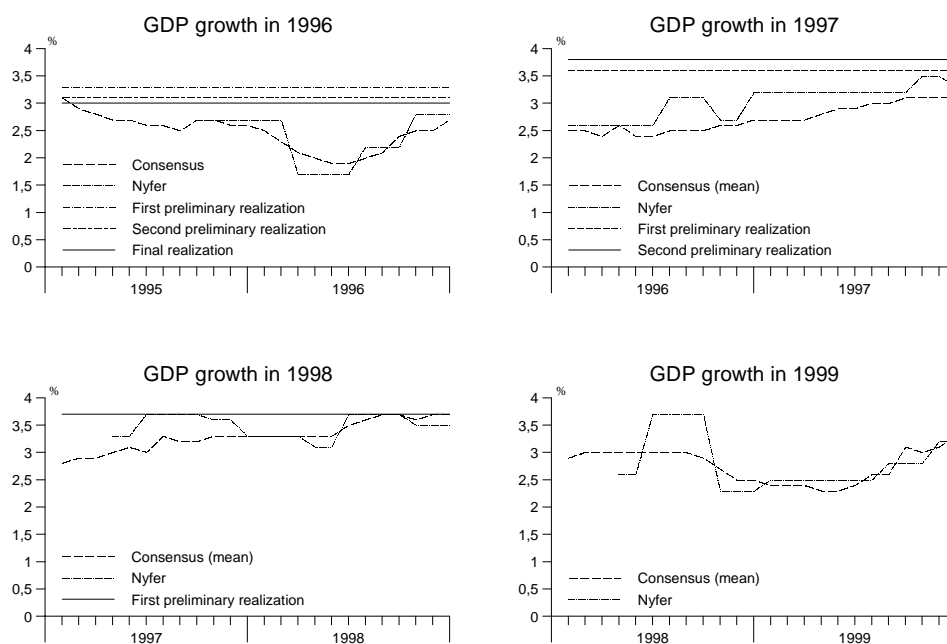


Source: CPB; several issues of Consensus Forecasts (Consensus Economics, London).

The issue here is not smoothing, but the economic coherence of changes in forecasts of GDP growth as derived from movements in financial markets. Arbitrage on future markets forces future rates to satisfy the criterion of logical coherence, even for nonlinear identities (like the dollar per DM rate is the inverse of the DM per dollar rate). But whether efficient market expectations satisfy economic coherence is an open question. Different agents may have different views on its implications, and the market outcome may well deviate from any single coherent economic theory.

Yet, efficient markets do pose a problem for forecasters: how can a forecaster credibly deviate from market expectations? If he truly believes his forecast is superior, why doesn't he "put his money where his mouth is" and get rich? When asked for forecasts of financial market variables, the standard answer at our institute is that, if we knew, then we wouldn't be at CPB but at the Bahama's. Yet we do publish some interest rates and exchange rates together with our macroeconomic forecast. And these published rates reflect our view on an economically coherent forecast, rather than currently expected

Figure 3 NYFER and consensus forecasts for Dutch GDP growth, 1996-1999



Source: CPB; several issues of Consensus Forecasts (Consensus Economics, London); several issues of Nyfer kwartaalbericht (Sdu, The Hague).

rates implied by future markets. Our story then could be that market expectations may diverge, because our forecast is conditional on current official government policy and/or on some particular assumption for a major uncertainty in the near future (like the impact of the Asian crisis). Also, we could point to the possible incoherence and the instability (volatility) of market expectations, and explain the advantages of a coherent and stable forecast. As someone observed: of the last ten turning points, the stock market predicted twenty.

Why, then, aren't we rich? As De Grauwe et al. (1993) have shown, plausible models of foreign exchange markets yield chaotic behaviour for the exchange rate, which supports the idea that it is inherently unpredictable. De Grauwe et al. (1993) find that there is scope for forecasting for the (very) short term, based on past behaviour. For the horizons common in macroeconomic forecasting, neither past behaviour nor the fundamental driving forces provide a reliable basis for forecasting. A similar argument can be applied to the stock and bond markets, as well as to several commodity markets.

Luckily, such markets are not dominant in macroeconomics. Hence there is some real job for macroeconomic forecasting, but one which does not bring exceptional profits.

Some markets are so imperfect that our forecasts may influence the outcome. New forecasts for economic growth may affect the "animal spirits" of entrepreneurs, and influence investment spending. Wage formation in The Netherlands is largely by collective wage bargaining, and it is true that current economic forecasts, in particular from CPB, are taken into account when the bargaining parties prepare their positions. In this case, I would expect our forecasts on productivity, profitability, unemployment and inflation (all interpreted conditionally on our wage forecast) to be more important than our wage forecast itself. However that may be, I find it hard to think of a proper procedure to test whether our prophecies might be to some extent self-fulfilling or self-denying.

5. Communicating with the client

"Past performance provides no guarantee for future results"
(required warning in Dutch commercial investment ads)

As I argued in section 2, the purpose of a forecast is to help clients make a decision in a situation of uncertainty. Because the clients' loss functions are often unknown, we cannot determine the best decision but must communicate our information on the future state of the world s through some forecast f , conditional on a baseline decision d . Let us assume here that the client knows how alternative decisions will affect his targets (if he does not, we'll be happy to help him by providing the proper policy analysis). If certainty equivalence applies, all the decision maker needs is our unbiased forecast.¹⁰ But in the more common case that certainty equivalence does not apply, we should not just communicate our best forecast, but also the uncertainty that comes with it. How can we do that?

The simple answer is: supply the (characteristics of the) distribution of the forecast error, so the client can solve his own decision problem. Too simple an answer, in my view. As I explained in section 4, the statistical properties of the forecast (error) can only be derived on more or less strong (and largely untested) assumptions about the match between model and real world. What the decision maker needs, is the unconditional distribution of the forecast error, which includes the effects of model uncertainty. Even in a Bayesian setting that is a highly impractical demand.

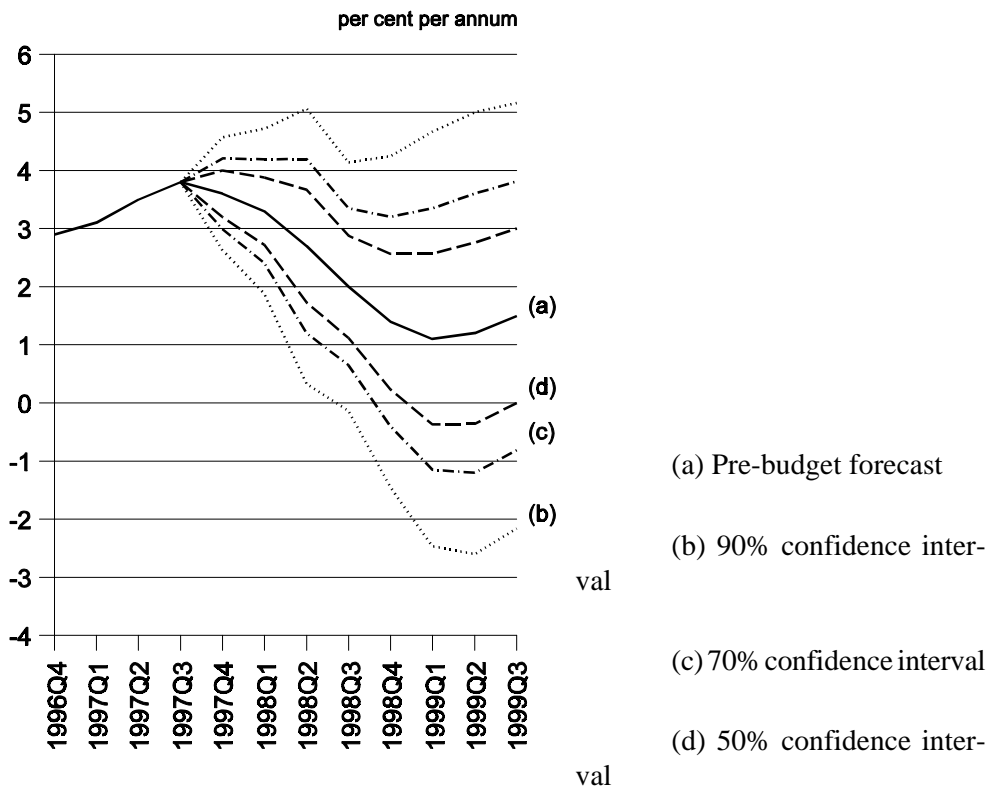
In spite of these fundamental problems, some forecasting institutes have recently started to supply error margins or standard errors with the main variables in their forecast. In its quarterly Inflation Report, the Bank of England (BoE) includes a "fan chart" very much like the one pictured for a Treasury forecast in Figure 4. And the National Institute of Economic and Social Research (NIESR), in its quarterly Review, includes a table stating probability distributions for the growth and inflation forecasts.

There are differences in the types of uncertainty communicated by the BoE and NIESR. BoE specifically allows for subjective views of future uncertainty, see Britton et al. (1998). NIESR is more inclined to rely only on historical forecast errors, though in the issue 3/98 of its Review (page 23-24) it also uses Monte Carlo stochastic simulation of their model to show that forecast uncertainty is not independent of the

¹⁰ This is not as simple as it may seem. As we will see below, it is very hard to specify the proper probability space in which this unbiasedness should hold.

policy regime. For the client, the different types of forecast error margins are hard to distinguish. There is no uniform answer to some important questions: What is the underlying probability space? How is uncertainty about model specification and validity handled? What model axioms are taken for granted? I am afraid that the clients are bound to misinterpret the reported error margins as unconditional probability statements, while in fact they are highly conditional on largely untested model assumptions and/or subjective assessments.

Figure 4 GDP growth confidence intervals for HM Treasury Pre-Budget Forecast



Source: Melliss and Whittaker (1998); Crown Copyright 1998. Reproduced by permission of Her Majesty's Stationary Office.

Yet, it was a bold move of NIESR and the Bank of England to publish the error margins that they have computed, considering their size. In Don (1994) I observed that forecast errors tend to be uncomfortably large for policy makers. This is an understatement if you look at the huge margins that both institutes have presented. Figure 4 gives an adequate illustration, though it refers to a Treasury forecast. One wonders what the clients make of this.

And would not the arguments I just gave suggest that those conditional standard errors still underestimate the true unconditional uncertainty margins? Yes they would. But there is also a possible source of overestimation in applying the standard classical statistical theory to assessing forecast error margins. Classical statistics is too optimistic in posing the model axiom, but too pessimistic in suggesting that all information that we have on the model coefficients comes from the current data sample used for estimation. More information usually is available and used, be it informally, in model selection and parameter choice. This information ranges from tested economic theory to *a priori* insights in what constitutes a plausible parameter value and what does not. Also, non-model information tends to be used in the actual preparation of a forecast. At CPB, the standard example of the latter is sectoral information on investment plans. In principle a careful Bayesian analysis could take care of all that, but this is practically impossible.

In my study of forecast uncertainty (Don, 1994), I compared standard errors from observed forecast errors with standard errors computed from a Monte Carlo exercise. Both approaches led to highly similar numbers for the standard errors (see table 1). The Monte Carlo exercise included four types of uncertainty: errors in preliminary data, parameter uncertainty, forecast errors in (non-policy) exogenous variables, and error terms in behavioural equations. It was argued that for the sample at hand, errors in policy variables would, if anything, most likely have increased the computed standard errors. On the other hand, interference from expert opinion (non-model information) should have reduced them.¹¹ While explicit effects of model uncertainty were not included, one might argue that they had crept in through parameter uncertainty and error terms. That study would suggest, then, that forecast errors can be reasonably assessed through such a Monte Carlo exercise.

The study also showed that, from the four types of uncertainty that were included, by far the most important contributor to standard forecast error is forecast errors in (non-policy) exogenous variables, especially on medium term. Table 2 shows

¹¹ In fact, I pointed to expert opinion for an explanation why short term forecasts for investment and inflation showed substantially smaller observed than computed standard errors. Others have also observed that expert opinion tends to reduce forecast errors, *e.g.* Klein (1981) and Wallis and Whitley (1991).

the percentage shares of the different types in total forecast error. For the very open Dutch economy, it comes as no surprise that forecasting the external variables (most notably real world trade, world market prices and exchange rates) is all important to the quality of the domestic forecast. The same fact is illustrated by figure 5, which shows that one year ahead GDP forecast errors would have been substantially smaller if the external variables would have been known in advance. With knowledge of external variables, the MSE of the forecast is reduced by 40%.¹²

Table 1 Observed and computed standard errors for CPB forecasts

	One year ahead forecast		Four years ahead forecast		
	observed ^a	computed ^b	observed ^a		computed ^b
<i>real growth rates</i>	<i>% per annum</i>				
private consumption	2.3	2.3	1.7	(1.9)	1.5
business investment	6.4	9.6	3.9	(3.7)	4.0
non-energy exports	4.5	5.0	3.0	(1.9)	3.2
production enterprises	2.2	2.5	1.1	(1.6)	1.4
<i>nominal growth rates</i>	<i>% per annum</i>				
consumption price	1.6	2.5	1.7	(1.9)	3.5
wage rate enterprises	2.4	2.1	3.4	(1.8)	3.6
<i>other indicators</i>	<i>% NNP (end year)</i>				
government surplus	na	1.3	na	(1.9)	6.0
current account balancena		2.6	na	(na)	7.0

^a Based on Westerhout (1990); numbers between parentheses based on Hers (1993).

^b Based on Monte-Carlo exercises with four sources of uncertainty.

Source: Don (1994).

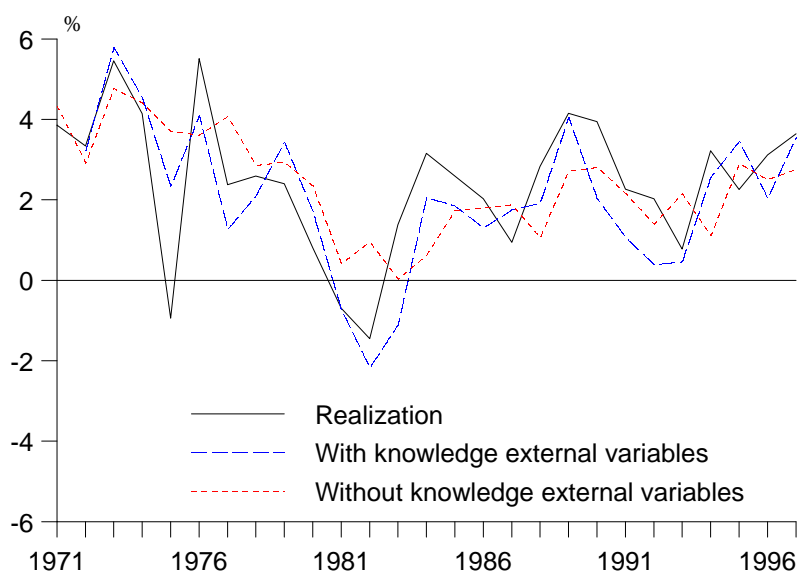
¹² The reduction is measured in MSE rather than RMSE, because the former admits an additive decomposition.

Table 2 Shares of four types of uncertainty in total forecast error variance

	preliminary data	model parameters	non-policy exog. vars	error terms	total
<i>% share in forecast error variance</i>					
A One year ahead forecast					
<i>real growth rates</i>	3	5	33	59	100
private consumption	10	13	4	72	100
business investment	2	2	83	13	100
non-energy exports	3	5	49	42	100
production enterprises					
<i>nominal growth rates</i>					
consumption price	2	2	60	36	100
wage rate enterprises	3	20	27	50	100
<i>other indicators</i>					
government surplus	14	1	70	15	100
current account balance	5	3	67	25	100
B Four years ahead forecast					
<i>real growth rates</i>	4	12	48	36	100
private consumption	4	11	34	51	100
business investment	2	6	62	30	100
non-energy exports	6	11	45	38	100
production enterprises					
<i>nominal growth rates</i>					
consumption price	0	1	91	8	100
wage rate enterprises	1	10	75	13	100
<i>other indicators (end-year)</i>					
government surplus	1	2	93	4	100
current account balance	1	2	89	9	100

Source: Don (1994).

Figure 5 Forecasting Dutch GDP growth with and without knowledge of external variables



Source: CPB

The main problem with applying the Monte Carlo approach in an *ex ante* rather than an *ex post* setting, is that it is very hard to generate error bands for the *ex ante* forecasts of the exogenous variables. Do we feel that past forecast errors on the exchange rate or the oil price provide a proper perspective on current forecast uncertainty? Sometimes yes, sometimes no. While we have learned that oil price shocks are far from unique, in some years the uncertainty in that area appears to be larger, in other smaller than might be derived from forecast errors observed in the past. A similar statement holds for the other external variables. This probably is the most important reason for including subjective elements in the Bank of England's fan charts (Britton et al., 1998).

If we would have a model to generate the forecasts for the exogenous variables, then that model might help us assessing the relevant *ex ante* error bands. However, such a model is likely to be conditioned on some of the variables that are determined on (nearly) efficient markets, like exchange rates and commodity prices. For such variables

it is impossible to derive a probability distribution for the current *ex ante* forecast uncertainty¹³.

Although I sympathize with the idea of formulating subjective probabilities, this Bayesian approach puts high demands on the subject formulating them. As Leamer (1978) puts it: "A formal Bayesian encounters insurmountable difficulties in constructing meaningful prior distributions. Thus, most uncertain judgments elude precise quantification" (*op. cit.*, p. 2). Also it remains unclear whether the client understands this type of subjective uncertainty communicated by the forecaster.

Then how can we properly communicate the uncertainty that comes with our forecasts? Two instruments are commonly used: (1) provide uncertainty variants with the forecast, to highlight some relevant risks; and (2) provide a number of alternative scenarios rather than a single forecast. The first is standard practice for most forecasting institutes. At CPB we have also had some successful experiences¹⁴ using a small number of scenarios rather than a single forecast.

In studies of long term prospects, *e.g.* CPB (1992) and CPB (1997a), we used three or four scenarios to help our clients manage the uncertainty on a 25 year horizon. For the medium term outlook (4 or 5 years ahead), it is now standard practice at CPB to publish two scenarios, one favourable and one cautious in terms of projected economic growth. What is the econometric basis for these scenarios? There is no formal link to forecast error margins or something like that. The scenarios are derived from the model for the domestic economy, feeding it with two different sets of numbers for the exogenous (largely international) variables. These two sets rather informally attempt to capture a reasonable band-width, based on historical trends and an assessment of the economic situation in the base year in the different regions of the world (CPB, 1997b; for a summary in English, see the sections on the medium term outlook in CPB, 1997c). The idea is that these scenarios "show between what margins economic growth in The Netherlands for the projection period is likely to lie, barring extreme conditions" (CPB, 1998, p. 9, my translation). There is no numerical probability statement; rather the flavour is informal and subjective, but coming from independent experts.

Fast adjustment of policy decisions to actual economic developments tends to be difficult, if only because uncertainties are large also for a short horizon. Hence the

¹³ Option prices contain information about market expectations of volatility. On the value of such information, see section 4.

¹⁴ Not all experiences were successful, though. In the early eighties, we prepared two medium term scenarios as the basis for political negotiations on economic policy. A political committee was asked to figure out how to handle them. The result was that the mean growth rate was chosen as the baseline forecast, and policy discussions proceeded on the assumption of certainty equivalence.

advice to policy makers is to prepare, in each policy field, for that scenario which holds the strongest challenge. This means using the cautious scenario in drafting budgetary policies, while at the same time one should prepare for *e.g.* the environmental challenges posed in the scenario with higher GDP growth. While the new coalition governments of 1994 and 1998 did indeed base their budgetary agreements on the cautious scenario, it proved more difficult to simultaneously use the favourable scenario for policies on the environment and infrastructure. Indeed, this scenario approach is a far cry from certainty equivalence, which is so much easier to swallow. Policy makers have a strong tendency to choose a single scenario and use it on the assumption of certainty equivalence, rather than use the different scenarios as a test on the robustness and flexibility of policy choices.

6. Some final remarks

"The only difference between forecasts of economists and fortune-tellers is that economists are better in explaining why they were wrong"

(from M.F.M. Canoy, PhD thesis Tinbergen Instituut, Amsterdam 1993)

Whatever way they are assessed, forecast errors are uncomfortably large for policy makers. Some uncertainty in macroeconomics is related to possibly chaotic processes on financial and commodity markets. Surprises in exogenous variables and error terms will continue to dominate forecast uncertainty. At the same time, continuous efforts are required to keep the econometric models up to date and the contribution of error terms at current levels. In my opinion, therefore, it is unlikely that forecast errors can be significantly reduced. We will have to live with them, and help our clients handle the uncertainty in their decision making process.

Some observers take the relatively large forecast errors as evidence that most econometric policy analysis is useless, because estimated effects of policy options tend to be smaller than the medium term forecast error margin. Apart from the lack of a viable alternative, the uncertainty argument is not valid. I claim that the estimated policy effects are more reliable than the forecasts. As we have seen, the large error margins characteristic of macroeconomic forecasts are largely determined by the uncertainty in non-policy exogenous variables. But most effects of policy options depend only little, if at all, on the external macroeconomic environment. For policy analysis, the more important sources of uncertainty are related to parameter values, model specification and model invariance. However sizable the errors from those sources may be, this is only a subset of the sources of uncertainty that affect the macroeconomic forecasts.

With a similar argument, Hendry and Mizon (1999) question the value of selecting the best forecasting model as the basis for empirical policy analysis: "since forecast failure often results from factors unrelated to the policy change in question, the econometric model may continue to characterize the response of the economy to the policy, despite its forecast inaccuracy".

References

Britton, E., P. Fisher and J. Whitley (1998), "The Inflation Report projections: understanding the fan chart", Bank of England Quarterly Bulletin, February, 30-37.

Clements, M.P. (1995), "Rationality and the role of judgement in macroeconomic forecasting", *Economic Journal* 105, 410-420.

Clements, M.P. (1997), "Evaluating the rationality of fixed-event forecasts", *Journal of Forecasting* 16, 225-239.

Chow, G.C. (1975), "Analysis and control of dynamic economic systems", Wiley, New York.

CPB (1992), "Scanning the Future: a long-term scenario study of the world economy, 1990-2015", Sdu, The Hague.

CPB (1997a), "Economie en Fysieke Omgeving: beleidsopgaven en oplossingsrichting en 1995-2020", Sdu, The Hague.

CPB (1997b), "Economische verkenning voor de volgende kabinetsperiode", Sdu, The Hague.

CPB (1997c), CPB Report 1997/4, Sdu, The Hague.

CPB (1998), "Keuzes in kaart: economische effecten van vijf verkiezingsprogramma's", Sdu, The Hague.

De Grauwe, P., H. Dewachter and M. Embrechts (1993), "Exchange rate theory: chaotic models of foreign exchange markets", Blackwell, Oxford, Cambridge (Mass.).

Don, F.J.H. (1994), "Forecast uncertainty in economics", in: J. Grasman and G. van Straten (eds.), "Predictability and nonlinear modelling in natural sciences and economics", Kluwer, Dordrecht, 569-581.

Hendry, D.F. and G.E. Mizon (1999), "On selecting policy analysis models by forecast accuracy", in the Festschrift in Honour of Michio Morishima, STICERD, London School of Economics.

Hers, J.F.P. (1993), "De voorspelkwaliteit van de middellange-termijn prognoses van het CPB", Research memorandum 107, CPB, Den Haag.

Hildreth, C. (1963), "Bayesian statisticians and remote clients", *Econometrica* 31, 422-438.

Klein, L.R. (1981), "Econometric Models as Guides for Decision-Making", The Charles C. Moskowitz Memorial Lectures no. XXII, The Free Press, New York.

Kuhry, B. (1998), "Trends in onderwijsdeelname: van analyse tot prognose", *Sociale en Culturele Studies* 25, SCP, Rijswijk.

Laster, D., P. Bennett, I.S. Geoum (1999), "Rational bias in macroeconomic forecasts", *Quarterly Journal of Economics* 114, 293-318.

Leamer, E.E. (1978), "Specification searches: ad hoc inference with nonexperimental data", Wiley.

Lucas, R.E. (1976), "Econometric policy evaluation: a critique", in: K. Brunner and A.H. Meltzer (eds.), "The Phillips curve and labor markets", North Holland, 19-46.

McCloskey, D.N. (1994), "Knowledge and persuasion in economics", Cambridge University Press.

Melliss, C. And R. Whittaker (1998), "The Treasury forecasting record: some new results", *National Institute Economic Review*, April, 65-79.

Nordhaus, W.D. (1987), "Forecasting efficiency: concepts and applications", *Review of Economics and Statistics* 69, 667-674.

Smith, R. (1998), "Emergent policy-making with macro-econometric models", *Economic Modelling* 15, 429-442.

Theil, H. (1958), "Economic Forecasts and Policy", North Holland, Amsterdam.

Van den Berg, P.J.C.M. (1986), "De betrouwbaarheid van macro-economische voorspellingen", Economisch Statistische Berichten 71, nr. 3583, 1153-1157.

Wallis, K.F. and J.D. Whitley (1991), "Sources of error in forecasts and expectations: U.K. economic models, 1984-8", Journal of Forecasting 10, 231-253.

Westerhout, E.W.M.T. (1990), "Voorspelling en realisatie", Discussienota 9001, Ministerie van Economische Zaken, Den Haag.

Abstract

Forecasts are made in order to help decision making. As loss functions are usually unknown, and certainty equivalence is unlikely to prevail, a forecaster must communicate not just his forecast but also enable the client to form his opinion on the associated uncertainty. Communicating statistical properties tends to be very difficult, because of implicit or explicit conditioning on other variables, on model parameters or specification, and sometimes on subjective probability assessments.

Three non-statistical criteria for forecast quality are put forward: logical coherence, economic coherence, and stability. Against this background, alternative forecasting techniques are discussed. For communication with the client, uncertainty variants and alternative scenarios appear adequate. Still policy makers have a strong tendency to choose a single scenario and use it on the assumption of certainty equivalence, rather than use different scenarios as a test on the robustness and flexibility of policy choices.