Formalizing judgemental adjustment of model-based forecasts*

Philip Hans Franses†

Econometric Institute, Erasmus University Rotterdam

April 24, 2006

Econometric Institute Report 2006-19

*This paper has been prepared for possible inclusion in the special issue of the International Journal of Forecasting on The Use of Judgement in Forecasting, to be edited by Paul Goodwin and Mathew Parackal. I thank Fergus Bolger, Henk Don, Dick van Dijk, Robert Fildes, Richard Paap, Teun Kloek, Chris Sims and Peter Vlaar for helpful discussions. A presentation of this paper was given at the Netherlands Bureau for Economic Policy Analysis (2003), the Netherlands Central Bank (2004) and the University of Groningen (2004). Simone Jansen and Christine Santoso provided useful research assistance, and Merel van Diepen has been very helpful with the computer programs.

†Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, The Netherlands, `email: franses@few.eur.nl`

Abstract

In business and in macroeconomics it is common practice to use econometric models to generate forecasts. These models can take any degree of sophistication. Sometimes it is felt by an expert that the model-based forecast needs adjustment. This paper makes a plea for a formal approach to such an adjustment, more precise, for the creation of detailed logbooks which contain information on why and how model-based forecasts have been adjusted. The reasons for doing so are that such logbooks allow for (i) the preservation of expert knowledge, (ii) for the possible future modification of econometric models in case adjustment is almost always needed, and (iii) for the evaluation of adjusted forecasts. In this paper I put forward an explicit mathematical expression for a judgementally adjusted model-based forecast. The key parameters in the expression should enter the logbook. In a limited simulation experiment I illustrate an additional use of this expression, that is, looking with hindsight if adjustment would have led to better results. The results of the simulation suggest that always adjusting forecasts leads to very poor results. Also, it is documented that small adjustments are better that large adjustments, even in case large adjustments are felt necessary.

Key words: Forecasting, Judgemental adjustment
JEL code: C53

1 Introduction

In many practical situations in business and in macroeconomics it occurs that forecasts are based both on the outcome of an econometric model and on the added touch of an expert. This practice of adjustment appears to be indifferent from the perceived quality of the econometric model, and also simple trend extrapolation techniques experience the human factor. When the addition of expert knowledge occurs after the expert has seen the forecast, this is called judgemental adjustment and the associated forecast is called a conjunct forecast. This definition is crucial as such adjustment differs from rules where the forecast has not been seen before. For example, in forecast combinations, see Clemen (1989) and Timmermann (2006) one simply combines a variety of forecasts without at first dismissing a few. In this paper, I address the issue of how to understand the creation of an expert-adjusted model-based forecast.

The key question is how one can understand the added value of an expert to a forecasting model. The discussion in this paper concerns (i) the way the expert adds information to the forecast, (ii) the way the expert could gain trust, and (iii) the way one can evaluate the forecasts afterwards. An attempt is made to formally describe how such expert-adjusted forecasts are made, where simplicity of the explicit representation of the judgemental adjustment process is important.

The formal model is empirically illustrated for forecasting US monthly inflation, where a simple time series model is the starting point. I review various ways an expert could have modified the forecast. The simulation results show that judgemental adjustment can lead to better forecasts, but an important conclusion is THAT it is useful only once in a while. In fact, the so-called 50-50% rule, balancing the model and the expert as advocated by Blattberg and Hoch (1990), is shown to lead to very poor performance.

The outline of the paper is as follows. In Section 2, I discuss more details of judgmental adjustment of model-based forecasts. This must all be seen as a prelude to the material in Section 3 where more formal expressions are given. These formal expressions are felt as necessary, as it allows one to document in a logbook why and how adjustment happened[1]. In Section 4 a potential use of the formal expression is given by performing a limited simulation experiment. When forecasting monthly US inflation, model-based forecasts are adjusted using various rules. The simulation results suggest that always adding expert knowledge leads to the poorest results, and that small adjustments are better than large adjustments, even when the latter are deemed necessary. These results all the more show that logbooks are important. Section 5 concludes this paper with various topics for further research.

2 Motivation

Econometric models are often used for forecasting. Independent of their level of sophistication, it is common practice that forecasts from these models are not always mechanically accepted and that, say, an expert might redirect the forecast based on his or her opinion, see Turner (1990) and Donihue (1993), among others. This judgemental adjustment is not considered to be bad practice showing a lack of confidence in the model, but merely it indicates that a model cannot incorporate all possible features, and sometimes recent and important information can lead to an adaptation of the model forecast. This strategy is widely practiced in governmental forecasting, and also in business and industry. Interestingly enough, to my knowledge there are no studies which consider the precise construction and evaluation of such, what are called, conjunct forecasts. This paper aims to put forward some thoughts that might lead to formalizing the added value of an expert to a forecasting model. This formal approach is relevant as it allows preserving expert knowledge, as it might lead to persistent modification of the model, and as it allows for an evaluation of the quality of the conjunct forecast.

[1]Conversations with professional forecasters associated with the Netherlands Bureau of Economic Policy Analysis (CPB) indicated that this Bureau keeps logbooks since 1999, where detailed minutes are kept on why their main macroeconomic model-based forecasts were adjusted. I thank Rocus van Opstal, Henk Kranendonk and Debby Lanser of the CPB for these conversations.

What is a conjunct forecast?

In basic notation, judgemental adjustment amounts to the following. Suppose one has an econometric model for a variable y_t, were $t = 1, 2, ..., n$, and n marks "now", and suppose it gives the one-step ahead forecast $\hat{y}_{n+1}$. An expert might believe that this forecast needs to be adjusted *once* the expert has seen the forecast. This adjustment can be denoted as $\hat{a}_{n+1}$, where the ˆ indicates that the size of the adjustment needs to be estimated too. Together we then have the conjunct forecast f_{n+1} which equals $\hat{y}_{n+1} + \hat{a}_{n+1}$.

The emphasis on *once* is important here. When an adjustment always takes place, without having a look at the model-based forecast, then that simply is a forecast combination. Amongst the reasons for adjustment, once the model-based forecast has been seen, are that sometimes past patterns cannot be extrapolated, that relationships may not remain constant, that an event is known to happen soon which has never happened before, and that the model-based forecast is unpleasant or inconvenient for the manager or the minister[2]. Basically, the user or the seller of the forecast wants to exercise control, and does want not to leave all to the model. So, the starting point here is a model-based forecast of which an expert believes it needs adjustment[3].

To be precise, the issue at stake in this paper differs from the notion of combining forecasts. This notion assumes that there can be two or more models for the same phenomenon, where one of the "models" is the forecast of the expert. These models can both be used for forecasting. One may want to combine the forecasts from these two models, and there is a large literature on this. A basic premise of the combination of forecasts is that these models are considered equally important and that the decision to combine or not is independent from the forecast outcome of one of the models. The literature on forecast combination is large, and a recent excellent

[2]Sometimes forecasts are said to be adjusted because an input variable is estimated too high or too low. However, I feel that if such a situation occurs, one should modify the value of that input variable *before* generating a model-based forecast.

[3]A concept related to the creation of a conjunct forecast is intercept correction, which is reviewed in for example Clements and Hendry (1998). The authors provide a taxonomy of forecast errors and their origins, and suggest to use intercept correction for potential improvement. To my knowledge, they nowhere discuss how to do intercept correction and ho to determine its size

survey appears in Timmermann (2006).

Literature review

Judgemental adjustment makes sense, as Webby and O'Connor (1996) state "(it) implies that judges will outperform models when they have contextual information to help them comprehend discontinuities in series." Contextual information is defined as "information, other than the time series and general experience, which helps in the explanation, interpretation and anticipation of time series behaviour".

To comprehend the merits of such information, it seems relevant to write down what this information is, and how it is implemented. Indeed, Goodwin (2002, p. 129) remarks that "This suggests that using a formal decomposition model to structure judgemental adjustments to statistical forecasts may lead to improvements over informal adjustment. Surprisingly little research has been carried out to date to examine this possibility". This formal decomposition is even more relevant as a survey amongst academics, the results of which are documented in Franses (2004), shows that most editorial board members of academic forecasting journals believe that forecasts can be improved when expert knowledge is combined with model-based forecasts.

There are some studies that focus on more formal approaches to combine forecasts from models and experts. The relevant literature includes Saaty (1990) with the Analytical Hierarchy Procedure (AHP), see also the discussion of this method in Salo and Bunn (1995) and in Belton and Goodwin (1996), among others. The main finding is that AHP might sometimes yield better results, but it is not exactly clear why this is the case. Flores and Olson (1992) have another method than AHP, although there are some similarities. Still, their method is rather complicated, even though it only deals with the linear combinations of the variables that one uses to adjust.

In contrast, in this paper I propose a rather simple tool to formalize the conjunct forecast. It contains a few components that can simply be disentangled from the final forecast, and can be stored and analyzed afterwards.

Further motivation

Why is it important to formalize the practice of judgemental adjustment? This formalization can help to document the answers to the following relevant questions in practice: (i) why and when do we want to adjust?, (ii) how do we adjust? and (iii) how can we evaluate the resultant conjunct forecast?

It is important to study the properties of conjunct forecasts, as the study of forecast errors in informative for future modifications. Indeed, a crucial reason to use formal econometric models and to analyze their forecast errors is that this allows us to learn from our mistakes, see Clements and Hendry (1998). Hence, if it is unknown how the expert changed the model-based forecast, there is no way to learn from past forecast errors. When both the model forecast and the conjunct forecast would be published, then it would be possible to disentangle the relative forecast errors. However, to my knowledge, this rarely happens, if at all. This might seem odd, as along these lines one cannot only learn from past errors, but also the expert cannot build up trust from fellow forecasters or from those who use the forecast for decision making. Indeed, it would be favourable for the expert to see that his or her alteration of the model-based forecasts would indeed lead to substantial improvement.

When one accepts the idea that an expert adds something to a model forecast, this does not mean that the econometric model should be discarded or that the expert assumes the model is wrong or useless. Merely, one can think that the model summarizes some key features of the data, and that an expert might feels the need to redirect the model outcomes. It is unlikely that the expert knowledge is always needed, and that all forecasts should be conjunct forecasts. Hence the 50-50 % rule advocated in Blattberg and Hoch (1990) is merely a plea for forecast combination then for adjustment[4]. In sum, the decision to construct a conjunct forecast depends on the model outcome.

The main goal of this paper is to stress that for conjunct forecasts to be useful,

[4]Personal communication with various professional governmental forecasters suggests that in reality only sometimes the model outcomes are adjusted.

one needs to clearly document how one has arrived at these forecasts. Without such documentation, one cannot learn from past errors. The idea of this paper is to propose a very simple and stylized version of how forecast adjustment basically works. There may be many modifications needed to this basic framework, but that can be relegated to future work.

3 Formalizing expert adjustment

This section puts forward an explicit expression of a conjunct forecast. The ingredients of the expression can be recorded. The main issues that are addressed are (i) the ways one can combine forecasts given that the expert's opinion is only added to the model forecasts *once* this last forecast is observed and evaluated on its merits, (ii) a description of what exactly the expert's knowledge is, and (iii) the size of the modification to the model-based forecast.

When do we modify model-based forecasts?

Consider a variable y_t for which one has observations $t = 1, 2, .., n$. The goal is to forecast one-step ahead[5]. So, one wants to find the best possible forecast f_{n+1} for the true observation y_{n+1}.

Suppose one has an econometric model, and let it be given by

$$y_t = X_t\beta + \varepsilon_t, \tag{1}$$

where it is assumed that the expected value of ε_t is zero. Here I choose for a linear regression model, but other types of models are feasible too. The main notion is that one intends to forecast a continuous variable, like inflation, sales, GDP, unemployment and so on, and not a categorical variable[6]. Further, the X_t can contain various explanatory variables, but it may also contain terms associated with exponentially weighted moving averages. In other words, there is no need to assume

[5]Adjusting multiple-step ahead forecasts seems to be a bit different than a one-step ahead forecast, as then the model really misses something like long-run trends or foreseeable regime shifts. More research is needed here.

[6]Forecast adjustment of, say, a binary variable seems to be different as the forecast is a probability.

that the regression is the best possible model. However, when it is, it means that the parameters can be estimated assuming a squares loss function, and the forecast from this model is based on the same assumption.

If one would use the regression model for forecasting the $(n+1)$-th observation, one could use

$$\hat{y}_{n+1} = X_{n+1}\hat{\beta}_{[1,2,..,n]}, \tag{2}$$

where $\hat{\beta}_{[1,2,..,n]}$ means that the parameters are estimated using the sample data from 1 to n. When a new data point becomes available, the sample shifts to $n+1$, and in many practical situations the parameters are estimated again. When the values of X_{n+1} are known, for example as they include lags, the forecasts are easy to make. When the values of the explanatory variables X_{n+1} are unknown, they have to be estimated too, most likely also using a regression-type model. When these forecasts of X_{n+1} seem implausible, these values should be modified prior to generating forecasts of y_{n+1}. Further, it is assumed that the parameters are estimated in the best possible way. Hence, one cannot wish to modify a model-based forecast because of a lack of confidence in the parameter estimates, as one better first seek an appropriate estimation method or modify the model.

As an example, one which will return below in Section 4, consider a monthly inflation rate, seasonally adjusted, and suppose that a forecasting model is given by an autoregression of order 1, AR(1). Hence, the model is $y_t = \rho y_{t-1} + \varepsilon_t$, and the one-step ahead forecast (under mean squared error loss) for $n+1$ is $\hat{\rho} y_n$, where $\hat{\rho}$ is obtained through ordinary last squares [OLS]. The one-step ahead forecast error is ε_{n+1}.

The final forecast can be the model-based forecast, that is, f_{n+1} might be set equal to $\hat{y}_{n+1}$. However, sometimes the value of $\hat{y}_{n+1}$ is such that an expert wants to modify this. Such a modification of the forecast could happen when $\hat{y}_{n+1}$ exceeds some upper threshold value, say, c_2, or be below some bottom threshold value, say c_1, where c_1 may differ from $-c_2$ and where also both values can be positive or be negative. One needs to scale $\hat{y}_{n+1}$ by a standard deviation, which might be set at that of the variable y_t itself. Denote this standard deviation as σ_y, where its value gets estimated using the sample running from 1 to n. For the sake of clarity, c_1 and c_2 are assumed as fixed, but one can of course allow for time-varying boundaries.

To obtain an expression for the decision to adjust the model-based forecast, define the indicator function

$$F_y(\hat{y}_{n+1}) = \frac{1}{1 + \exp[-\gamma_y(\frac{\hat{y}_{n+1}}{\hat{\sigma}_y} - c_1)(\frac{\hat{y}_{n+1}}{\hat{\sigma}_y}) - c_2)]}, \tag{3}$$

where the positive parameter γ_y measures the speed at which this modification is considered to be relevant. A graphical example of this function in given in Figure 1. This function takes values in between 0 and 1. Symmetry is not imposed as a too high a forecast can be considered as less bad than too low a forecast. The values of $F_y(\hat{y}_{n+1})$ are close to 0 when $c_1 < \frac{\hat{y}_{n+1}}{\hat{\sigma}_y} < c_2$, and the function values approach 1 when the scaled $\hat{y}_{n+1}$ is in excess of these boundaries. When γ_y is large, the change from 0 to 1 occurs instantaneously. The values of the function $F_y(\hat{y}_{n+1})$ also indicate the degree of which the expert knowledge is added. If the function would always be equal to 1, then there apparently is always a reason to adjust, and then one has the 50-50 % rule. Note that when there is always a reason to adjust, *once* the forecast has been seen, the model forecast apparently is always beyond pre-set boundaries, which really means that the model is considered as unreliable. The model apparently always gives too extreme a forecast, and hence expert judgement is always felt as necessary to dampen the forecast more towards the mean[7]. Again, one might also want to allow the c_1 and c_2 parameters to change over time, although fixing them to values as -2 and 2 would match with rules where forecasts get adjusted when the boundaries of more than two times the standard deviations are crossed.

To continue the example, suppose the inflation rate has a mean value of 2 percent and a standard deviation of 0.5 percent, and suppose the current value is 3 percent. If the model predicts 5 percent, one may be inclined to adjust this forecast. One

[7]It may be that the model gives too conservative forecasts, and that one wishes to adjust such that more extreme values can appear. One can then simply replace the expression for $F_y(\hat{y}_{n+1})$ by another one. Other rules are also possible. One might think of adjusting model-based forecasts when they exceed two standard errors of the regression from the current value of y_n. Simple alternative expressions for $F_y(\hat{y}_{n+1})$ can be proposed. Finally, it may be that some other source of information should be included in a switching function. As $F_y(\hat{y}_{n+1})$ may be associated with a probability, when the likelihood that an important event might take place, like a war that starts or the death of a president, is above some threshold, the function value switches from 0 to 1.

might also want to do that when the forecast would be 0.5 per cent. Of course, one can consider sub-samples to compute those means and standard deviations and perhaps take only the last few years into account.

Which information do we use to modify the forecast?

Suppose the decision is taken to modify the forecast, there must also be a sizeable reason to do so. Indeed, when one intends to add to the forecast the most recent change in the oil price, but that change is zero, there is not much to adjust. Hence, the information that one intends to use for adjustment must be relevant and significant. If one just adds the recent value of a variable that was not in the model, then in fact it could have been included in the model in the first place.

Suppose the expert aims to add to the model-based forecast a linear combination of variables Z_{n+1} with weights γ, that is, $Z_{n+1}\gamma$, where Z_{n+1} is assumed to be known at n (perhaps because they concern lagged variables). It seems plausible to assume that only when the value of $Z_{n+1}\gamma$ is somehow relevant one considers adjusting the model-based forecast. So, only when there is additional information of some importance, one may want to add it to the model-based forecast.

For example, in times of increasing interest rates, one might believe that forecasts for inflation need to be adjusted upwards. When interest rates unexpectedly decrease, one may want to adjust inflation forecasts downwards, but perhaps not in a similar way as when interest rates increase.

Also for this decision one can rely on a specific mathematical expression. In fact, one can use a similar function as above, where now σ_z denotes the deviation of $Z_{n+1}\gamma$ (measured up to and including n), that is, one can consider the indicator function

$$F_z(Z_{n+1}\hat{\gamma}) = \frac{1}{1 + \exp[-\gamma_z(\frac{Z_{n+1}\hat{\gamma}}{\hat{\sigma}_z} - d_1)(\frac{Z_{n+1}\hat{\gamma}}{\hat{\sigma}_z}) - d_2)]}, \tag{4}$$

where again d_1 and d_2 are bottom and upper thresholds, with $d_2 > d_1$ and where again symmetry is not needed. For example, when oil prices go up, the economy shows different responses than when oil prices go down. So, when $d_1 < \frac{Z_{n+1}\hat{\gamma}}{\hat{\sigma}_z} < d_2$, there might be felt a need to adjust the model-based forecast but apparently there is no relevant information for the size of that adjustment. The question is of course how the expert gets values of γ. One possibility is that the expert has another model for that purpose.

The conjunct forecast

Finally, one needs to decide on the size of the adjustment $\hat{a}_t$ itself. For example, it may be that the change in the oil price is beyond 10 percent. This may provide sufficient reason to do an adjustment, but the adjustment itself can then be as small as 0.1. The value of $\hat{a}_t$ depends on various factors, most notably on the type and size of the variable to be forecast. Also, one might want to have small-sized adjustments, also as making large adjustments does not show great confidence in the model.

Taking all expressions together, a simple expression for the final conjunct forecast is then

$$f_{n+1} = X_{n+1}\hat{\beta}_{[1,2,..,n]} + F_y(\hat{y}_{n+1})F_z(Z_{n+1}\hat{\gamma})Z_{n+1}\hat{\pi}, \tag{5}$$

where $Z_{n+1}\hat{\pi}$ is the value to be added[8]. Of course, one may also consider multiplication instead of summing, although in practice such re-scaling is rarely seen. The idea of this expression is that the contribution of $Z_{n+1}\hat{\pi}$ gets weighted by the relevance of $Z_{n+1}\hat{\gamma}$ and by the degree to which $\hat{y}_{n+1}$ is outside a certain range.

An important by-product of (5) is that by comparing the f_{n+1} with the true value y_{n+1}, one can decompose the forecast errors. The overall forecast error is

$$y_{n+1} - X_{n+1}\hat{\beta}_{[1,2,..,n]} - F_y(\hat{y}_{n+1})F_z(Z_{n+1}\hat{\gamma})Z_{n+1}\hat{\pi}, \tag{6}$$

and one can easily compute various error measures based on comparing (6) with

$$y_{n+1} - X_{n+1}\hat{\beta}_{[1,2,..,n]}, \tag{7}$$

being the model-based forecast error.

What should be documented in the logbook?

Table 1 contains the key parameters that should be included in the minutes in the logbook. One needs to have information on (i) why the model-based forecasts needed

[8]Basically, what the Analytical Hierarchy Process does is to quantify the $Z_{n+1}\hat{\pi}$. Otherwise, it is assumed that adjustment is always necessary, that is, it is assumed that $F_y(\hat{y}_{n+1})$ is equal to 1, and that $Z_{n+1}\hat{\pi}$ is such that $F_z(Z_{n+1}\hat{\gamma})$ is also equal to 1.

adjustment (the function $F_y(\hat{y}_{n+1})$ with key parameters γ_y, c_1 and c_2), (ii) why an adjustment is relevant and which variables and factors establish this relevance (the function $F_z(Z_{n+1}\hat{\gamma})$ with its key parameters), and (iii) the adjustment itself ($Z_{n+1}\hat{\pi}$).

4 Illustration

The expression of the conjunct forecast is useful for keeping track of expert knowledge, that is, the values of $F_z(Z_{n+1}\hat{\gamma})$ and of $Z_{n+1}\hat{\pi}$, for keeping track of reasons why forecasts were adjusted in the first place, that is, $F_y(\hat{y}_{n+1})$ and the forecast errors, which can be assigned separately to the model and to the expert.

In this section the expression in (5) is used for another purpose, that is, it can be used to see, afterwards, whether adjustment would have led to better final forecasts, and which type of adjustments would have been most beneficial. Such an exercise can also be useful to modify the model.

To continue with the example on forecasting inflation, the question that can now be answered with hindsight, is whether it could have been beneficial to include adjustments based on for example changes in interest rates.

Set-up of experiment

To illustrate, I consider monthly data on US inflation, all items, covering CPI data for 1959.01-1999.12. The data have been seasonally adjusted. I assume that a simple first order autoregression [AR(1)] is used to generate model-based forecasts[9]. The forecasts are created in two ways. The first is to fit the model for the first 120 observations, to forecast one-step ahead, and then to add the 121-th observation, to re-estimate the parameters, and to make again a one-step forecast, and so on. This strategy yields forecasts based on recursive samples that each time increase with one observation. In sum, there are 370 forecasts. The second strategy is similar, but now when adding the 121-th observation, the first observation is discarded. Hence, model parameters are now re-estimated each time for 120 observations. This second

[9] Inflation is computed as $\log CPI_t - \log CPI_{t-1}$, and an AR(1) model absorbs again one observation.

strategy involving rolling samples might be viewed as more adaptive to breaks or changing trends in the data.

The basic model is the AR(1) model, but it is assumed that adjustments to the forecasts could have occurred, where the adjustment would be based on a one-month lagged unemployment rate or a one-month lagged change in the 3-month treasury bill interest rate. This defines Z_{n+1}. I take a one-month lag as such observations would be actually available by the time one might want to create a conjunct forecast for inflation.

The parameter configurations for the expression in (5) appear in Tables 2 and 3. In all cases, the γ_y and γ_z parameters are set at 100. For c_1 and c_2 in the function concerning the decision to adjust the forecast are set at -2, -1 and 0, and 2, 1 and 0, respectively. Roughly speaking the -2 and 2 values indicate that one-step ahead model-based forecasts beyond the 95% confidence interval are corrected, while the 0 values mean that these forecasts are always adjusted. For d_1 and d_2 in the function concerning the relevance for adjustment, that is whether there are any noticeable relevant values for interest rates and unemployment rates, the values are also set at -2, -1 and 0, and 2, 1 and 0, respectively, with similar interpretation as above. Hence, when d_2 and d_1 are zero, any change in these two variables is considered relevant enough to adjust the forecast of inflation. Finally, the values for π depend on the scale of measurement of these two variables in Z_{n+1}, but clearly, the larger they are in absolute sense the larger is the adjustment. Tables 2 and 3 give the ratios of the Root Mean Squared Prediction Error [RMSPE] of the conjunct forecast over the model-based forecast, when averaged over 370 forecasts. Clearly, ratios smaller than 1 indicate the superiority of the conjunct forecast.

Results

Table 2 gives the results for the quality of conjunct forecasts over model-based forecasts when the adjustment is based on changes in the unemployment rate, while Table 3 concerns changes in the interest rate.

A couple of conclusions can be drawn. When the two main panels concerning the two types of samples are compared, it is clear that differences across these two panels are small, although in general the ratios are larger for the rolling sample forecasts. Hence, adding expert knowledge seems less beneficial in cases where parameter estimates are allowed to vary more over time, and hence the model-based forecast is not that bad in the first place.

A second conclusion, which holds across the two tables, is that for larger absolute values of c_1 and c_2 the expert's added value can yield better forecasts. In fact, always adding an expert's opinion, which concerns the cases where these two parameters are zero, and where this expert's adjustment is large, leads to the worst forecasts. This suggests that if adjustment is beneficial, such adjustment should be either small and/or rare.

A third conclusion can be drawn when comparing the results for differing values of the d_1 and d_2 for the same values of the c_1 and c_2 parameters. When the absolute values of the d parameters get smaller, meaning that an expert more frequently sees a reason for adjustment, the quality of the conjunct forecast gets lower. This suggests that it may seem wise to be a little reluctant to adjust, even though a focal variable takes unexpected values.

A fourth and final conclusion to be drawn from both tables is that the size of the adjustment π should better be small. In this case of forecasting inflation, the added value of the expert by including 0.1 times the one-month lagged interest rate only once in a while can lead to a ratio like 0.958, meaning that a conjunct forecast could have been about 4% better than a model-based forecast, at least for this variable[10].

Although this exercise is merely a simulation experiment to illustrate how one can use the explicit expression for a conjunct forecast, it does suggest some results that might carry over to other settings. These results are that conjunct forecasts can be better than model-based forecasts in case it is not often felt that the model-based forecast needs adjustment, that only significant changes in the outside variable are taken as a sign to adjust, and that only small adjustments are allowed.

[10] In this particular case, adding the one-month lagged interest rate to the AR(1) model yields a 5% significant parameter with estimated value 0.078, which is indeed close to 0.1. Here this might result in a modification of the model, also as the forecast improvement holds for the cases where the c and d parameters are all zero.

5 Discussion

This paper has put forward an explicit, though very simple, expression for a conjunct forecast, which is an adjusted model-based forecast. Adjustment only takes place when a model-based forecast has been seen and evaluated. The expression also formalizes what might go on in the mind of the expert. Together, it allows for a documentation of how judgemental forecasts were established, such that one can learn from mistakes or find suggestions to modify the model. A limited simulation experiment on forecasting monthly inflation showed that even a simple time series model might not be easy to beat by adding an expert's opinion, and that always adding an expert's adjustment can lead to very poor results, in particular when the size of such adjustment is large. Another intriguing finding is that even when the OLS estimate of an added variable is significant, improved forecast performance can only sometimes be observed. This suggests an interesting area for further research, which is that also for regression models with all significant parameters one might only once in a while use the variables for out-of-sample forecasting. So, perhaps the inclusion of functions like those in (5) in regular linear regression models, when it comes to forecasting, might also be beneficial to forecast quality.

Lessons for practice

The key lessons from this paper for everyday practice of forecasting is that it pays off to keep track of model-based forecasts, the reasons for their adjustment and the precise construction of the adjusted forecast. These track records can be used to, afterwards, see if other types of expert adjustment could have been better (like the exercise on forecasting inflation, where unemployment rate and interest rate were considered). Also, it allows one to learn from forecast errors, and to see if models can be improved, or if expert adjustments need focus or fine tuning. In the best case, the track record can show that the expert opinion really is beneficial, providing trust in the capabilities of the expert.

Further work

Besides the suggestion above, I see various issues as interesting topics for further research. First, when it so turns out that the added value of an expert in only noticeable once in a while, perhaps we should not anymore use forecast evaluation criteria that equally treat all observations. Van Dijk and Franses (2003) have proposed evaluation criteria that allow for zooming in on only a few relevant data points for non-linear models, but further work in the present area seems relevant.

Second, in the simulation experiment I simply used a few parameter configurations in the two switching functions only for illustrative purposes. In practice, we need of course methods to decide on the values of for example γ and π. Additionally, the experiment relied on a single variable in Z, and the question of course is how one should incorporate more than one variable. Perhaps the application of principal components analysis can be useful here, but further work is needed.

Finally, model-based forecasts usually come with confidence bounds, and it also seems important to construct such bounds around conjunct forecasts. Most likely is that one needs simulation methods to compute such bounds. These bounds can then also be used to evaluate the post-hoc quality of the conjunct forecast. Indeed, while the model-based forecast is usually created under the assumption of a mean squared error loss function, the conjunct forecast must have another loss function. It is unlikely that one can know this last function, and hence other ways of evaluating expert-adjusted forecasts seem needed.

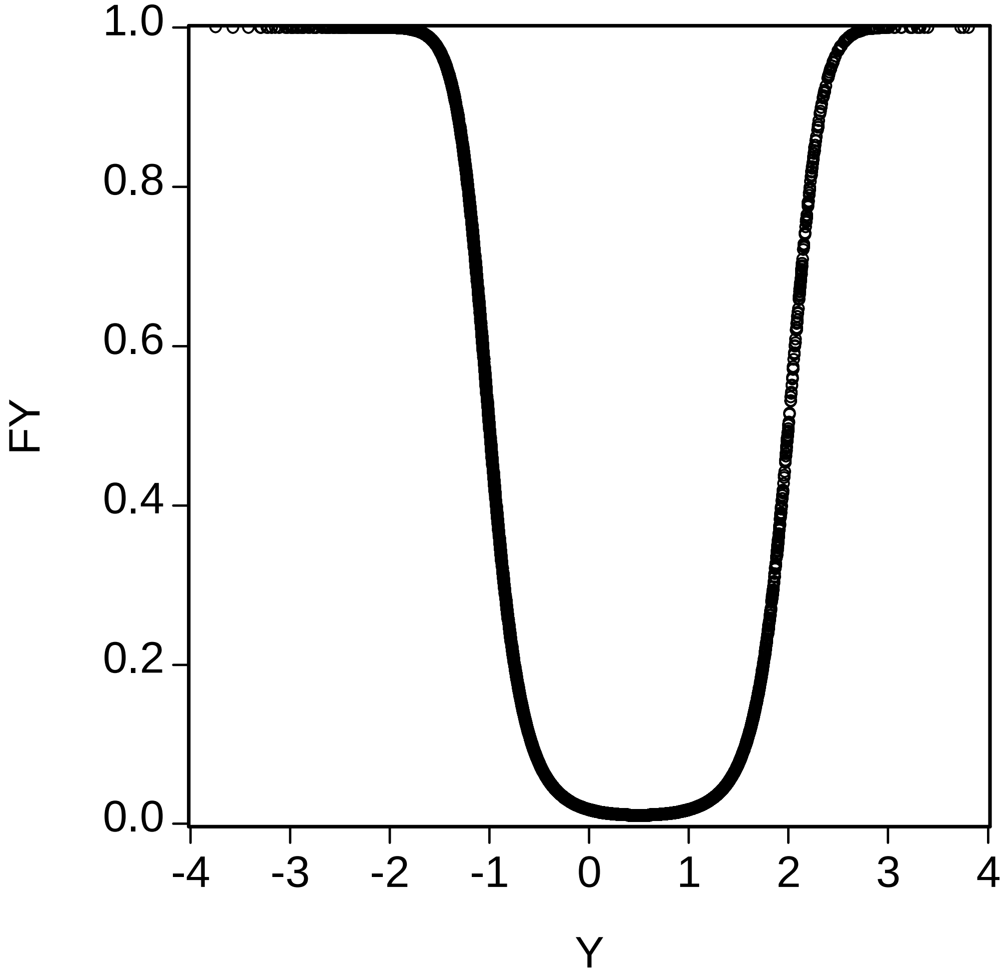

Figure 1: An example of the indicator function

Table 1: Components of the logbook

Function	Parameter
$F_y(\hat{y}_{n+1})$	γ_y c_1 c_2
$F_z(Z_{n+1}\hat{\gamma})$	Z_{n+1} $\hat{\gamma}$ γ_z d_1 d_2
$Z_{n+1}\hat{\pi}$	$\hat{\pi}$

Table 2: Ratios of RMSPEs for the forecasts with expert adjustment over the forecasts from the model only: The case of one-month lagged changes in the unemployment rate

Parameters				π					
c_1	c_2	d_1	d_2	-2	-1	-0.2	0.2	1	2
Recursive sample									
-2	2	-2	2	1.064	0.995	0.993	1.010	1.078	1.229
-2	2	-1	1	1.239	1.053	1.000	1.005	1.079	1.291
-1	1	-2	2	1.598	1.131	0.999	1.014	1.205	1.747
-1	1	-1	1	2.302	1.335	1.016	1.009	1.298	2.230
0	0	-2	2	1.659	1.148	1.001	1.014	1.215	1.793
0	0	-1	1	2.819	1.484	1.029	1.005	1.367	2.585
0	0	0	0	3.274	1.596	1.033	1.011	1.485	3.053
Rolling sample, $n = 120$									
-2	2	-2	2	1.262	1.037	0.992	1.015	1.151	1.491
-2	2	-1	1	1.499	1.115	1.001	1.009	1.154	1.577
-1	1	-2	2	1.787	1.178	1.001	1.016	1.253	1.937
-1	1	-1	1	2.787	1.465	1.024	1.010	1.396	2.654
0	0	-2	2	1.836	1.195	1.003	1.014	1.251	1.947
0	0	-1	1	3.048	1.545	1.032	1.006	1.414	2.786
0	0	0	0	3.520	1.669	1.039	1.008	1.513	3.207

Table 3: Ratios of RMSPEs for the forecasts with expert adjustment over the forecasts from the model only: The case of one-month lagged changes in the 3-month interest rate

Parameters				π					
c_1	c_2	d_1	d_2	-0.5	-0.2	-0.1	0.1	0.2	0.5
				Recursive sample					
-2	2	-2	2	1.997	1.242	1.095	0.958	0.968	1.313
-2	2	-1	1	1.996	1.236	1.091	0.963	0.980	1.356
-1	1	-2	2	2.326	1.308	1.117	0.957	0.988	1.526
-1	1	-1	1	2.425	1.325	1.122	0.960	1.001	1.614
0	0	-2	2	2.554	1.346	1.127	0.965	1.023	1.747
0	0	-1	1	2.684	1.367	1.132	0.970	1.042	1.870
0	0	0	0	2.885	1.416	1.152	0.961	1.035	1.933
				Rolling sample, $n = 120$					
-2	2	-2	2	1.975	1.240	1.095	0.955	0.960	1.274
-2	2	-1	1	2.077	1.249	1.094	0.966	0.993	1.436
-1	1	-2	2	2.326	1.315	1.120	0.957	0.989	1.544
-1	1	-1	1	2.751	1.386	1.140	0.964	1.034	1.872
0	0	-2	2	2.408	1.318	1.118	0.963	1.008	1.633
0	0	-1	1	2.808	1.390	1.139	0.972	1.055	1.971
0	0	0	0	3.073	1.457	1.167	0.958	1.039	2.029

References

Belton, V. and P. Goodwin (1996), Remarks on the application of the analytic hierarchy process to judgemental forecasting, *International Journal of Forecasting*, 12, 155-161.

Blattberg, R.C. and S.J. Hoch (1990), Database models and managerial intuition: 50% model + 50% manager, *Management Science*, 36, 887-899.

Clemen, R.T. (1989), Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting*, 5, 559-583.

Clements, M.P. and D.F. Hendry (1998), *Forecasting Economic Time Series*, Cambridge: Cambridge University Press.

Donihue, M.R. (1993), Evaluating the role judgement plays in forecast accuracy, *Journal of Forecasting*, 12, 81-92.

Dyer, J.S. (1990), Remarks on the analytic hierarchy process, *Management Science*, 36, 249-258.

Flores, B.E., D.L. Olson and C. Wolfe (1992), Judgmental adjustment of forecasts: A comparison of methods, *International Journal of Forecasting*, 7, 421-433.

Franses (2004), Do we think we make better forecasts than in the past? A survey of academics, *Interfaces*, 34, 466-468.

Goodwin, P. (2002), Integrating management judgment and statistical methods to improve short-term forecasts, *Omega*, 30, 127-135.

Saaty, T.L. (1980) *The Analytic Hierarchy Process*, New York: McGraw-Hill.

Saaty, T.L. (1990), An exposition of the analytic hierarchy process in reply to Dyer (1990), *Management Science*, 36, 259-268.

Salo, A.A. and D.W. Bunn (1995), Deomposition in the assessment of judgemental probability forecasts, *Technological Forecasting and Social Change*, 49, 13-25.

Timmermann, A. (2006), Forecast combinations, in Graham Elliott, Clive Granger and Allan Timmermann (eds.), Handbook of Economic Forecasting, Amsterdam: Elsevier (2006), in print.

Turner, D. (1990), The role of judgement in macroeconomic forecasting, *Journal of Forecasting*, 9, 315-346.

Van Dijk, D.J.C. and P.H. Franses (2003), Selecting a nonlinear time series model using weighted tests of equal forecast accuracy, *Oxford Bulletin of Economics and Statistics*, 65, 727-744.

Webby, R. and M. O'Connor (1996), Judgemental and statistical time series forecasting, *International Journal of Forecasting*, 12, 91-118.

Wolfe, C. and B. Flores (1990), Judgemental adjustment of earnings forecasts, *Journal of Forecasting*, 9, 389-405.