Evaluating the Rationality of Managers' Sales Forecasts

Bert de Bruijn
Philip Hans Franses*

Econometric Institute
Erasmus University Rotterdam

November 14, 2011

Econometric Institute Report 2011-36

Abstract

This paper deals with the analysis and evaluation of sales forecasts of managers, given that it is unknown how they constructed their forecasts. Our goal is to find out whether these forecasts are rational. To examine deviations from rationality, we argue that one has to approximate how the managers could have generated the forecasts. We describe several ways to construct these approximate expressions. The analysis of a large set of a single manager's forecasts for sales of pharmaceutical products illustrates the practical usefulness of our methodology.

Keywords: Evaluating forecasts, Sales forecasts, Rationality, Intuition, Fixed-event forecasts, Forecast updates.

JEL classifications: C22, C53.

*Econometric Institute, Erasmus School of Economics, PO Box 1738, NL-3000 DR Rotterdam, The Netherlands, `email: franses@ese.eur.nl` (corresponding author)

1 Introduction

Managers' sales forecasts often include a subjective component (Sanders and Manrodt, 2003). This subjective component could be the source of biases or inefficiency but could also improve the forecasts using information that a model could not incorporate. For this reason there is much literature on evaluating forecasts which are not purely statistical, but partly or fully judgmental (for an early review, see Webby and O'Connor (1996), while Lawrence, et al. (2006) provide a more recent review). For sales forecasting, most attention has been given to judgmental adjustments of statistical forecasts (Mathews and Diamantopoulos, 1986, 1989, 1990, 1992; Sanders and Ritzman, 1995, 2001; Fildes et al. 2009; Franses and Legerstee, 2009, 2010; Syntetos et al. 2009; Davydenko et al. 2010a, 2010b; and Eroglu and Croxton, 2010). The literature on judgmental forecasts has also addressed how to combine them with statistical forecasts (Sanders and Ritzman, 2004; Franses and Legerstee, 2011, among others).

In this paper we will evaluate managers' sales forecasts without knowing how these forecasts were created. As we do not have statistical model-based forecasts, evaluating deviations from statistical forecasts is not an option, and hence another approach is needed. We propose various methods in this paper.

To illustrate, we examine the sales forecasts created by a manager of an international pharmaceutical company for hundreds of products for one country. These forecasts are the responsibility of a single individual only and it is unknown whether he has used a statistical model as a basis for these forecasts or whether these forecasts are entirely judgmental. We investigate whether these forecasts can be considered as rational, where rationality implies that the forecaster is able to optimally use available information given a certain loss function.

The remainder of this paper is as follows. Section 2 contains a literature review. Section 3 puts forward some theory of rational forecasts and of possible deviations from rationality. In Section 4, we illustrate this theory for the sales forecasts of the pharmaceutical company. Section 5 contains the conclusion, discussion and suggestions for further research.

2 Literature review

The forecasts that we evaluate are so-called fixed-event forecasts, that is, they concern forecasts made at origins $t - h$ to $t - 1$ for horizon t. Nordhaus (1987) introduced a test for forecast rationality for fixed-events which is based on weak efficiency. Clements (1997) considers the pooling of this test across multiple realization dates and he documents substantial negative autocorrelation in series of revisions. This is consistent with the negative first-order autocorrelations of the sales forecast revisions found by Lawrence and O'Connor (2000).

One prominent reason for inefficient or irrational forecasts concerns the inclusion of judgment. For several decades researchers have been interested in forecasts with a judgmental component (see Adam and Ebert, 1976; Armstrong, 1983; Lawrence et al. 1986; Lawrence and Makridakis, 1989; O'Connor et al. 1993; Lim and O'Connor, 1995; and Goodwin, 2002) For more complete reviews of the literature see Webby and O'Connor (1996) and Lawrence et al. (2006).

Most recent literature has focused on subjective adjustments of statistical forecasts. For example, Franses and Legerstee (2009) discuss properties of managers' adjustments of model forecasts and find that the adjustments are frequent and often upwards. They also find that the adjustments are predictable using earlier adjustments and earlier forecast errors, and that the adjustments are not independent of current model-based forecasts. Franses and Legerstee (2010) find that expert-adjusted forecasts of sales of stock-keeping units (SKU) are not better and often worse than model-based forecasts. On the other hand, Fildes et al. (2009) report that adjustments can improve forecasting performance, although errors are often made when making small or upward adjustments.

Less attention has been given to the analysis of sales forecast revisions, which are the changes of the forecasts for the same realization date when the forecast horizon decreases. While there have been several studies on forecast revisions for macroeconomic variables (Ashiya, 2003) and financial variables (Amir and Ganzach, 1998), we are aware of only one study that deals with SKU-level sales. Lawrence and O'Connor (2000) find that the accuracy improves only little over time and that the revisions have negative first-order autocorrelations.

Additional to evaluating their actions directly using actual forecasts data, another way to find out what managers drives in making their forecasts is by simply asking what they do. This has led to many surveys concerning sales forecasts over many decades (Modigliani and Weingartner, 1958; Cerullo and Avila, 1975; Dalrymple, 1987; Sanders and Manrodt, 1994, 2003; Klassen and Flores, 2001; Boulaksil and Franses, 2009). The general conclusion is that the managers feel that they sometimes need to adjust or ignore statistical model forecasts in order to make a proper forecast. This is despite the fact that there are many studies suggesting that these actions can also have large negative effects on forecasting performance. This suggests that judgment may lead to deviations from rationality.

The novelty of the present study is that we propose a methodology to elicit deviations from rationality of managers' forecasts in case only their forecasts are available. For this purpose we construct artificial model-based forecasts, and with these we can compute the adjustments (deviations). If the constructed model forecasts are similar to the unknown factual model forecasts, the constructed adjustments should have properties similar to the typical adjustments as found in judgmental adjustment literature.

3 Theory

In this section we will discuss some theory on forecasts and on the evaluation of forecasts. We will discuss properties that define a rational forecast in Section 3.1. In practice, there might be a deviation from a rational forecast. Section 3.2 discusses the models and methods that we will use to reconstruct this deviation in the case the analyst, who should evaluate the quality of managers' forecasts, does not have access to statistical model-based forecasts.

3.1 What is a rational forecast?

Suppose the variable to be predicted is y_t. At an earlier time, say h periods earlier, a forecast is made. Denote this forecast as $\hat{y}_{t|t-h}$.

Forecast errors

The forecast will in general not be equal to the realization at time t, and there will be a forecast error $f_{t|t-h} = y_t - \hat{y}_{t|t-h}$. Ideally, one would like the forecast to be unbiased, which is defined as the error having a mean equal to zero, that is, $E(f_{t|t-h}) = 0$.[1]. This is desirable in an unconditional sense (overall unbiasedness) and in a conditional sense (unbiasedness in typical situations). Also, one wants the variance $\sigma^2_{f_{t|t-h}}$ to be as small as possible. If the forecast is truly incorporating all relevant information, this forecast error variance should preferably be smaller than the forecast error variance of any other forecast. As one could always re-use the forecast of one horizon earlier, the new forecast should also have a variance equal to or smaller than the last forecast, that is, $\sigma^2_{f_{t|t-h}} \leq \sigma^2_{f_{t|t-h-1}}$. Another way to phrase this is that forecasts should improve in accuracy over time.

The speed of this improvement depends on the variable under consideration. For example, consider an AR(1) process with parameter ϕ and unconditional mean μ, that is, $y_t = \mu + \phi(y_{t-1} - \mu) + \varepsilon_t$. Assuming that ϕ and μ are known, the error series $..., \varepsilon_{t-1}, \varepsilon_t$ is the only source of forecasting error, which means that the smaller the forecast horizon, the more information is known to improve upon the unconditional forecast μ. The development of this information over the forecast horizon is introduced by Galbraith (2003) as forecast content, with an application in for example Galbraith and Tkacz (2007). The main pattern is that the information gain per horizon increases as the evaluation date approaches, which is due to the fact that later forecast origins often contain more information. Finally, errors that do not have an overlap between forecasting event and origin should be uncorrelated, such as $f_{t|t-h}$ and $f_{t-h-k|t-h-p}$ for any $0 < k < p$. Overlapping errors can be correlated, as they do not include the same news, such as $f_{t|t-2}$ and $f_{t-1|t-3}$, which both do not have the error (news) at $t-1$.

[1] This is the case if one uses a symmetric and quadratic error function. There could be situations where one could consider alternative loss functions, but this is not pursued here.

Forecast revisions

Interesting for our application are the properties of the forecast revisions[2] defined by $\delta_{t|t-h} = \hat{y}_{t|t-h} - \hat{y}_{t|t-h-1}$ made at $t-h$ for event t. If the forecast $\hat{y}_{t|t-h-1}$ is assumed to be unbiased, upward revisions should occur just as often as downward revisions, given a symmetric loss function. This does not only hold for revisions made at adjacent origins, but also for more separated origins. Moreover, the probability of an upward revision should not depend on the last revision, and should always be the same. The same holds for the size of the revision (irrespective of the sign), that is, it should not depend on the sign or size of the earlier revisions[3]. This is similar to the absence of correlation between the forecast updates.

Forecast errors and revisions combined

There are also some interesting properties of the cross-correlations of the errors and the revisions in case of rational forecasts. The following discussion is summarized in Table 1. First, in the case of a rational forecast, there should not be a correlation between the revision $\delta_{t|t-h}$ and the error of its associated forecast $f_{t|t-h}$, as in that case the revision could have been improved. The same holds for the errors of all forecasts on or after the revision date. Thus, $\delta_{t|t-h}$ is uncorrelated with $f_{t-k|t-h+p}$ for all $0 \leq k \leq h-1$ and $0 \leq p \leq h-k-1$. Also, forecast errors concerning a realization before the revision moment of the current revision should not correlate with the current revision, which means that $\delta_{t|t-h}$ is also uncorrelated with $f_{t-h-k|t-h-p}$ for all $0 < k < p$. On the other hand, the errors that have just been known on $t-h$ can be rationally correlated with $\delta_{t|t-h}$, as they can contain relevant information that has not been used already. Thus, $\delta_{t|t-h}$ can be correlated with $f_{t-h|t-h-p}$ for all $p > 0$.

In Appendix A, we will give more details of the rationality properties along with more formal statistical notation. The properties are illustrated for a stationary time

[2]Forecast revisions are also known as forecast updates. Both terms will be used in this paper.

[3]It could depend on the forecast horizon as more information is contained in the last origin before realization. This could result in those revisions being larger in (absolute) size due to the fact that they contain more information. The revisions could also be smaller as the cumulative information until that moment also has increased, making the new information a smaller piece of the total amount of available information.

series model.

3.2 Reconstructing deviations from rational forecasts

Using managers' sales forecasts only, we can calculate relevant sample statistics and evaluate whether there is a deviation from rationality. However, we cannot quantify the size of the deviation from rationality, as the rational (model-based) forecasts are not available. Our solution is to create artificial statistical forecasts for which the rational properties should hold. These artificial forecasts can be given the interpretation of model forecasts, resulting in a bridge between the deviations from these artificial forecasts (which will be given the interpretation of *intuition* later on) and the judgmental adjustments of statistical forecasts discussed in the literature review. When the artificial forecasts are similar to the statistical forecasts, intuition should have similar properties as judgmental adjustment. It would then be plausible that this is the approach that could have been taken by the manager, that is, first to create model forecasts and then to adjust them.

Notation

Again, denote realizations and managers' forecasts as y_t and $\hat{y}_{t|t-h}$. Now suppose there are also model forecasts $z_{t|t-h}$ and denote the difference as

$$\nu_{t|t-h} = \hat{y}_{t|t-h} - z_{t|t-h}, \tag{1}$$

which can be interpreted as intuition. Rational forecasts imply that the updates are unbiased and uncorrelated, which leads to $\alpha = 0$ and $\beta = 0$ in the auxiliary regression

$$\delta_{t|t-h} = \alpha + \beta\delta_{t|t-h-1} + w_{t,h}. \tag{2}$$

Suppose we assume that intuition $\nu_{t|t-h}$ obeys

$$\nu_{t|t-h} = \lambda\nu_{t|t-h-1} + \eta_{t,h} \tag{3}$$

(with $\eta_{t,h}$ iid with mean 0 and variance σ_ν^2), then Appendix B shows that $\beta = \frac{\lambda-1}{2}$, or equivalently $\lambda = 2\beta + 1$. For β to be 0, λ must be 1. The λ can also take some other interesting values. First, if $\lambda = 0$ (equivalent with $\beta = -\frac{1}{2}$), the intuition is uncorrelated with past intuition, meaning that previous intuition is immediately discarded. If $\lambda = -1$ ($\beta = -1$), the intuition of the previous update is reversed on average, which would result in highly varying intuition over time.

Creating artificial forecasts

In the common case when there are not many data points, we would suggest the following five alternative ways of defining or estimating either $z_{t|t-h}$ or $\nu_{t|t-h}$. Approach A defines $z_{t|t-h}$ to be the fit of $y_t = \alpha_h + \beta_h \hat{y}_{t|t-h} + u_{t,h}$. Approach B simply takes the Random Walk (RW) forecast and thus assumes $z_{t|t-h} = y_{t-h}$. The third model forecast (C) is a slight variation, in which a possible temporary swing is incorporated into the model, that is, $z_{t|t-h} = \pi y_{t-h} + h(1-\pi)(y_{t-h} - y_{t-h-1})$. The π parameter indicates how important the last realization is, as compared to the trend information in the forecast. To avoid too much influence of temporary swings, this parameter has to be close to 1. We will use $\pi = 0.95$ in our empirical work below.

Before defining the final two forecasts, we need a few derivations. As the $z_{t|t-h}$ are supposed to be rational, changes in model forecasts should not be predictable from earlier forecasts, that is,

$$z_{t|t-h} = z_{t|t-h-1} + \omega_{t,h}$$

with $\omega_{t,h}$ having white noise properties. Rewriting (1) to $z_{t|t-h} = \hat{y}_{t|t-h} - \nu_{t|t-h}$ gives

$$\hat{y}_{t|t-h} - \nu_{t|t-h} = \hat{y}_{t|t-h-1} - \nu_{t|t-h-1} + \omega_{t,h},$$

or, after re-arranging

$$\hat{y}_{t|t-h} - \hat{y}_{t|t-h-1} = \nu_{t|t-h} - \nu_{t|t-h-1} + \omega_{t,h}.$$

The left hand side is equal to $\delta_{t|t-h}$. Adding (3) results in the following set of equations:

$$\begin{aligned} \delta_{t|t-h} &= \nu_{t|t-h} - \nu_{t|t-h-1} + \omega_{t,h} \\ \nu_{t|t-h} &= \lambda \nu_{t|t-h-1} + \eta_{t,h} \end{aligned} \tag{4}$$

For the fourth model forecast (D) we will fix λ at $2\beta+1$, where β is estimated using (2). For the fifth and final model forecast (approach E), λ is estimated along with the rest of the model. In contrast to the earlier approaches, these last two methods directly result in $\nu_{t|t-h}$, of which then $z_{t|t-h}$ can be calculated.

Both the actual managers' forecasts and the constructed model forecasts can then be evaluated on their rational properties. Also, the resulting λ can be compared for the different model forecasts. Finally, the deviations between actual forecasts and the model forecasts can be further investigated. For example, it is interesting to check whether these estimates of intuition have similar properties as the judgmental adjustments documented in the judgmental adjustment literature, such as for example in Franses and Legerstee (2009, 2010).

4 Illustration

This section applies our methodology to a large and novel dataset of managers' forecasts. First, we discuss the data and then we check rationality properties. After that, we create artificial model forecasts and hold these against the managers' forecasts.

4.1 Data

We use data from a large multinational Germany-based pharmaceutical company. This company has a manager for each country who produces sales forecasts for hundreds of products. We have data for a single country with forecasts up to one year ahead. This results in data of the form of Table 2, with 88 forecasts and 11 realizations for each product (at the stock keeping unit SKU level). We will pool all data after standardization.

One characteristic of the data set is that the sales volumes (and their forecasts) vary per product. We therefore standardize the data as follows. We first calculate per product the averages of the forecasts and of the realizations and take an equal-weighted average of this as the standardized mean, such that forecast information and realization information is equally important. We also calculate the standard deviation of both sets and take the square root of the product of these as the standardization standard deviation. Finally, we calculate the standardized forecasts and realizations by subtracting the standardized mean from each original value and by dividing by the standardized standard deviation: $y^* = \frac{y-\mu_S}{\sigma_S} = \frac{y-\frac{1}{2}\mu_r-\frac{1}{2}\mu_f}{\sqrt{\sigma_r\sigma_f}}$, with subscripts S, r and f indicating the standardization, realization or forecast variables. Standardization approaches based on only the forecast information or the realization information yield similar results. In the following, we will use forecast and error to denote standardized forecast and standardized error.

4.2 Rational properties

Table 3 shows descriptive statistics of the forecast errors for different horizons. For each horizon, the mean error is significantly smaller than zero, which indicates that the forecasts are on average too high and thus biased. The errors approach zero as the time to the event decreases. The mean squared error (and similarly the unreported variance) also declines as the forecast horizon decreases, indicating that the forecasts become more accurate as the realization date gets nearby. We also present the forecast content as defined by Galbraith (2003). This shows a different pattern than what is usually found. Usually, the forecast content quickly drops when forecast origin and event are more distant, but here, the decline only really occurs for horizons larger than 10. This suggests that about half of the deviation from the unconditional sales mean is already known by the forecaster about 10 months before the sales have realized. After that, the forecast accuracy increases only a little, which is consistent with the results of Lawrence and O'Connor (2000). The short horizons have more observations, which is due to the data format displayed in Table 2, but this does not seem to influence the results (as we have verified with alternative samples, but in unreported tables).

Evaluating the correlation of the errors can be done in many different ways, as the features that can be held as constant include forecast horizon, realization moment and forecast moment. Table 4 shows the correlation between forecast errors for the same realization date across the first eight horizons. It can be seen that the closer the two events are to each other, the closer the correlation is to 1, which makes sense under rationality as less distant forecasts share a larger part of their unknown error.

Table 5 shows the correlation between forecast errors corresponding to the same horizon for different realization dates. For example, the entry in the third column and fourth row (0.216) gives the correlation between all 2-step-ahead forecast errors with a difference of 3 months in realization dates. Hence, on average, a positive 2-months-ahead forecast error is followed by a positive 2-months-ahead forecast error 3 months later.

Revision properties

From Table 6 it can be seen that the average forecast update is downward as (almost) all mean revisions are negative. The shape of the mean revision over the horizons is similar to a parabola. On average, the largest revisions occur for the middle horizons, while the revisions one year before realization and the final revisions are not significantly different from zero. A similar pattern can be seen in the variance of the revision. There is hardly any final non-zero variance, while the revision at horizon 7 has a variance of 0.322.

Table 7 contains statistics relating to the direction of the revision. It can be seen that most revisions are downward. Table 8 contains the number of revisions in a certain direction after a revision in any other direction has occurred. This shows that positive updates are more often followed by negative updates (since $239 > 86$), while negative updates are more often followed by positive updates ($249 > 95$). This means that the distribution of the sign of the revision seems dependent on the earlier updates.

Concerning the autocorrelation of the revisions, consider Table 9. This table shows the autocorrelation and cross-correlation of the revisions and the squared revisions for various lags. It can be seen that the first-order autocorrelation of the revisions is negative, which is consistent with the results in Table 8. This thus violates the rational assumption (see equation 13 in Appendix A) but is consistent with the results of Lawrence and O'Connor (2000). There is also a large negative contemporaneous correlation between revision and squared revisions, indicating that if the updates are large in size, they are often downward.

A second way of evaluating the autocorrelation of the revisions results in the numbers in Table 10, where we display the correlations between forecast revisions for various realization dates but the same forecast horizon. Under rationality, all cells in this table should be zero, as none of the paired revisions are made simultaneously and the revisions should not be predictable over time. It can be seen that the one-step-horizon revisions contain substantial autocorrelation, averaging at 0.170, which is in conflict with rationality. If the current one-step revision is known, future one-step revisions can be predicted from it, and this should not be the case under rationality.

The relationship between errors and revisions

Table 11 contains the cross-correlations between the revisions and the errors for the same realization. It can be seen that the contemporaneous correlation is not large, and this is also the case for the cross-correlation with a lagged error. They are slightly negative for small horizons, and this suggests that a positive revision now is paired with a negative error, not only for this forecast horizon, but also for upcoming horizons. This makes sense as a positive revision makes the forecasts larger as compared to the realization, and the forecast errors are defined as the realization minus the forecast. The cross-correlations in which the error leads the revision are far away from zero. This indicates that a current large error results in an upward revision in the future, or vice versa. This is consistent with more information gradually becoming more used, although this also means that the information is smoothed and not used immediately when it is available.

Table 11 also contains estimates of θ_h as in (12), using the one-step-forecast error $f_{t-h|t-h-1}$ as an approximation for the shock ε_{t-h} (which is valid under the assumption of rationality). As the estimate of θ_1 is not significantly different from zero, the results suggest that the latest news before realization has no influence on the realization of the sales. All other θ's are significantly different from 0 (at the 1% level), indicating that the shocks at these horizons have a significant effect on current sales. The sum of these parameters is 0.811, which can only be interpreted as a lower bound of the total future effect of a shock as for horizons above 10 the θ's might also be significantly different from zero.

In sum, of all the rational properties discussed in section 3.1 and Appendix A, only (8) and (14) stand up tall in the data. We can thus safely conclude that the manager's forecasts are not rational (under a symmetric loss function). Now it is interesting to study how deviations from rationality look like.

4.3 Artificial model forecasts and manager's forecasts

Method A, in which we define $z_{t|t-h}$ to be the fit of $y_t = \alpha_h + \beta_h \hat{y}_{t|t-h} + u_{t,h}$, is straightforward to apply. Using Ordinary Least Squares (OLS) to fit this model for each h, we then calculate the $\nu_{t|t-h}$ as the difference of $z_{t|t-h}$ and $\hat{y}_{t|t-h}$. Interestingly, most of the β's are estimated to be negative (and significantly different from zero), indicating that positive forecasts correspond to negative realizations and vice versa. We check if a few outliers are the culprit of these findings, but when we apply a robust estimation method (Least Absolute Deviation or Iteratively Reweighed Least Squares) we obtain negative β's of similar magnitude.

Methods B and C create artificial forecasts $z_{t|t-h}$ using actual sales data only and ignore information in the manager's forecasts. Note that both methods imply a loss of information at the maximum horizon length. Indeed, we only have 11 realizations per product, and hence the maximum horizon for these methods is 10.

Methods D and E are based on the two equations in (4). Method D assumes that λ can be fixed at $2\beta + 1$, with β estimated using (2), while approach E estimates λ. For the first forecast of each event, we do not have an earlier forecast and thus no earlier intuition, so we will also estimate the first $\nu_{t|t-k}$ (with k the maximum forecast horizon for that t). For that, we impose an identical unconditional distribution for all products. The estimation process for these methods is done using OpenBUGS 3.2.1, which allows for Bayesian estimation, as an alternative method would be difficult due to the thousands of $\nu_{t|t-h_{max}}$ that have to be estimated. This Bayesian method directly results in a distribution of all $\nu_{t|t-h}$ and, in the case of approach E, of λ. We have only used flat priors. The starting values for the precision have been varied and that did not influence the results. For both methods, 10000 iterations have been used, after discarding the first 1000 iterations for convergence. We have used the means of the posterior distributions as estimation values, neglecting the rest of the distribution information. Methods D and E are different from the other three in that they directly model $\nu_{t|t-h}$ instead of $z_{t|t-h}$.

Results

The five methods (A to E) are applied to construct both model forecasts $z_{t|t-h}$ and the intuition $\nu_{t|t-h}$. Table 12 shows the estimated λ parameters. The first three are based on the estimation of (3), the fourth on $\lambda = 2\beta + 1$ with β as in $y_t = \alpha_h + \beta_h \hat{y}_{t|t-h} + u_{t,h}$ and using the Delta method (Greene, 2002) for the confidence intervals and the fifth on the posterior distribution of λ after estimating the set of equations (4). Clearly, all estimates are significantly different from both 0 and 1, implying that the manager's past intuition is not immediately discarded and that future intuition is predictable.

Figure 1 shows the mean forecast error for each approach, together with the forecast error of the original forecasts as reported in Table 3. The errors for approach A are all zero by construction, as that is a property of OLS. What can be seen is that each modeling approach results in forecast errors that are smaller than those of the original manager's forecasts. For short horizons the RW forecasts have smaller errors than the dynamic equations forecasts, but for more distant horizons the situation is reversed. Approach B has smaller errors than approach C, and approach E has smaller errors than approach D. But still, all errors are significantly different from zero, meaning that even these forecasts violate a rationality property. Figure 2 shows a similar picture for the variance of the forecast errors. Three methods (A, D and E) result in a forecast error variance which is always smaller than the original manager's variance, although only for method A this difference is significant. Approaches B and C sometimes have a larger variance and sometimes a significantly smaller variance. Mostly, the variance seems to decline as the horizon gets close.

In sum, the constructed model forecasts can also not be considered as fully rational forecasts. This could be due to our simple intuition specification, due to not allowing for enough flexibility to properly model the intuition process, or due to excluding some of the intuition correlation in the artificial model forecasts.

It is evident that approach A results in the 'best' forecasts, but this is not surprising considering it makes direct use of information that is only available after the realizations are available. This means that the results for approach A should be interpreted as an upper bound (or lower bound, depending on the statistic that is being discussed). Methods B and C are in general applicable, as they are only based on available data. Their downside is that they do not directly model the behavior of the forecaster. Methods D and E are also fairly applicable beforehand, assuming one knows the λ and has an idea of the initial bias. More importantly, they are directly based on the behavior of the forecaster, as only the δ's are used in the process of estimation. Of these two, approach E appears the most realistic as it estimates λ jointly with the intuition $\nu_{t|t-h}$. This is why we will further analyze this approach in what follows.

4.4 Interpretation of intuition as judgmental adjustment

In the judgmental adjustment literature, several properties of adjustments have been documented. We will now compare the properties of the estimated intuition of approach E with the adjustment properties as found by Franses and Legerstee (2009, 2010).

First, these authors document that managers often adjust statistical forecasts. They report an overall adjustment percentage of 89.5%. In other words, the judgmental adjustment of statistical forecasts is often found to be non-zero. Here, with method E this is also the case, but of course this is due to how intuition is estimated. The probability that the simulated values of the intuition is on average exactly zero is zero. When we use a threshold of 0.05, that is, only adjustments that are (in an absolute sense) larger than this value are considered as non-zero adjustments, we obtain an average adjustment percentage of 98.8%. Raising the threshold to 0.5 results in an overall adjustment percentage of 57.1%, meaning that more than half of the adjustments have a size of at least as large as half the standard deviation of the sales and the corresponding forecasts, which is due to the standardization procedure discussed in Section 4.1. Evidently, intuition is thus often a large component of the manager's final forecast. In order to get a similar adjustment percentage as in Franses and Legerstee (2009), the threshold has to be put equal to 0.285, and this threshold will be used to study other statistics of intuition.

Franses and Legerstee (2009) report that for 53.5% of the forecasts, the manager adjusts upwards. Correcting for non-zero adjustments, this percentage increases to 59.8%, showing that there are almost 50% more upward adjustments than downward adjustments. Using the threshold 0.285, we find that intuition is positive in 88.8% of the cases and for the non-zero adjustments this increases to 99.2%. Hence, the manager under scrutiny has a strong tendency to have positive-valued intuition.

Franses and Legerstee (2009) also evaluated the predictability of adjustments using the history of the adjustments and of the statistical forecasts. They found an average R^2 of 44.3% in regressions using data until the 7th lag. Due to our short history of forecasts we cannot replicate their findings. However, when we use one lag as in $\nu_{t|t-h} = \lambda_0 + \lambda_1 \nu_{t|t-h-1} + \lambda_2 \nu^2_{t|t-h-1}$ we obtain an average R^2 of 68.5%, indicating there is a strong predictability of intuition, which corresponds with the results on λ in Table 12.

Next, Franses and Legerstee (2009) report a negative correlation between adjustment and model-based forecast. In the regression $\hat{y}_{t|t-h} = \alpha + \gamma z_{t|t-h} + u_{t|t-h}$ they document that γ is 0.424 on average, and that this is significantly different from both 0 and 1. In our situation there are three ways to estimate γ: 1. $\hat{y}_{i,t|t-h} = \alpha + \gamma z_{i,t|t-h} + u_{i,t|t-h}$, with a single γ for all products i; 2. $\hat{y}_{i,t|t-h} = \alpha + \gamma_i z_{i,t|t-h} + u_{i,t|t-h}$, with a γ per product; 3. $\hat{y}_{i,t|t-h} = \alpha + \gamma_{i,t} z_{i,t|t-h} + u_{i,t|t-h}$, with a γ per product per forecast event. For the second approach we consider only the one-step-ahead forecasts ($h = 1$). For all three approaches we find (the average) γ to be around 1.05 to 1.09 and significantly different from 1 (and 0). This means that γ^* in $\nu_{t|t-h} = \alpha + \gamma^* z_{t|t-h} + u_{t|t-h}$ is around 0.05 to 0.09 and significantly different from 0, indicating that if the statistical forecast is large and positive, the intuition is upward, suggesting double-counting. This is in contrast with the result of Franses and Legerstee (2009), for which the γ^* on average is around $0.424 - 1 = -0.576$, which for their data means that if the statistical forecasts are large and positive, the adjustments are downward, thus dampening the forecast.

It is also interesting to study the effect of current intuition on future intuition. Concerning forecast adjustments, Franses and Legerstee (2009) reported an average persistence of 0.773. In our case, persistence of intuition is measured by the value of λ, which for approach E is equal to 0.919, which is quite close to 0.773.

Finally, Franses and Legerstee (2010) reported ρ in $y_t = \alpha + \beta z_{t|t-1} + \rho(\hat{y}_{t|t-1} - \gamma z_{t|t-1}) + u_{t,1}$ to be 0.247, on average, which entails that there is significant forecast information in the adjustments (where γ is defined to be the correlation between the model forecasts and the managers' forecasts, as above). Here, using the results for γ as earlier, we find ρ to be equal to 0.898 and significantly different from zero (but not from 1), indicating that the forecasting contribution of the information in the intuition $\nu_{t|t-h}$ is quite large. Interestingly, the β is found to be almost zero, implying that the model forecasts do not provide extra information additional to the intuition. The R^2 of the regression is 2.85%, while the R^2 of the regression without the intuition is only 0.01%.

All above results are summarized in Table 13. We note that the constructed intuition seems to have properties similar to the properties of judgmental adjustments as reported by Franses and Legerstee (2009, 2010). Most percentages and parameters deviate from the rational norm in the same direction, with γ being the only exception.

5 Conclusion

We have proposed a methodology to analyze managers' forecasts when only these forecasts are given and it is unknown how managers created their forecasts. We illustrated this methodology for a large range of forecasts for a single manager, but for hundreds of products. Generally, we conclude that these sales forecasts violate rationality. Deviations from model forecasts have been constructed using five different methods and these variables seem to have properties similar to those reported in the judgmental adjustment literature. Hence, we conjecture that it is plausible that the manager's forecasts have been constructed by adjusting available model forecasts. Our artificial model forecasts turned to lead to smaller forecast errors. So, also here we see that a manager's adjustment of model forecasts does not necessarily lead to better forecasts.

Managers might have non-standard beliefs and might be confronted with non-standard decision making situations, and this may lead to irrational forecasts. In the future, more research on this topic can be done given the availability of larger data sets containing more managers, more products per country and a longer time span per product. This would allow a comparison of forecasting practices.

Our methodology cannot directly be used to improve managers' forecasts. What could be done is to inform managers about the properties of their forecasts, and to prevent them from making systematic errors in their adjustments or revisions. By simply showing the biases the managers appear to have, they might be able to (partially) change their behavior. Another possibility is to use the information of their forecasts in a different way and to incorporate the adjustments or revisions of the forecasters into the model forecast. This combined forecast benefits from the extra information that the forecasters might have and the unbiasedness of a statistical model.

References

Adam, E.E. and R.J. Ebert (1976). A comparison of human and statistical forecasting. *AIIE Transactions*, 8, 120-127

Amir, E. and Y. Ganzach (1998). Overreaction and underreaction in analysts' forecasts. *Journal of Economic Behavior & Organization*, 37, 333-347

Armstrong, J.S. (1983). Relative accuracy of judgemental and extrapolative methods in forecasting annual earnings. *Journal of Forecasting*, 2, 437-447

Ashiya, M. (2003). Testing the rationality of Japanese GDP forecasts: the sign of forecast revision matters. *Journal of Economic Behavior & Organization*, 50, 263-269

Boulaksil, Y. and P.H. Franses (2009). Experts' stated behavior. *Interfaces*, 39, 168-171

Cerullo, M.J. and A. Avila (1975). Sales forecasting practices: a survey. *Managerial Planning*, September/October, 33-39

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

Dalrymple, D.J. (1987). Sales forecasting practices: results from a United States survey. *International Journal of Forecasting*, 3, 379-391

Davydenko, A., R. Fildes and J.T. Arenas (2010). Judgmental adjustments to demand forecasts: accuracy evaluation and bias correction. *Working paper 2010/003, Lancaster University Management School*

Davydenko, A., R. Fildes and J.T. Arenas (2010). Measuring the accuracy of judgmental adjustments to SKU-level demand forecasts. *Working paper 2010/026, Lancaster University Management School*

Eroglu, C. and K.L. Croxton (2010). Biases in judgmental adjustments of statistical forecasts: the role of individual differences. *International Journal of Forecasting*, 26, 116-133

Fildes, R., P. Goodwin, M. Lawrence and K. Nikolopoulos (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, 25, 3-23

Franses, P.H. and R. Legerstee (2009). Properties of expert adjustments on model-based SKU-level forecasts. *International Journal of Forecasting*, 25, 35-47

Franses, P.H. and R. Legerstee (2010). Do experts adjustments on model-based SKU-level forecasts improve forecast quality? *Journal of Forecasting*, 29, 331-40

Franses, P.H. and R. Legerstee (2011). Combining SKU-level sales forecasts from models and experts. *Expert Systems with Applications*, 38, 2365-2370

Galbraith, J.W. (2003). Content horizons for univariate time-series forecasts. *International Journal of Forecasting*, 19, 43-55.

Galbraith, J.W. and G. Tkacz (2007). Forecast content and content horizons for some important macroeconomic time series. *Canadian Journal of Economics*, 40, 935-953.

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

Greene, W.H. (2002). *Econometric Analysis. 5th edition*, Upper Saddle River, NJ: Prentice Hall

Klassen, D., B. E. Flores (2001). Forecasting practices of Canadian firms: survey results and comparisons. *International Journal of Production Economics*, 70, 163-174

Lawrence, M., R. Edmundson and M. O'Connor (1986). The accuracy of combining judgemental and statistical forecasts. *Management Science*, 32, 1521-1532

Lawrence, M., P. Goodwin, M. O'Connor, D. Önkal (2006). Judgmental forecasting: a review of progress over the last 25 years. *International Journal of Forecasting*, 22, 493-518

Lawrence, M. and S. Makridakis (1989). Factors affecting judgmental forecasts and confidence intervals. *Organizational Behavior and Human Decision Processes*, 43, 172-187

Lawrence, M and M. O'Connor (2000). Sales forecasting updates: how good are they in practice? *International Journal of Forecasting*, 16, 369-382

Lim, J.S. and M. O'Connor (1995). Judgemental adjustment of initial forecasts: its effectiveness and biases. *Journal of Behavioral Decision Making*, 8, 149-68

Mathews, B.P. and A. Diamantopoulos (1986). Managerial intervention in forecasting. An empirical investigation of forecast manipulation. *International Journal of Research in Marketing*, 3, 3-10

Mathews, B.P. and A. Diamantopoulos (1989). Judgemental revision of sales forecasts: a longitudinal extension. *Journal of Forecasting*, 8, 129-140

Mathews, B.P. and A. Diamantopoulos (1990). Judgemental revision of sales forecasts: effectiveness of forecast selection. *Journal of Forecasting*, 9, 407-415

Mathews, B.P. and A. Diamantopoulos (1992). Judgemental revision of sales forecasts: the relative performance of judgementally revised versus non-revised forecasts. *Journal of Forecasting*, 11, 569-576

Modigliani, F. and H.M. Weingartner (1958). Forecasting uses of anticipatory data on investment and sales. *Quarterly Journal of Economics*, 72, 23-54

Nordhaus, W.D. (1987). Forecasting efficiency: concepts and applications. *The Review of Economics and Statistics*, 69, 667-74

O'Connor, M., W. Remus and K. Griggs (1993). Judgemental forecasting in times of change. *International Journal of Forecasting*, 9, 163-172

Sanders, N.R. and K.B. Manrodt (1994). Forecasting practices in US corporations: survey results. *Interfaces*, 24, 92-100

Sanders, N.R. and K.B. Manrodt (2003). Forecasting software in practice: use, satisfaction and performance. *Interfaces*, 33, 90-93

Sanders, N.R. and L.P. Ritzman (1995). Bringing judgment into combination forecasts. *Journal of Operations Management*, 13, 311-321

Sanders, N.R. and L.P. Ritzman (2001). Judgmental adjustment of statistical forecasts. *In: Armstrong, J.S. (Ed.). Principles of forecasting: a handbook for researchers and practitioners.* Kluwer Academic Publishers, Boston, pp. 405-416.

Sanders, N.R. and L.P. Ritzman (2004). Integrating judgmental and quantitative forecasts: methodologies for pooling marketing and operations information. *International Journal of Operations & Production Management*, 24, 514-529

Syntetos, A.A., K. Nikolopoulos, J.E. Boylan, R. Fildes and P. Goodwin (2009). The effects of integrating management judgement into intermittent demand forecasts. *International Journal of Production Economics*, 118, 72-81

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

Wold, H.O.A. (1954). *A study in the analysis of stationary time series.* Stockholm: Almqvist and Wiksell Book Co.

A Properties of rational forecasts for a stationary time series

We will illustrate the properties of rational forecasts for a stationary time series. As rational expectations $E^r(.)$ are the same as statistical expectations $E(.)$ based on the true model, we will use statistical expectations. Assume that the data x_t are generated by a stationary time series model without a time trend. Also assume that the relevant parameters are known or estimated very accurately. Denote the h-step ahead forecast as $\hat{x}_{t|t-h}$ and the forecast error as $f_{t|t-h} = x_t - \hat{x}_{t|t-h}$. We will use the result of Wold's theorem (Wold, 1954), which states that it is possible to decompose any covariance-stationary time series as the sum of an infinite moving average process (MA(∞)) and a deterministic component. In the case of our stationary model the deterministic component is only the unconditional mean, which we will denote as μ. This means that

$$x_t = \mu + \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}, \tag{5}$$

where we will assume that $\varepsilon_t \sim \mathbb{N}(0, \sigma_\varepsilon^2 I)$ (and thus linearly uncorrelated) and $\theta_0 = 1$. This means that $f_{t|t-h}$ can be defined as

$$f_{t|t-h} = \sum_{i=0}^{h-1} \theta_i \varepsilon_{t-i}, \tag{6}$$

as the errors at origin $t-h$ or older are known.

1. The forecast errors should be unbiased:

$$E(f_{t|t-h}) = 0. \tag{7}$$

Proof. Using Wold's theorem: $E(f_{t|t-h}) = E(\sum_{i=0}^{h-1} \theta_i \varepsilon_{t-i}) = 0$, as the errors all have expectation equal to zero.

2. The forecast accuracy should improve as the horizon decreases:

$$Var(f_{t|t-h}^2) < Var(f_{t|t-k}^2) \tag{8}$$

for $h < k$.

Proof. $Var(f_{t|t-h}) = Var(\sum_{i=0}^{h-1} \theta_i \varepsilon_{t-i}) = \sum_{i=0}^{h-1} \theta_i^2 \sigma_\varepsilon^2$ and similarly $Var(f_{t|t-k}) = \sum_{i=0}^{k-1} \theta_i^2 \sigma_\varepsilon^2$. As $h < k$, $Var(f_{t|t-k}) - Var(f_{t|t-h}) = \sum_{i=h}^{k-1} \theta_i^2 \sigma_\varepsilon^2 > 0$, which means that $Var(f_{t|t-h}^2) < Var(f_{t|t-k}^2)$.

3. Nonoverlapping forecast errors should be uncorrelated:

$$E(f_{t|t-h} f_{t-h-k|t-h-p}) = 0 \tag{9}$$

for any $0 < k < p$.

Proof. $E(f_{t|t-h} f_{t-h-k|t-h-p}) = E(\sum_{i=0}^{h-1} \theta_i \varepsilon_{t-i} \sum_{i=h+k}^{h+p} \theta_i \varepsilon_{t-i}) = 0$, as all relevant ε are uncorrelated.

4. The revisions should be as often upward as downward, irrespective of earlier revisions:

$$P(\delta_{t|t-h} > 0) = P(\delta_{t|t-h} < 0) \tag{10}$$

and

$$P(\delta_{t|t-h} > 0|t-h-k) = P(\delta_{t|t-h} > 0|t-h-p) \tag{11}$$

for any $0 < k < p$.

Proof. $P(\delta_{t|t-h} > 0) = P(\hat{y}_{t|t-h} - \hat{y}_{t|t-h-1} > 0) = P(\theta_h \varepsilon_{t-h} > 0) = \frac{1}{2}$, as $\varepsilon_t \sim \mathbb{N}(0, \sigma_\varepsilon^2 I)$. Similarly, $P(\delta_{t|t-h} < 0) = \frac{1}{2}$. This derivation is the same for any (earlier) information set, as for this model the information set that is known does not affect the extra information that will be known at the moment of the revision. Indeed, at $t-h+1$, the probability $P(\delta_{t|t-h} > 0)$ would not be equal to $\frac{1}{2}$, as at that moment the revision is already known. Before $t-h$ there is no information on the error at $t-h$, irrespective of how far in the future $t-h$ is.

Corollary.

$$\delta_{t|t-h} = \theta_h \varepsilon_{t-h}, \tag{12}$$

which means that the θ's of a Wold's decomposition can be extracted from the data using the revisions and the one-step-ahead forecast errors concerning realizations at $t-h$ (which are proportional to the corresponding shocks ε_{t-h}).

5. The revisions should have no autocorrelation:

$$E(\delta_{t|t-h}\delta_{t|t-h-p}) = 0 \tag{13}$$

for any $p > 0$.

Proof. $E(\delta_{t|t-h}\delta_{t|t-h-p}) = E[(\hat{y}_{t|t-h} - \hat{y}_{t|t-h-1})(\hat{y}_{t|t-h-p} - \hat{y}_{t|t-h-p-1})] = E(\theta_h\theta_{h-p}\varepsilon_{t-h}\varepsilon_{t-h-p}) = 0$, as the ε are independent.

6. The cross-correlations between errors and revisions should be zero for all errors with the time span from forecast origin to event entirely before or after the revision moment:

$$E(\delta_{t|t-h}f_{t-h-k|t-h-p}) = 0 \tag{14}$$

for any $0 < k < p$ and

$$E(\delta_{t|t-h}f_{t-h+p|t-h+k}) = 0 \tag{15}$$

for any $0 < k < p$.

Proof. $E(\delta_{t|t-h}f_{t-h-k|t-h-p}) = E(\theta_h\varepsilon_{t-h}\sum_{i=h+k}^{h+p-1}\theta_i\varepsilon_{t-i}) = 0$, as the ε's all occur before the ε_{t-h} and the ε are uncorrelated. $E(\delta_{t|t-h}f_{t-h+p|t-h+k}) = E(\theta_h\varepsilon_{t-h}\sum_{i=h-p}^{h-k-1}\theta_i\varepsilon_{t-i}) = 0$ since the ε's all occur after the ε_{t-h} and the ε are uncorrelated.

B Proof of $\beta = \frac{\lambda-1}{2}$.

The first step is to derive the unconditional variance and first-order autocovariance of the intuition. The second step is to use those to derive the variance and first-order autocovariance of the intuition update. The final step is to use that to calculate the first-order autocorrelation. This can then be linked to β.

First, the unconditional variance of the intuition is

$$Var(\nu_{t|t-h}) = \lambda^2 Var(\nu_{t|t-h-1}) + Var(\eta_{t,h})$$

and as the first two variances are equal (unconditionally) it follows that

$$Var(\nu_{t|t-h}) = \frac{\sigma_\nu^2}{1-\lambda^2},$$

with σ_ν^2 the variance of the error term.

To derive the first-order autocovariance of the intuition, we will need

$$\begin{aligned} Covar(\nu_{t|t-h}, \nu_{t|t-h-1}) = E(\nu_{t|t-h}\nu_{t|t-h-1}) - E(\nu_{t|t-h})E(\nu_{t|t-h-1}) = \\ \lambda E(\nu_{t|t-h-1}^2) + E(\nu_{t|t-h-1}\eta_{t,h}) = \frac{\lambda\sigma_\nu^2}{1-\lambda^2} \end{aligned}$$

and

$$\begin{aligned} Covar(\nu_{t|t-h}, \nu_{t|t-h-2}) = E(\nu_{t|t-h}\nu_{t|t-h-2}) - E(\nu_{t|t-h})E(\nu_{t|t-h-2}) \\ = \lambda E(\nu_{t|t-h-1}\nu_{t|t-h-2}) + E(\nu_{t|t-h-1}\eta_{t,h}) = \lambda^2 E(\nu_{t|t-h-2}^2) = \frac{\lambda^2\sigma_\nu^2}{1-\lambda^2}. \end{aligned}$$

The variance of the intuition update:

$$\begin{aligned} Var(\nu_{t|t-h} - \nu_{t|t-h-1}) = E[(\nu_{t|t-h} - \nu_{t|t-h-1})(\nu_{t|t-h} - \nu_{t|t-h-1})] = \\ 2Var(\nu_{t|t-h}) - 2Covar(\nu_{t|t-h}, \nu_{t|t-h-1}) = \frac{2\sigma_\nu^2}{1-\lambda^2} - \frac{2\lambda\sigma_\nu^2}{1-\lambda^2} = \\ \frac{2(1-\lambda)\sigma_\nu^2}{(1-\lambda)(1+\lambda)} = \frac{2}{1+\lambda}\sigma_\nu^2. \end{aligned}$$

The first-order autocovariance of the intuition update is

$$\begin{aligned} Covar(\nu_{t|t-h} - \nu_{t|t-h-1}, \nu_{t|t-h-1} - \nu_{t|t-h-2}) = E[(\nu_{t|t-h} - \nu_{t|t-h-1})(\nu_{t|t-h-1} - \nu_{t|t-h-2})] = \\ 2Covar(\nu_{t|t-h}, \nu_{t|t-h-1}) - Var(\nu_{t|t-h-1}) - Covar(\nu_{t|t-h}, \nu_{t|t-h-2}) = \frac{2\lambda - 1 - \lambda^2}{1-\lambda^2}\sigma_\nu^2 = \\ \frac{-(\lambda-1)^2}{(1-\lambda)(1+\lambda)}\sigma_\nu^2 = \frac{\lambda-1}{1-\lambda}\sigma_\nu^2. \end{aligned}$$

Finally, the first-order autocorrelation of the intuition update is thus given by

$$Corr(\nu_{t|t-h} - \nu_{t|t-h-1}, \nu_{t|t-h-1} - \nu_{t|t-h-2}) =$$
$$\frac{Covar(\nu_{t|t-h} - \nu_{t|t-h-1}, \nu_{t|t-h-1} - \nu_{t|t-h-2})}{Var(\nu_{t|t-h} - \nu_{t|t-h-1})} = \frac{\frac{\lambda-1}{1-\lambda}\sigma_\nu^2}{\frac{2}{1+\lambda}\sigma_\nu^2} = \frac{\lambda - 1}{2}.$$

For an α equal to zero, which is the case for (unconditionally) unbiased forecasts (and even more so for rational forecasts), this correlation is equal to β, thereby proving $\beta = \frac{\lambda-1}{2}$.

C Tables and figures

Table 1: In this table the $f_{s|s-k}$ (with event s and forecast horizon k) which are allowed to correlate with $\delta_{t|t-h}$ under the assumption of rationality are marked with O and the rest is marked with X. The correlation does not depend on the forecast horizon k of the forecast error, but only on the moment of the error.

k:	1	2	3	4	5	6	...
...							
s = t-h-1	X	X	X	X	X	X	
s = t-h	O	O	O	O	O	O	
s = t-h+1	X	X	X	X	X	X	
...							

Table 2: The data format for a single variable. Across the columns the timing of the event changes, and across the rows the forecast origin changes. F1 means a 1-step-ahead forecast, F2 a 2-step-ahead forecast, and so on. The bottom row shows the realizations (R).

Realization	4	5	6	7	8	9	10	11	12	13	14
1	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13
2	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
3	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
4		F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
5			F1	F2	F3	F4	F5	F6	F7	F8	F9
6				F1	F2	F3	F4	F5	F6	F7	F8
Forecast origin 7					F1	F2	F3	F4	F5	F6	F7
8						F1	F2	F3	F4	F5	F6
9							F1	F2	F3	F4	F5
10								F1	F2	F3	F4
11									F1	F2	F3
12										F1	F2
13											F1
R	R	R	R	R	R	R	R	R	R	R	R

Table 3: The number of observations, the mean forecast error (with the standard error of the mean in parentheses), the mean squared forecast error and the forecast content for the first 13 horizons for the standardized forecasts.

Horizon	N	Mean Error (SE)	Mean Squared Error	Forecast content
1	3729	-0.720 (0.023)	2.456	0.730
2	3729	-0.720 (0.023)	2.478	0.728
3	3729	-0.757 (0.023)	2.620	0.713
4	3390	-0.736 (0.025)	2.642	0.710
5	3051	-0.775 (0.027)	2.821	0.690
6	2712	-0.891 (0.029)	3.149	0.654
7	2373	-0.949 (0.033)	3.451	0.621
8	2034	-0.986 (0.037)	3.770	0.586
9	1695	-1.041 (0.042)	4.099	0.550
10	1356	-1.147 (0.049)	4.569	0.499
11	1017	-1.482 (0.056)	5.408	0.407
12	678	-1.817 (0.069)	6.522	0.284
13	339	-2.538 (0.089)	9.112	0.000

Table 4: The correlation between forecast errors corresponding to different forecast horizons for the same event.

	H=1	H=2	H=3	H=4	H=5	H=6	H=7	H=8
H=1	1.000							
H=2	0.990	1.000						
H=3	0.875	0.883	1.000					
H=4	0.730	0.738	0.864	1.000				
H=5	0.593	0.602	0.729	0.861	1.000			
H=6	0.528	0.534	0.589	0.724	0.844	1.000		
H=7	0.449	0.455	0.493	0.574	0.702	0.845	1.000	
H=8	0.331	0.336	0.368	0.441	0.527	0.691	0.843	1.000

Table 5: The correlation between forecast errors corresponding to the same horizons for the different events. A row beginning with $\Delta t = k$ indicates that the time span between events, for which the correlation is calculated, is equal to k.

	H=1	H=2	H=3	H=4	H=5	H=6	H=7	H=8
$\Delta t = 1$	0.267	0.262	0.311	0.328	0.337	0.333	0.317	0.278
$\Delta t = 2$	0.259	0.258	0.325	0.318	0.322	0.308	0.275	0.238
$\Delta t = 3$	0.222	0.216	0.236	0.240	0.263	0.248	0.209	0.162
$\Delta t = 4$	0.131	0.130	0.151	0.146	0.173	0.172	0.141	0.106
$\Delta t = 5$	0.181	0.174	0.161	0.125	0.129	0.119	0.090	0.051
$\Delta t = 6$	0.182	0.179	0.191	0.120	0.090	0.073	0.034	
$\Delta t = 7$	0.088	0.084	0.124	0.081	0.031	0.022		
$\Delta t = 8$	0.121	0.121	0.123	0.092	0.034			
$\Delta t = 9$	0.135	0.128	0.150	0.044				
$\Delta t = 10$	0.036	0.041	0.055					

Table 6: The number of observations, the mean forecast revision (with the standard error of the mean in parentheses) and the variance of the revision.

Horizon	N	Mean Revision (SE)	Variance of Revision
1	3729	-0.001 (0.002)	0.017
2	3729	-0.037 (0.008)	0.210
3	3390	-0.051 (0.009)	0.252
4	3051	-0.054 (0.009)	0.274
5	2712	-0.068 (0.010)	0.281
6	2373	-0.065 (0.011)	0.289
7	2034	-0.087 (0.013)	0.322
8	1695	-0.058 (0.012)	0.264
9	1356	-0.066 (0.014)	0.265
10	1017	-0.055 (0.015)	0.235
11	678	-0.027 (0.014)	0.142
12	339	0.000 (0.019)	0.118

Table 7: The number of upward, downward and no-change revisions, along with the percentage of revisions that result in a change (either upward or downward) and the percentage of upward revisions given that there is a change.

Horizon	N	Up	No Change	Down	% Changed	% Up | Change
1	3729	852	1910	967	48.8 %	46.8 %
2	3729	472	2708	549	27.4 %	46.2 %
3	3390	380	2455	555	27.6 %	40.6 %
4	3051	437	2006	608	34.3 %	41.8 %
5	2712	425	1702	585	37.2 %	42.1 %
6	2373	374	1517	482	36.1 %	43.7 %
7	2034	271	1284	479	36.9 %	36.1 %
8	1695	257	1114	324	34.3 %	44.2 %
9	1356	176	891	289	34.3 %	37.8 %
10	1017	157	572	288	43.8 %	35.3 %
11	678	196	367	115	45.9 %	63.0 %
12	339	156	109	74	67.8 %	67.8 %

Table 8: The number of positive, no-change or negative updates after a positive, no-change or negative recent update.

	$Update_{t-2} > 0$	$Update_{t-2} = 0$	$Update_{t-2} < 0$	Sum
$Update_{t-1} > 0$	86	517	249	852
$Update_{t-1} = 0$	147	1558	205	1910
$Update_{t-1} < 0$	239	633	95	967
Sum	472	2708	549	3729

Table 9: The second column concerns the $k-th$ order autocorrelation of the forecast revisions and the third column contains the autocorrelation of the squared revisions. The fourth and fifth column contain the cross-correlations of both, in which the squared revision either leads (+) or lags (-) the revision.

Lag	Revision	Squared Revision	Revision,Sq.Revision(+)	Revision,Sq.Revision(-)
1	-0.111	0.095	-0.222	-0.222
2	0.058	0.049	-0.054	0.080
3	-0.010	0.018	-0.046	-0.042
4	-0.033	-0.006	-0.032	0.027
5	-0.020	0.002	0.012	0.011
6	-0.051	0.023	-0.010	0.015
7	-0.035	-0.009	0.012	0.017
8	0.010	0.005	0.038	0.014
9	0.024	-0.003	0.029	0.018
10	-0.074	0.036	-0.020	0.018

Table 10: The correlation between forecast revisions concerning the same horizons for varying realization dates. A row beginning with $\Delta t = k$ indicates that the time between events, for which the correlation is calculated, is equal to k.

	H=1	H=2	H=3	H=4	H=5	H=6	H=7	H=8
$\Delta t = 1$	0.250	0.010	0.079	0.001	-0.026	-0.009	-0.013	-0.036
$\Delta t = 2$	0.190	0.012	0.063	0.028	0.037	0.052	0.049	0.028
$\Delta t = 3$	0.153	0.028	0.096	0.008	0.017	-0.001	0.001	-0.004
$\Delta t = 4$	0.106	-0.020	0.045	-0.009	-0.010	-0.007	-0.003	-0.004
$\Delta t = 5$	0.145	-0.056	0.017	0.035	0.024	0.003	-0.006	
$\Delta t = 6$	0.201	-0.021	-0.036	0.013	-0.015	-0.007		
$\Delta t = 7$	0.234	-0.031	0.022	0.002	-0.005			
$\Delta t = 8$	0.175	0.029	-0.061	-0.017				
$\Delta t = 9$	0.080	0.085	-0.011					
$\Delta t = 10$	0.167	-0.083						

Table 11: The cross-correlation between forecast revisions and forecast errors for the same realization, in which the forecast error either leads (+) or lags (-) the revision with the horizon difference given in the first column. The final column contains the estimates of θ_h as in (12), except for the first entry, which is 1 by definition.

Horizon	Revision,Error(+)	Revision,Error(-)	θ_h
0	-0.051	-0.051	1
1	0.286	-0.028	0.000
2	0.270	-0.048	0.076
3	0.242	-0.043	0.072
4	0.239	-0.032	0.086
5	0.199	-0.025	0.104
6	0.126	-0.007	0.104
7	0.096	0.011	0.118
8	0.062	0.016	0.085
9	0.057	0.003	0.076
10	0.123	0.030	0.090

Table 12: The λ estimates, along with upperbounds and lowerbounds, for each of the five methods.

Approach	λ	95 % LB	95 % UB
A	0.843	0.839	0.847
B	0.441	0.432	0.458
C	0.288	0.274	0.303
D	0.861	0.848	0.873
E	0.919	0.904	0.932

Table 13: The percentages and parameter estimates on the judgmental adjustments of Franses and Legerstee (2009,2010) compared with the same percentages and parameter estimates for the constructed intuition of method E.

	Franses and Legerstee (2009,2010)	Method E
Percentage adjustments $\neq 0$	89.5 %	89.5 %
Perc. upward adj.	53.5 %	88.8 %
Perc. upward adj. given nonzero	59.8 %	99.2 %
R^2 of adjustments	44.3 %	67.7 %
γ	0.424	1.051
Persistence	0.773	0.919
ρ	0.247	0.898

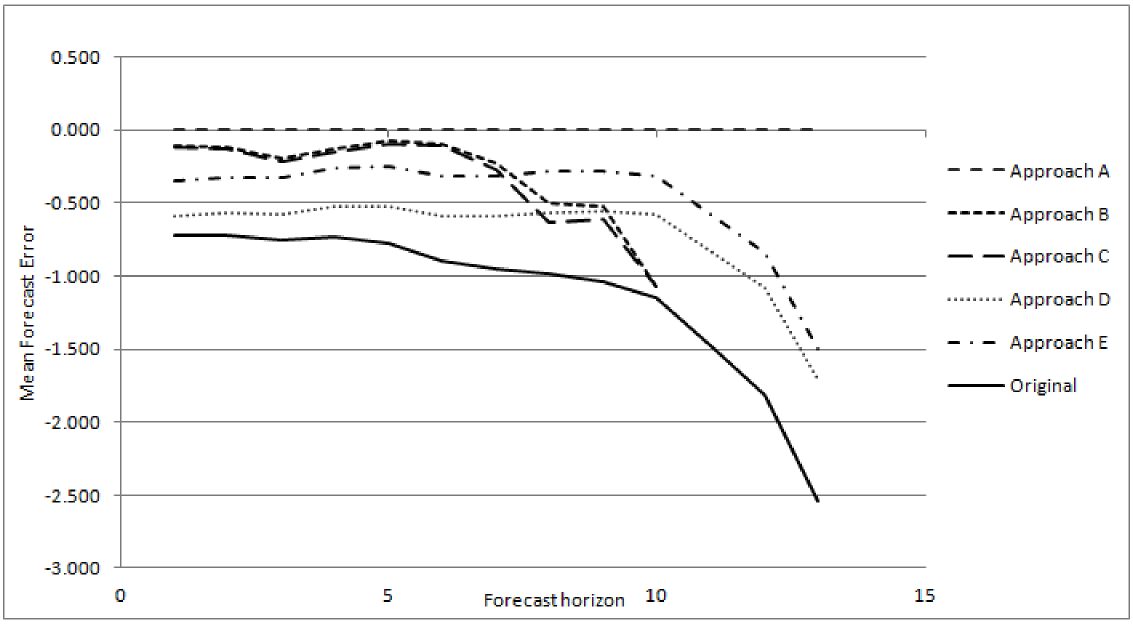

Figure 1: Mean forecast error of the five approaches to construct rational forecasts, together with the mean forecast error of the original forecasts, for all horizons.

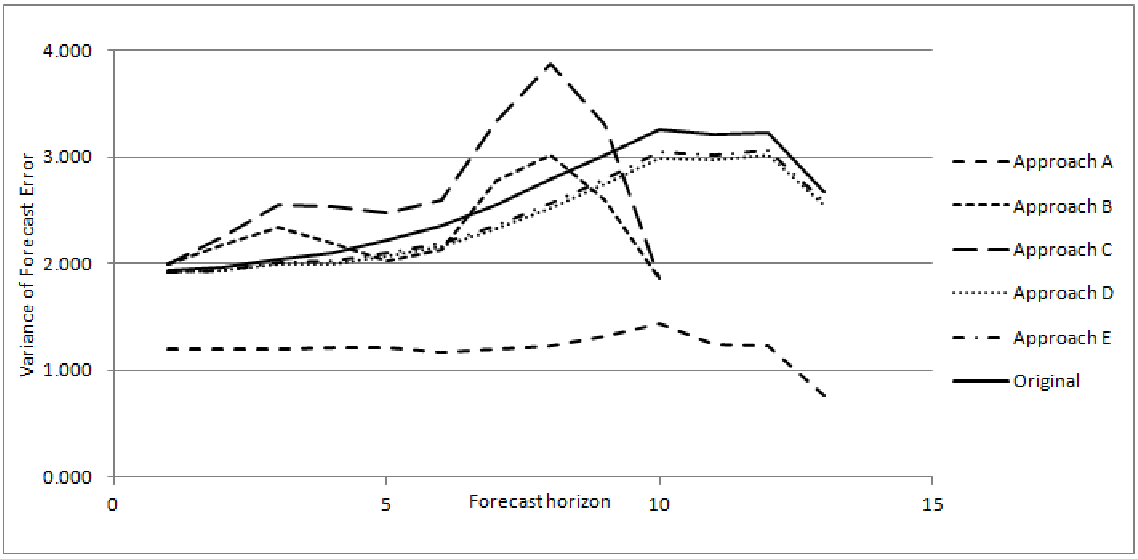

Figure 2: Variance of the forecast error of the five approaches to construct rational forecasts, together with the variance of the forecast error of the original forecasts, for all horizons.