What Determines Forecasters’ Forecasting Errors?

Ingmar Nolte* Sandra Nolte†
Lancaster University Lancaster University
CoFE

Winfried Pohlmeier ‡
University of Konstanz,
CoFE, RCEA, ZEW

This version: May 11, 2015

* Lancaster University Management School Bailrigg, Lancaster LA1 4YX, UK. email: I.Nolte@lancaster.ac.uk.
† Corresponding author: Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK, e-mail: S.Nolte@lancaster.ac.uk. Earlier versions of this paper circulated under the title 'The Good, the Bad and the Ugly: Analyzing Forecasting Behavior within a Quantal Response Framework with Misclassification'. The usual disclaimer applies. For helpful comments we like to thank Karim Abadir, Angelo de Melino, James Mitchell, Christiane Baumeister, Wojciech Charemza, Mark Salmon, Stephen Taylor, Mark Shackleton, Steve Young, as well as the seminar participants at the EEA 2010, CIRET 2010, IAAE 2014 and ESEM 2014.
‡ Department of Economics, Box 124, University of Konstanz, 78457 Konstanz, Germany. email: Winfried.Pohlmeier@uni-konstanz.de.
Abstract

This paper contributes to the growing literature in macroeconomics and finance on expectation formation and information processing by analyzing the relationship between expectation formation at the individual level and the prediction of macroeconomic aggregates. Using information from business tendency surveys we present a new approach of analyzing qualitative forecasting errors made by forecasters. Based on a quantal response approach with misclassification we define qualitative mispredictions of forecasters in terms of deviations from the qualitative rational expectation forecast and relate them to individual and macro factors driving these mispredictions. Our approach permits a detailed analysis of individual forecasting decisions allowing for the introduction of individual and economy wide determinants that affect the individual forecasting error process. Our findings show that professional forecasters in our dataset are able to revise their mispredictions in the long run but not in the short run. We also shed light not only on the heterogeneity of the forecasts among groups of forecasters but also on the heterogeneity of expectations formation for different economic variables. Hence, the paper also contributes to the fairly new literature about models with partial and delayed information processing, as well as to the literature on rational inattention.

JEL classification: C23, C25, D84, E27

Keywords: Expectations, Tendency Survey, Forecasting Errors, Misclassification, GLARMA
1 Introduction

Expectation formation by economic agents is a key element of models in macroeconomics and finance. Although the rational expectations assumption has served as the dominant work horse in most models, the critique on this behavioral assumption is as old as the rational expectations hypothesis itself. The recent literature in macroeconomics and finance emphasizes information rigidities and heterogeneity in information processing by economic agents in order to explain observed deviations from theoretical predictions based on models with homogeneous agents with rational expectations. In particular models with partial (individuals only observe a noisy information signal, Woodford (2002), Sims (2010)) and delayed (only a share of individuals receives up-to-date information, c.f. Calvo (1983), Mankiw & Reis (2002), Reis (2006a, b) and Lorenzoni (2010)) information have been developed which provide more realistic mechanisms for information processing. A detailed survey on these models and their extensions can be found in Mankiw & Reis (2010).

Models of aggregate stock market behavior deviating from the rational expectation behavior include learning models as well as preference and belief based models. In their capital asset pricing model Barberis et al. (2015) assume, besides traditional rational traders, traders who extrapolate their beliefs on the expected returns of the stock market as a weighted average of past price changes. Interestingly the theoretical predictions of models with heterogeneous agents and information rigidities are much more in line with the empirical evidence obtained from expectation surveys (e.g. Coibion & Gorodnichenko (2012) and Greenwood & Shleifer (2015)) than traditional rational expectations based models.

This paper contributes to the literature on expectation formation and information processing by shedding some light on the validity of these approaches. We present a new empirical approach of systematically assessing expectation errors at the individual level.

\[1\] Alternative explanations to why human beliefs deviate from an optimal processing of objective data as assumed by the rational expectations hypothesis have also been investigated in psychological theories. Challenges in this tradition include among many others the works on (cumulative) prospect theory (Kahneman & Tversky (1979), Tversky & Kahneman (1992)), models with distorted beliefs as in the decision model under cognitive dissonance proposed by Akerlof & Dickens (1982) and models taking into account non-cognitive aspects of decision making (Bénabou & Tirole (2000) and Bénabou & Tirole (2004)). Brunnermeier & Parker (2005) propose the concept of optimal expectations, where forward-looking agents take into account expected future utility flows which imply a higher current felicity if agents are optimistic.

\[2\] Barberis, Greenwood, Jin & Shleifer (2015), Table 1 surveys a range of models of aggregate stock market behavior.
level using tendency survey data. Based on a dynamic quantal response approach with misclassification we are able to define qualitative mis-predictions in terms of deviations from the qualitative rational expectations forecast and relate them to individual and macro factors driving these individual mis-predictions.

The use of tendency data to analyze expectation formation has a rather long tradition in economics. Nevertheless, financial economists, like Cochrane (2011) for example, are still rather sceptical about the quality of survey data, arguing that survey responses depend heavily on the way the questions are framed and the employed wording. But then, why are surveys interesting or in other words why are they attractive to researchers? And why are we witnessing an increasing availability of survey data on investors’, consumers’, professionals’ expectations and beliefs about future stock market returns, the economy and macroeconomic variables? for example see Greenwood & Shleifer (2015). First of all, business tendency surveys represent a timely source of information (Lui, Mitchell & Weale (2010)). Second, as pointed out by Mankiw & Reis (2010), these survey data provide a good opportunity for empirical work on imperfect information models. Third, the predictions of different expectations formation models can be gauged by using measures of agents’ forecasts directly available in those survey data (Coibion & Gorodnichenko (2012)). Fourth, while pure time series approaches require some functional form assumptions on the expectation formation process (e.g. linear conditional mean as the best predictor function in the mean squared error sense, definition of the underlying information set, etc.) tests of individual expectation formation on the basis of tendency survey data take the directional forecasts of individuals as the benchmark for the comparison of predictions with realizations. Therefore, no behavioral hypothesis on the expectation formation process has to be imposed ex-ante. Fifth, survey data perform reasonably well in predicting stock returns (Barberis et al. (2015)), or macro variables like inflation (Ang, Bekaert & Wei (2007)). And last but not least, heterogeneity in beliefs and information sets can, in principle, be accounted for due to the richness of the data in terms of their high cross-sectional and time series dimension. Early empirical studies attempting to test the rational expectations hypothesis using business tendency survey data include Nerlove (1983) and König, Nerlove & Oudiz (1981). These studies compare

---

3One needs to be aware that there are some limitations to particular survey data: while professional forecasters’ surveys are the richest in terms of time span, different forecasting horizons, and number of predicted macroeconomic variables, surveys of other agents might be more restricted: consumer forecasts do not ask to predict a particular index, but the whole economy; firm forecasts may only include large institutions and may not be available at the individual level, and the forecasting horizons might be limited as well.
qualitative predictions (e.g. on changes in the levels of inventories, order backlogs or on price changes) with their qualitative assessments of corresponding outcomes. Using log-linear probability models these early studies measure the bivariate association between $j$ different expectations and $j$ different outcome variables at the firm level.

Alternatively, quantification methods like the probability method tracing back to Anderson (1951) and further developed by Carlson & Parkin (1975) or the regression method (e.g. Pesaran 1984, 1985, 1987) were used to measure aggregate expectations on a continuous macro variable at one point in time using $j$ qualitative responses at the individual level. These approaches depend heavily on the aggregation rule (the quantification method), implicit homogeneity assumptions and identifying restrictions (e.g. assumptions on the threshold parameters), so that a rejection of a certain expectation formation hypothesis may always be the result of the underlying identification restrictions and aggregation rules.

More than thirty years after Marc Nerlove’s (1983) attempt to peek into the “black box” of expectation formation at the individual level, we present in this paper an alternative approach of analyzing expectation formation at the individual level using qualitative information from tendency surveys. Our approach differs from previous attempts in various dimensions. Although looking at qualitative forecasts at the individual level we compare in our approach $j$ qualitative predictions with the corresponding qualitative macro outcome at a given time period. Through the introduction of a dynamic Markov type misclassification matrix our approach accounts for individual heterogeneity in forecasting behavior.

Our approach permits a detailed analysis of individual forecasting decisions allowing for the introduction of individual and economy wide determinants affecting the individual forecasting error process. Our findings show that professional forecasters in our dataset are able to correct/revise their mis-predictions in the long run but not in the short run. In this sense, this paper also contributes to the literature on delayed information processing or the theory of “inattentiveness” proposed by Reis (2006a, b), highlighting that individual forecasters need time to process new information. New information is only progressively incorporated in their predictions. Moreover, this

---

4 See for example Mitchell, Smith & Weale (2013) who present a framework for quantifying and aggregating qualitative survey responses of firms. Using an approach closely related to traditional forecast combination techniques, they relate the qualitative answer of each firm to the overall official growth rate based on that company’s financial report, in order to produce an early indicator for official output data.
paper also contributes to the literature on heterogeneity of expectations related to the dispersion of information, leading to the so called “rational inattention” concept proposed by Sims (2003). Indeed, our specifications allow us to distinguish between the predictions of managers, experts from insurance and industry firms from those of bankers. The results shed light on the heterogeneity of the forecasts between these groups, and their “rational inattention”. Bankers are more “attentive” than the other groups and therefore generate less misclassification/mis-predictions then managers and experts from insurance and industry firms. The latter groups may have only partial information (rational inattention), in comparison to bankers, as their day-to-day business is less related to the financial variables of interest in this study.

Like the early studies on expectation formation using tendency data our approach is consistent with the rather general definition of expectations as subjectively held beliefs by individuals. Special behavioral assumptions like rational expectations are reflected in a special form of the misclassification matrix (see Gourieroux & Pradel (1986)). Our model allows for the estimation of individual specific misclassification matrices, but due to the linearity between forecasts at the aggregate level and individual forecasts we can aggregate the misclassification matrix to obtain a measure for aggregate expectation errors in the sense of Pesaran & Weale (2006).

Comparing qualitative individual responses on expectations with qualitative outcomes at the macro level via the misclassification matrix turns out to be an elegant way to solve the interpretational problems related to the no-change interval in qualitative survey data, which has no exact mathematical counterpart in the rational expectations framework. In business survey data the no-change category may simply serve a practical purpose by allowing forecasters a statement about the future state of the world with little effort. The underlying assumption behind the no-change interval is the behavioral hypothesis that there exists a range of imperceptibility, in which individuals do not react to signals. Additional assumptions (e.g. threshold symmetry) to identify the underlying continuous variable to be forecasted, are no longer needed, if the comparisons remain at the (nonparametric) discrete level.

The novel econometric approach developed in this paper depends on a logistic generalized ARMA type structure for the misclassification matrix. It extends previous misclassification approaches for qualitative dependent variables (Hausman, Abreyava & Scott-Morton (1998), Hausman (2001), and Dustmann & van Soest (2004)) to a

---

5Experimental psychologist refer to the thresholds as 'difference limen' or 'differential threshold’. They reflect the just noticeable difference to the smallest change in stimulation a person can detect.
dynamic framework. The model is estimated with data from the Financial Markets Survey of the Centre for European Economic Research (ZEW), a monthly qualitative survey of around 330 financial experts, giving six-month-ahead predictions of major macroeconomic aggregates and financial indicators observed over 16 years.

The outline of the paper is as follows: In Section 2 we develop the general framework. In particular, we define the data generating process at the macro-level and its relationship to the predictions at the individual level. The econometric implementation of the theoretical set-up for quantal response data with a no-change category is presented in Section 3. Here we derive the dynamic misclassification matrix which relates the objective probabilities for the discrete outcome of the macro series with the qualitative forecasts at the micro level. Section 4 discusses the empirical findings. Section 5 presents a number of robustness checks, while Section 6 concludes and gives an outlook on further research.

2 Expectations at the Micro and the Macro Level

The basic idea for our model set-up relies on the definition of forecasts for the discrete counterparts of a continuous macro variable based on different information sets at the macro and at the micro level. We restrict our model set-up to models of expectation formation which have a straightforward statistical formulation. However, the framework can be easily generalized to models based on other behavioral assumptions.

Let \( Y^*_t \) be a continuous time series process at the macro level with \( t = 1, \ldots, T \) and define \( \{S_k, k = 1, \ldots, K\} \) and \( \{G_l, l = 1, \ldots, L\} \) as two given partitions of the outcome space of \( Y^*_t \). Let \( \tilde{Y}_t \) be the discrete counterpart of \( Y^*_t \) within a threshold crossing ordinal response model taking on the form of the \( k^{th} \) unit-vector for \( Y^*_t \in S_k \), i.e., if \( Y^*_t \in S_k \) the \( k^{th} \) component of \( Y_t, Y_{k,t} \), is equal to one, while all other components are equal to zero. Correspondingly define for the second partition, \( \tilde{Y}_t \) as the \( l^{th} \) unit-vector of dimension \( L \) indicating that \( Y^*_t \in G_l \). For example, if \( Y^*_t \) denotes a first difference or a return rate series possible partitions could be a two states partition (growth vs. no growth) or a three states world (large growth rate above some positive threshold, weak growth around zero called the “no-change” state, large negative growth rate below some negative threshold). The distinction between two different partitions is reasonable for many tendency data if the survey contains a no-change interval, while at the macro level a reasonable partition could also be a binary one. Assume a forecasting horizon of length \( h \) and a partition \( \{S_k\} \) and let \( \mathcal{F}_t \) be the information set given at time \( t \),
then for $Y_{t+h}$ there exists a $K$ dimensional vector $P_{t+h}$ with the typical element

$$\Pr[Y_{t+h}^* \in S_k|\mathcal{F}_t] = \Pr[Y_{k,t+h} = 1|\mathcal{F}_t],$$

that given the information set $\mathcal{F}_t$ the outcome $Y_{t+h}^*$ at time $t+h$ lies in $S_k$. Since $E[Y_{k,t+h}|\mathcal{F}_t] = \Pr[Y_{k,t+h} = 1|\mathcal{F}_t]$ the $k$-th element of $P_{t+h}$ is the continuous rational expectations forecast of state $k$ in the sense of Gourieroux & Pradel (1986).

Consider now the same process forecasted by an individual forecaster $i$, $i = 1, \ldots, N_t$, where $N_t$ denotes the number of participants in the survey at time $t$. Each forecaster is assumed to make a qualitative assessment at time $t$ about the development of the macro variable $Y_{t+h}^*$. $P_{i,t+h}$ denotes the $L$ dimensional vector of probabilities

$$\Pr[Y_{t+h}^* \in G_l|\mathcal{F}_{it}] = \Pr[\tilde{Y}_{l,t+h} = 1|\mathcal{F}_{it}],$$

that state $l$ occurs in $t+h$ given individual $i$-th information set $\mathcal{F}_{it}$.

The forecasts at the micro level may differ because of different information sets $\mathcal{F}_{it}$ (information disparity) but also because of different individual probability measures (belief disparity). In principle, our setup allows for both kinds of disparities, but for the ease of notation to cover the standard case of information disparity at the micro level, we stick to the assumption of belief homogeneity.

Expectations at the macro level, $P_{t+h}$, and expectations for forecaster $i$, $P_{i,t+h}$, are related through the individual misclassification matrix $\Pi_{i,t+h}$ by:

$$P_{i,t+h} = \Pi_{i,t+h}^t P_{t+h}, \quad (1)$$

where the components of $\Pi_{i,t+h}$ are given by $\pi_{kl}^{i,t+h} = \Pr[\tilde{Y}_{l,t+h}|\mathcal{F}_{it} = 1|Y_{k,t+h}|\mathcal{F}_t = 1]$.

The individual misclassification matrix, $\Pi_{i,t+h}$, defined in (1) relates the objective expectation formation process at the macro level to the individual (possibly subjective) expectation formation process at the micro level. If $P_{t+h}$ denotes the predictions of the state indicator vector under rational expectations, the misclassification matrix measures the deviation of a forecaster’s beliefs from the true data generating process. For identical partitions $S_k = G_k, K = L$, the forecaster’s expectations are rational

6For both, informational disparity and disparity in beliefs $\Pr[\tilde{Y}_{l,t+h} = 1|\mathcal{F}_{it}] = E_{it}[\tilde{Y}_{l,t+h}|\mathcal{F}_{it}]$, where $E_{it}[\cdot]$ indicates that the expectation is taken under the $i$-th forecasters (subjective) probability measure at $t$.

7We refer to objective expectation and objective probabilities as those implied by the (theoretical) model on the macro level, see Muth (1961).
in the sense of Muth (1961), if the misclassification matrix is the identity matrix, \( \Pi_{i,t+h} = I_K \). Moreover, at the macro level specific expectation formation schemes such as static expectations, adaptive or error learning expectations can be imposed. In this case the misclassification matrix can be used to measure deviations of the individual forecaster’s beliefs from a given macroeconomic model world. For instance let \( P_{i,t+h} \) be the probability forecasts for the exchange rate changes based on a purchasing power parity model. In this case \( \Pi_{i,t+h} \) measures the extent to which the beliefs of forecaster \( i \) differ from the purchasing power parity hypothesis.

**Average Expectations**

In the general set-up developed above \( \Pi_{i,t+h} \) is individual specific (e.g. reflecting ability or experience of forecaster \( i \)) as well as time specific (e.g. evolution of macroeconomic uncertainty over time). In a world of heterogeneous rational expectations, however, agents build their expectations on the expectations of other agents. Binder & Pesaran (1998) show, that a unique solution exists, if the forecasts of the agents are only based on common knowledge. Such a solution does not appear to be attractive in the context of survey expectations. Their merits particularly lie in the heterogeneity of the information sets of the forecasters. Pesaran & Weale (2006) propose a weaker concept of rational expectations in the presence of heterogeneity, called average rational expectations. This concept relies on the weighted average of the individual conditional densities of the continuous macro variable to be forecasted. Average expectations are consistent with the existence of heterogeneity in beliefs and allow for systematic deviations from rational expectations at the individual level. Let \( \bar{P}_{t+h} \) be the average expectation over the state vector \( Y_{t+h} \) defined as

\[
\bar{P}_{t+h} = \sum_{i=1}^{N_t} w_{it} P_{i,t+h} = \sum_{i=1}^{N_t} w_{it} E[Y_{t+h}|F_{it}] \tag{2}
\]

where the non-negative weights satisfy the conditions \( \sum_{i=1}^{N_t} w_{it} = 1 \) and \( \sum_{i=1}^{N_t} w_{it}^2 = O(\frac{1}{N_t}), \forall t \). Inserting Equation (1) gives

\[
\bar{P}_{t+h} = \sum_{i=1}^{N_t} w_{it} \Pi'_{i,t+h} P_{t+h} = \bar{\Pi}'_{t+h} P_{t+h}, \tag{3}
\]

where \( \bar{\Pi}'_{t+h} = \sum_{i=1}^{N_t} w_{it} \Pi'_{i,t+h} \). Let \( u_{t+h} = Y_{t+h} - P_{t+h} \) and \( u_{i,t+h} = Y_{t+h} - P_{i,t+h} \) be the expectation error vector for the macro and the individual level respectively. The two expectation errors are related by

\[
u_{i,t+h} = (I_K - \Pi'_{i,t+h}) P_{t+h} + u_{t+h}.
\]
Aggregating over the number of forecasters in survey wave $t$ gives the average expectation error $\bar{u}_{t+h} = \sum_{i=1}^{N_t} w_i u_{i,t+h}$

$$\bar{u}_{t+h} = (I_K - \bar{\Pi}'_{t+h}) P_{t+h} + u_{t+h}.$$ 

Under a set of sufficient conditions stated in Pesaran & Weale (2006), $\bar{u}_{t+h} \xrightarrow{\text{q.m.}} u_{t+h}$ for $N_t \to \infty$, this implies $I_K = \bar{\Pi}'_{t+h}$, so that average expectations equal rational expectations. Using the average misclassification matrix $\bar{\Pi}'_{t+h}$ we are able to test the expectation formation in terms of a 'consensus' or market concept. Note that $\Pi'_{t,t+h}$ can be used to test for a specific expectation hypothesis at the individual level for each forecaster separately in the sample. In this case $\bar{\Pi}'_{t+h}$ can simply be regarded as a summary statistic.

3 Econometrics and Empirical Set-up

3.1 Data

Our empirical analysis is based on the ZEW Financial Market Survey which has been conducted since December 1991 on a monthly basis and focuses on international financial market series. Those include the financial market in Germany, USA, Japan, Great Britain, France, Italy, and since January 1999, the Euro-Area. Each month representatives of the German Financial sector employed in banks, insurance companies or at finance departments or economic research departments in large industrial corporations - therefore called experts - are polled on their expectations regarding the developments in important international financial markets. From December 1991 to January 2012, e.g. 242 months, 1086 experts responded at least once on the survey. Since 1993 the number of participants are relatively stable, with around 300 experts responding to the questionnaire each month.

Participants are asked to give their six-months-ahead predictions for the economic activity, the inflation rate, the short and the long term interest rates, the exchange rates, and the profits of 13 German industries (banking, insurance, vehicles construction, chemicals and pharmaceutical, steel/non-ferrous metals, electrical engineering, mechanical engineering, consumer goods and retailing, construction, utilities, service providers, telecommunications, information technology), as well as the oil price. Up to November 1998, the exchange rate question relates to US-Dollar, Yen, UK-Pound and Swiss Franc per Deutsche Mark (DM) and since December 1998 per Euro. All

$q.m.$ stands for convergence in quadratic mean.
questions are asked with respect to the countries listed above.

The ZEW Financial Markets Survey is a purely qualitative survey, meaning that the respondents are asked to predict, whether in the next six months the price of the corresponding “financial market series” will go up, stay the same, or go down. A fourth possibility is to choose “no assessment”, if forecasters do not want or are unable to make a prediction. Responses’ probabilities for this category are rather small (on average less than 3 percent) and show no systematic correlations with the state of the macro economy. This category will therefore be ignored. At the beginning each questionnaire had to be returned on the third Friday of a month, but until October 2001 it changed to the second Friday of a month. The result of the questionnaire are published each month in the ZEW Financial Market Report. It includes a detailed listing of the changes in the percentages on the different response categories, as well as its standard deviation, for the inflation rate, the short and long term interest rates, the stock indices, the exchange rates, and the oil price. The ZEW Financial market report receives a lot of attention in German and European media and is closely observed by stock market investors.

Here we concentrate on 5 series asked with respect to Germany namely the inflation rate, the short and long term interest rates, DAX30, and the USD/EUR exchange rate. We decide to focus on 3 macro variables and 2 financial series in order to highlight whether forecasters form their expectations differently for different economic variables.

3.2 Estimation and Specification

Estimation

Our model is estimated within a standard panel framework based on the individual discrete forecasts for the change (return) of a given macro or financial series. We treat the discrete forecasts at the individual level as misclassified values (mispredictions) of the macro forecasts. The outcome probabilities at the individual level can be expressed as the sum of the misclassification matrix and the macro outcome probabilities. This leads to a log likelihood function, $\ln \mathcal{L}$, of the form

$$
\ln \mathcal{L} = \sum_{t=1}^{T} \sum_{i=1}^{N_t} \sum_{l=1}^{L} 1_{\{\tilde{Y}_{i,t+l} \mid \mathcal{F}_t = 1\}} \ln P_{i,t+l,l},
$$

(4)

where $P_{i,t+h,l} = \sum_{k=1}^{K} n_{i,t+h}^{kl} \cdot P_{t+h,k}$. In order to facilitate the computation of the likelihood $P_{t+h,k}$ is estimated in a first step. For the elements of the misclassification matrix
we choose a multinomial logistic form, where the log-odds ratios of the misclassification probabilities follow a generalized Autoregressive Conditional Moving Average (ARMA) type process.

For the Maximum Likelihood (ML) estimation of our model as outlined in equation (4), we need to specify a model for the macro probability vector \( \mathbf{P}_{t+h} \) for the discrete outcomes of the macro series and a specification for the misclassification matrix \( \Pi_{t+h} \). In the following we will concentrate on the case \( L = K = 3 \), i.e. a macro world with \( K = 3 \) states (up/same/down) that matches the \( L = 3 \) answer categories asked for in the ZEW survey. For each series analyzed, we use up (\( \tau_{\text{up}} \)) and down (\( \tau_{\text{down}} \)) threshold series asked for as a separate question by the ZEW, to determine the (up/same/down) partition of the respective macro series. We set \( h = 6 \) which corresponds to the 6 months forecasting horizon asked for in the ZEW survey.

### Macro World Probabilities

We rely on a dynamic quantal response strategy to derive \( \mathbf{P}_{t+h} \). We assume the following time series specification for \( Y_{t+h}^* \):

\[
Y_{t+h}^* = \mu_{t+h} + \varepsilon_{t+h},
\]

where \( \mu_{t+h} = \mathbb{E}[Y_{t+h}^*|\mathcal{F}_t] \) denotes the conditional expectation of \( Y_{t+h}^* \) given \( \mathcal{F}_t \). This model is sufficiently general to incorporate standard expectation formation models, such as static, adaptive or error learning specifications and also standard time series specifications such as ARMA models by choosing \( \mu_{t+h} \) appropriately. The \( k \)th component of \( \mathbf{P}_{t+h} \) can then be obtained by

\[
\mathbf{P}_{t+h,k} = F_{t+h}(\tau_{t+h}^k - \mu_{t+h}) - F_{t+h}(\tau_{t+h}^{k-1} - \mu_{t+h}), \quad \text{for} \quad k = 1, \ldots, K,
\]

where \( F_{t+h} \) denotes the conditional cumulative density function (c.d.f.) of \( \varepsilon_{t+h} \) given \( \mathcal{F}_t \) and \( \tau_{t+h}^k \) the \( k \)th threshold, where we set for convenience \( \tau_{t+h}^K = \infty \) and \( \tau_{t+h}^0 = -\infty \) for all \( t \). Assuming an ordered probit model we obtain:

\[
\mathbf{P}_{t+h,k} = \Phi \left( \frac{\tau_{t+h}^k - \mu_{t+h}}{\sigma_{t+h}} \right) - \Phi \left( \frac{\tau_{t+h}^{k-1} - \mu_{t+h}}{\sigma_{t+h}} \right), \quad \text{for} \quad k = 1, \ldots, K,
\]

where \( \Phi \) denotes the standard normal c.d.f. and \( \sigma_{t+h} = \mathbb{E}[\sigma_{t+h}|\mathcal{F}_t] \), with \( \sigma_t \) the standard deviation of \( \varepsilon_t \).

#### h-Step Ahead Forecast

In our empirical study we obtain the forecasts \( \hat{\mathbf{P}}_{t+h} \) by estimating \( \sigma_t \) on the grounds of a six month realized volatility estimator using weekly
observations and we assume AR(10) forecasting models with zero restrictions imposed upon the first 6 autoregressive coefficients for $\sigma_{t+h}$ and $Y^*_{t+h}$ to obtain the forecasted variables $\hat{\sigma}_{t+h}$ and $\hat{Y}^*_{t+h}$ at time $t$.

The Perfect Forecast Case For comparison reasons and to rule out forecasting model uncertainty caused by the use of a potentially misspecified forecasting model we also consider the following perfect forecast case under which we assume that we know the realizations of $\sigma_{t+h}$ and $Y^*_{t+h}$ already at time $t$. Hence, we obtain $\hat{P}_{t+h}$ by setting $\hat{\sigma}_{t+h} = \sigma_{t+h}$ and $\hat{Y}^*_{t+h} = Y^*_{t+h}$.

The above dynamic quantal response forecasting strategy should be understood as a starting point for the prediction of $\hat{P}_{t+h}$. An alternative strategy consists of deriving the probability vector $\hat{P}_{t+h}$ by dynamic quantile regression based forecasts of $Y^*_{t+h}$, in which the respective quantiles, which then yield $\hat{P}_{t+h}$, are chosen to match the up and down thresholds. A further possibility consists of transforming $Y^*_{t}$ with the help of the threshold series into their discrete counterparts $Y_{t}$ and to forecast $\hat{P}_{t+h}$ directly with the discrete information using for example an Autoregressive Conditional Multinomial (ACM) model of Russell & Engle (2002).

Misclassification Matrix

Recall the misclassification matrix $\Pi_{it+h}$, whose components are given by

$$\pi_{kl,i,t+h} = \Pr \left[ \hat{Y}_{l,t+h} | F_{i,t} = 1 \bigg| Y_{k,t+h} | F_{i} = 1 \right].$$  \hspace{1cm} (5)

Thus $\pi_{kl,i,t+h}$ is the probability that a participant $i$ gives the assessment $l$ although the true assessment should have been $k$, for $k, l = 1, 2, 3, l \neq k$. The probability that the assessment of the participant is correct, i.e., that no misclassification occurs, is then given by $\pi_{kk,i,t+h}$ and the sum over $l$ of $\sum_{l} \pi_{kl,i,t+h}$ is equal to one by definition.

In order to model these misclassification probabilities, we use the ACM model suggested by Russell & Engle (2002), with a logistic link function. This leads to a multinomial logit model of the form

$$\pi_{kl,i,t+h} = \frac{\exp \{ \Lambda_{kl,i,t+h} \}}{\sum_{l=1}^{3} \exp \{ \Lambda_{kl,i,t+h} \}}, \quad k, l = 1, 2, 3,$$

where the log-odds ratio $\Lambda_{kl,i,t+h}$ will be specified below. As normalization constraint, we use as the reference category the corresponding correct macro outcome category, i.e. $\Lambda_{kk,i,t+h} = 0$ for $k = l$, such that the odds ratios are defined as the quotient of a
given misclassification probability to the probability of a correct classification. The resulting vector of log-odds ratios given by

\[ \Lambda_{i,t+h}^k = \left( \Lambda_{i,t+h}^{kl}, \Lambda_{i,t+h}^{kl'} \right)' \]

\[ = \left( \ln \left[ \frac{a_{i,t+h}^{kl}}{a_{i,t+h}^{kk}} \right], \ln \left[ \frac{a_{i,t+h}^{kl'}}{a_{i,t+h}^{kk'}} \right] \right)' \]

for \( k, l, l' = 1, 2, 3 \), and \( k \neq l \neq l' \),
is specified as a multivariate ARMA type process, including explanatory variables statically

\[ \Lambda_{i,t+h}^k = G^k Z_{i,t+h} + \alpha_{i,t+h}^k \]

\[ \alpha_{i,t+h}^k = c^k + \sum_{j=1}^{q} A_j^k \xi_{i,t+h-j} \quad \forall k = 1, 2, 3 \]

with \( \{ A_j^k, j : 1 \to q \} \) being a matrix of dimension \( (2 \times 3) \) with the elements \( \{ a_{ij}^k \} \),
for \( l, l' = 1, 2, 3 \) with \( l \neq k \). \( c^k \) denotes the \((2 \times 1)\) vector of constants \( c^k = (c_1^k, c_2^k)' \),
for \( l, l' \neq k \). \( Z_{i,t+h} \) denotes the \( g \)-dimensional vector of explanatory variables which
are time and/or individual specific and are included statically with \( G^k \) as the corresponding coefficient matrix of dimension \((2 \times g)\). The misclassification indicator vector driving the ARMA type model is specified as

\[ \xi_{i,t+h} = (\xi_{i,t+h}^1, \xi_{i,t+h}^2, \xi_{i,t+h}^3)' \]

\[ = \left( \mathbb{1} \{ Y_{2,t+h|x_{i,t+h}} = 1 \} + \mathbb{1} \{ Y_{3,t+h|x_{i,t+h}} = 1 \} \right) \cdot \mathbb{1} \{ Y_{i,t+h} = 1 \} \]

which is the three-dimensional state vector of individual \( i \) at time \( t + h \) representing
whether that individual misclassified at time \( t \) the true change of the underlying variable. For example, if \( \xi_{i,t+h} = (1, 0, 0)' \) then individual \( i \) predicted either a no change or a negative change of the variable at time \( t \), although in reality the variable went up in the period from \( t \) to \( t + h \). Given that definition of the misclassification indicator vector a negative coefficient in \( A_j^k \) implies lower values in the corresponding log-odds ratio vector \( \Lambda_{i,t+h}^k \) and thus a specific learning effect. Consider for example
the case for \( A_j^1 \) which is given by

\[ A_j^1 = \begin{pmatrix} a_{41}^1 & a_{22}^1 & a_{31}^1 \\ a_{41}^1 & a_{22}^1 & a_{31}^1 \\ a_{41}^1 & a_{22}^1 & a_{31}^1 \end{pmatrix} \]

\(^{9}\text{We suppress } j.\)
and measures the influence of $\xi_{i,t+h}$ on $\Lambda_{i,t+h}^1$, which is given by

$$
\Lambda_{i,t+h}^1 = \left( \ln \left[ \frac{\pi_{12,i,t+h}}{\pi_{11,i,t+h}} \right], \ln \left[ \frac{\pi_{13,i,t+h}}{\pi_{11,i,t+h}} \right] \right),
$$

where $\pi_{11,i,t+h}, \pi_{12,i,t+h}, \pi_{13,i,t+h}$ are the misclassification probabilities that individual $i$ makes the assessment up/same/down given that the financial variable goes up. Assume that the first column of $A_j^1$ has negative coefficients then the impact of a previous misclassification (may it be same or down), in case that the variable went up before, implies a smaller log-odds ratio vector $\Lambda_{i,t+h}^1$ (in both components) and thus lower misprediction since the “correct” $\pi_{11,i,t+h}$ probability goes up while the other two probabilities go down and hence a clear learning effect from a misprediction (in the case of an upwards change) before. Similarly a negative second (third) column of $A_j^1$ represents a learning effect for an up assessment with regard to a misprediction for a previous no change (down) movement. For $A_j^2$ and $A_j^3$ we obtain the corresponding interpretations. In the case that a column in $A_j^k$ is positive we observe no learning but a more pronounced and accelerated degree of misprediction. In the case that we observe alternating signs for the coefficients in the columns of $A_j^k$ the interpretation regarding learning or misprediction is more complex, since at the same time both the basis misclassification category’s probability as well as that component’s misclassification probability might have gone down, for the component associated with the negative coefficient, while at the same time the other misclassification probability component’s associated with the positive coefficient will go up or stay at least the same. The only conclusion that can be drawn in such a case is that the misclassification probability corresponding to the negative coefficient goes down (or stays the same) while the misclassification probability corresponding to the positive coefficient goes up (or stays the same). No interpretation can be made with respect to the change of the basis category’s misclassification probability and thus we can make no statement about the learning effect just by examining the signs of the coefficients in the columns of $A_j^k$.

In our empirical study we employ ACM specifications with $q = 1$, and use the same set of explanatory variables for each state $k$ and assume that the coefficient matrices are equal across states, i.e. $G = G^1 = G^2 = G^3$ so that the corresponding coefficients reflect the general impact of the explanatory variables on misclassification.

Our set of explanatory variables can be divided in two classes: purely individual specific and both, time and individual specific variables.

We specify four individual specific explanatory variables, namely $\text{Insurance}_i$, $\text{Industry}_i$, 
Manager, and Reliability. Since our group of experts consists of experts from banks, insurance and industry companies we include dummy variables to figure out whether experts from particular groups have different abilities in forecasting the underlying series. Thus, we generate the dummy variables Insurance, and Industry to be equal to 1 if individual is a representative of an insurance or industry firm, respectively, and 0 otherwise. Hence, the group of bankers forms the basis category. The dummy variable Manager is equal to 1 if individual belongs to the management board and 0 otherwise and thus measures the forecasting ability of managers over the other groups of experts working in economic, security analysts, asset management or financial accounting divisions. Reliability is defined as the share of the number of questionnaires returned in time over the overall returned number of questionnaires for each individual. This explanatory variable allows us to characterize the general degree of punctuality or reliability of each participant, which we assume to be an overall characteristic of a forecaster and time independent.

In contrast Performance is a time and individual specific variable defined as the share of correct predictions over the last 12 months of individual and captures the effect of the historical forecasting performance of each participant on their future assessments. This variable allows us to examine a general individual specific long term learning process, while the misclassification indicator parameters reflect a short term error learning process with respect to specific states of the world, here \( k = 1, 2, 3 \) (up/same/down).

Note, that all explanatory variables have positive domains so that their influence on the vector of log-odds ratios and thus on the degree of misclassification can be interpreted in a straightforward way. A positive coefficient implies higher values for the components in the vector of log-odds ratios and hence reflects a higher degree of misclassification, since per construction for every state \( k \) the basis category corresponds to the “correct prediction” category.

4 Empirical Findings

4.1 Descriptive Figures

From the set of 1086 forecasters having responded during the life-time of the survey (242 months) at least once, we have selected those who have answered the questionnaire at least 12 times and thus shown a minimum regular interest in the survey.
Altogether we end up with roughly 300 participants per month, from which about 10% each are experts from industry firms and insurance companies. The share of managers returning the survey amounts to 15%. All of these figures remain relatively stable over the lifetime of the survey. The histogram in Figure 1 gives an idea about the distribution of the response rate across experts, which is measured as the number of questionnaire returns in time over the number of overall returns of the questionnaire for each individual.

**Figure 1:** Distribution of the response rate across experts. The response rate for each expert is computed as the number of questionnaire returns in time over the number of overall returns of the questionnaire.

In Figure 2 below we present for the 5 variables under consideration (German inflation rate, German short term interest rates, German long term interest rates, DAX 30, USD/EUR exchange rate) the graphs of the 6 months changes (inflation and interest rates) or the 6 months returns (DAX 30 and USD/EUR), the 6 months standard deviation estimated with a realized volatility estimator, the shares for up/same/down answered by the survey participants (sample probabilities), as well as the objective probabilities of up/same/down for both the perfect forecast and the 6 months ahead forecasting scenarios. In these graphs the up probability share is always drawn in blue, the same probability share is always drawn in grey and the down probability share is always drawn in yellow.

Whereas the perfect forecast case objective probabilities mimic the behavior of the series very closely (row 4), the objective probabilities from the 6 months ahead fore-
casting scenario (row 5) do anticipate their general behavior also quite well but to a less amplified extent and with variations only in a narrow band above stable baseline shares. This is of course due to the specific forecasting setup for probabilities that adds additional forecasting uncertainty and thus cloaked the true changes of the underlying series. The perfect forecast probabilities, in contrast, rule out any forecasting model uncertainty and constitute the limiting case in that respect that we assume that at time $t$ we already know the return (and its standard deviation) of the particular series from $t$ to $t + 6$.

From a pure inspection of the (up/same/down) sample probabilities of the survey participants (row 3), we also see that for the inflation and the interest rate series the group of experts seems to anticipate the changes of the underlying series very well. For the DAX 30 and the USD/EUR series, however, this intuition is much less clear. Under the assumption of (average) rational expectations these sample probabilities should correspond on average to the above discussed forecasted objective probabilities. However, we observe a certain amount of discrepancies between the sample and the objective probabilities yielding a specific amount of misclassification that we analyze within the proposed modeling framework.
Figure 2: The first row displays the graphs of the 6 months changes (inflation and interest rates) or the 6 months returns (DAX 30 and USD/EUR) of the actual series, second row: 6 months standard deviation estimated with a realized volatility estimator, third row: shares for up/same/down answered by the survey participants (sample probabilities), as well as the objective probabilities of up/same/down for both the perfect forecast and the 6 months ahead forecasting scenarios in the fourth and the fifth rows.
4.2 Estimation Results

In Table 1 we present for both macro probability scenarios the estimation results for the explanatory variables included in the ACM misclassification regressions for the 5 variables under consideration. The detailed regression outputs can be found in Appendix A.

A first important finding is that for the perfect forecast and the 6 months ahead forecasting scenarios we find no general qualitative differences in the effects of the explanatory variables on the degree of misclassification which implies a certain robustness of the results with regard to the choice of the specific forecasting model for the 6 months ahead prediction of the macro up/same/down probabilities.

Except for the long term interest rates in the perfect forecast scenario, where the coefficient is positively significant, we find that for all 5 series the effect of past prediction performance, which measures the degree of correct forecasts over the last 12 months, has a significant negative impact in both scenarios on the degree of misclassification implying a significant link between better historical prediction performance and more accurate future forecasting ability of individual \( i \). Hence, we observe that those forecasters who have given a better assessment in the past do continue to do so also in the future. They reveal a better understanding of the data generating process, may be better informed or may have access to better resources in making their predictions especially in the long run. This finding also allows an interpretation as a general long term learning of individual \( i \), in that sense that the individual forecaster learns from his past long run mistakes how to make more accurate forecasts which then imply increasing shares of correct predictions and thus less misclassification of future forecasts.

The effect of only the last mis-prediction and short term learning can be examined by looking at the coefficients in the \( A_1^k \) innovation term matrices of the ACM model. It is, however, less clear cut and we generally do not observe a uniform sign across coefficients. An interesting finding is that for both scenarios we observe for the DAX30 that the \( A_1^3 \) matrices have significant positive coefficients, implying that forecasters do not learn from short term past mis-prediction when these financial series are in a downward move. Moreover, they neither seem to learn when this series is in an upward move in a perfect forecast scenario. The results show on the contrary that their degree of mis-prediction is accelerated. To some extent this can be related to the concept of “trend following” proposed by Barberis et al. (2015). They provide evidence that many stock market investors exhibit a kind of “trend following behaviour” in the sense
that they form their predictions of future stock market returns based on past returns, meaning that they expect the stock market to perform well in the near future if it has recently performed well and vice versa. The professional forecasters we consider in our dataset may also follow past trends to form their beliefs/expectations about the future, and therefore they are unable to revise their predictions so rapidly to follow up and down movements. Additionally this results can be either related to bubbles observed for these series or be attributed to an optimistic bias. If we take a closer look at the graphics in the third row in Figure 2 for the DAX30, which represents the shares of up/same/down answered by the survey participants (sample probabilities), we see that approximately 60% of the participants always predict that the price of the DAX30 in the next 6 months will go up. This suggests that the participants are most of the time optimistic concerning the assessment of the 6 month ahead forecast of the DAX30, and might confirm the exhibition of an optimistic bias.

For the USD/EUR series, the short term interest rate respectively, we observe a similar but less pronounced effect for the upward state matrices $A_1^1$, for the downward state matrices $A_3^1$ respectively. On the contrary, we see that for the inflation series the $A_1^1$ matrix has significant negative coefficients, implying that forecasters do learn from short term past mis-prediction if this macro series is in an upward move. All results shed more light on the time horizon in learning and contradict to a certain extent previous findings in the literature. Many observations from psychology, political science, and organizational behavior indicate that people exhibit a taste for consistency. Meyvis, Ratner & Levav (2010) show that people are unable to recognize their forecasting error, due to the fact that they exhibit a tendency to recall their affective forecast to be in conformity with their actual experience. In this respect, they do not revise their beliefs and continue to rely on the same incorrect beliefs for their subsequent forecasts, so that they are unable to learn from past mis-predictions. This is in line with the findings of Wilson, Meyers & Gilbert (2001) and Fischhoff (1975) that people erroneously remember their past predictions.

Our observation that individuals tend to learn from their past long term mistakes but not from the short ones is in contrast more consistent with the information processing mechanism put forward in relatively new macroeconomic models. Calvo (1983), Mankiw & Reis (2002), Lorenzoni (2010) and Mankiw & Reis (2010) among others.

---

10 Optimism bias is the tendency for people to over-estimate the likelihood of positive events and under-estimate the likelihood of negative events. See Weinstein (1980), among others.

11 This is in line with Ang et al. (2007), who provide evidence that survey data can predict future inflation.
developed so called delayed information processing models in which only a share of individuals receives up-to-date information to provide more realistic mechanisms for information processing, whereas Woodford (2002) and Sims (2010) introduced models with partial information processing in which individuals only observe a noisy information signal. Our result, that professional forecasters are able to correct/revise their mis-predictions in the long run but not in the short run is also related to the theory of “"inattentiveness"" proposed by Reis (2006a,b), highlighting that individual forecasters need time to process new information, and that new information is only progressively incorporated in their predictions, due to their limitations in acquiring and processing information (Coibion & Gorodnichenko (2012)). This result is also supported by Andrade & LeBihan (2013) who show that participants of the ECB Survey of Professional Forecasters fail to update systematically their forecasts as a result of new information disclosure, and that they differ in their frequency of updating.

Interestingly, we find that in cases where the reliability variable is significantly different from zero, it is always positive implying that those forecasters who are more painstaking in returning the questionnaire in time do not necessarily provide better forecasts. For the DAX30 in both scenarios this result can be again related to the fact that 60% of the participants always predict that the price will go up. This result can in addition also find support in the delayed information processing models literature (Calvo (1983), Mankiw & Reis (2002), Lorenzoni (2010) and Mankiw & Reis (2010)). Those professional forecasters who delay the returning of the questionnaire have practically more time to process the available information, such that delaying the return of the questionnaire then plays in their favor. They take more time, and consequently make less mistakes and more accurate predictions.12

We find that experts from insurance and industry firms seem to generate a higher degree of misclassification (if significant the respective dummy variable coefficients are positive except for one case) which might be explained by less resources and information and by the fact that they might face a further distance in their day-to-day business to the variables of interest and thus have less experience with these variables than their colleagues from banks. A similar finding can be made with respect to the coefficient of the manager dummy variable which is if significant always positive (ex-

12Nevertheless, this result contradicts to a certain extent the literature investigating the relationship between conscientiousness and job success. Several psychological studies, indeed, have shown that conscientiousness is highly positively correlated with job success (e.g. Barrick & Mount (1991) and Robertson & Kinder (1993)), in the sense that employees having a higher interest in their job, and therefore have a higher degree of conscientiousness, are more successful.
cept for one case the USD/EUR series) which might also reflect a less time and effort effect. Recent literature on “rational inattention” (Sims 2010, 2003) suggests that heterogeneity of expectations/believes is related to dispersion of information. Our results shed, indeed, light on the heterogeneity of the forecasts between two groups: the bankers and the others, and their “rational inattention”. Bankers can be considered as being more “attentive” than the other group and therefore generate less misclassification/mis-predictions than managers, and experts from insurance and industry firms. This shows that the latter may have only partial information or so called “rational inattention”, given that their day-to-day business is less related to the macro and financial variables of interest in this study. Alternative explanations given in the literature suggest that domain knowledge and experience improves forecast accuracy (Harvey, Bolger & McClelland (1994)). Stickel (1992) analyzes the performance of security analysts on the Institutional Investor All-American Research Team relative to the performance of other analysts. He shows that the members of this team are more accurate in forecasting earnings, and forecast more frequently than other analysts, suggesting that experience has a positive impact on forecasting ability. In an experiment, where information is cumulatively distributed among traders, meaning that some investors know more than others by having the same plus some extra information, Huber, Kirchler & Sutter (2008) address the question whether having more information than others about the intrinsic value of an asset always leads to higher returns when trading in financial markets. They show that only the best informed traders outperform the less informed ones. In the same manner, Ackert, Church & Zhang (2002) show that well informed traders are able to exploit their informational advantage to outperform less informed ones. Those papers support our result that bankers may possess extra information that enable them to make a better assessment about the future development of the financial variables than experts from insurance and industry firms.

The detailed estimation outputs (Tables 7 and 8) can be found in Appendix A. Table 2 contains the mean misclassification matrices ˜Π, and we observe that all matrices differ from the identity matrix 14, which clearly shows that forecasters reveal a prediction

\[ \Pi_{i,t+h} = I_K. \]

\(^{13}\)Nevertheless, Patton & Timmermann (2010) come to a kind of opposite result in their study. Patton & Timmermann (2010) indeed, show in their empirical analysis of cross-sectional dispersion in forecasters’ predictions of macroeconomic variables that heterogeneity in forecasters’ information sets does not play a major role in explaining the cross-sectional dispersion in predictions of macroeconomic variables. Their result suggests that differences in predictions/believes cannot be explained by differences in information sets.

\(^{14}\)We defined previously in Section 2 that the forecaster’s expectations are rational in the sense of Muth (1961), if the misclassification matrix is the identity matrix, \( \Pi_{i,t+h} = I_K \).
behavior which is not consistent with average rational expectations.
### 6 months ahead forecasting scenario

<table>
<thead>
<tr>
<th>Variable</th>
<th>Inflation</th>
<th>Short Term Interest Rate</th>
<th>Long Term Interest Rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>t-stat</td>
<td>Est.</td>
<td>Est.</td>
<td>t-stat</td>
</tr>
<tr>
<td>Insurance (_i)</td>
<td>-0.0130</td>
<td>-0.1929</td>
<td>-0.0382</td>
<td>-0.0017</td>
<td>-0.0162</td>
</tr>
<tr>
<td>Industry (_i)</td>
<td>0.0518</td>
<td>0.6174</td>
<td>-0.1987</td>
<td>0.5528</td>
<td>4.4048</td>
</tr>
<tr>
<td>Manager (_i)</td>
<td>-0.0769</td>
<td>-1.1210</td>
<td>0.1261</td>
<td>0.0619</td>
<td>0.7651</td>
</tr>
<tr>
<td>Reliability (_i)</td>
<td>0.0225</td>
<td>0.1623</td>
<td>0.4594</td>
<td>-0.2700</td>
<td>-1.3595</td>
</tr>
<tr>
<td>Performance(_it)</td>
<td>-1.5250</td>
<td>-61.9919</td>
<td>-0.5405</td>
<td>-1.5109</td>
<td>-25.4360</td>
</tr>
</tbody>
</table>

### perfect forecast scenario

<table>
<thead>
<tr>
<th>Variable</th>
<th>Inflation</th>
<th>Short Term Interest Rate</th>
<th>Long Term Interest Rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>t-stat</td>
<td>Est.</td>
<td>Est.</td>
<td>t-stat</td>
</tr>
<tr>
<td>Insurance (_i)</td>
<td>0.0529</td>
<td>1.3634</td>
<td>0.0211</td>
<td>0.6394</td>
<td>-0.1311</td>
</tr>
<tr>
<td>Industry (_i)</td>
<td>0.3895</td>
<td>8.9540</td>
<td>0.0454</td>
<td>0.9458</td>
<td>0.3768</td>
</tr>
<tr>
<td>Manager (_i)</td>
<td>0.0938</td>
<td>2.4947</td>
<td>0.1272</td>
<td>3.4102</td>
<td>-0.0804</td>
</tr>
<tr>
<td>Reliability (_i)</td>
<td>0.3019</td>
<td>4.1470</td>
<td>0.0005</td>
<td>0.0076</td>
<td>0.3562</td>
</tr>
<tr>
<td>Performance(_it)</td>
<td>-0.2689</td>
<td>-19.6277</td>
<td>-0.0870</td>
<td>-10.6098</td>
<td>1.2670</td>
</tr>
</tbody>
</table>

**Table 1:** Estimation results for the covariates in the misclassification probability for inflation rate, short term interest rate, long term interest rate, DAX and USD/EUR FX-rate
### 6 months ahead forecasting scenario

<table>
<thead>
<tr>
<th>Inflation</th>
<th>Short Term Interest Rate</th>
<th>Long Term Interest Rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.2961 0.4449 0.2589)</td>
<td>(0.3334 0.2056 0.4608)</td>
<td>(0.4411 0.3829 0.1758)</td>
<td>(0.6024 0.2586 0.1389)</td>
<td>(0.2493 0.2691 0.4815)</td>
</tr>
<tr>
<td>(0.1419 0.5264 0.3316)</td>
<td>(0.6175 0.0468 0.3355)</td>
<td>(0.6287 0.0391 0.3320)</td>
<td>(0.2605 0.5267 0.2126)</td>
<td>(0.8334 0.1085 0.0579)</td>
</tr>
<tr>
<td>(0.3754 0.4189 0.2055)</td>
<td>(0.0874 0.7599 0.1525)</td>
<td>(0.2707 0.6500 0.0792)</td>
<td>(0.8607 0.0441 0.0951)</td>
<td>(0.2363 0.4284 0.3351)</td>
</tr>
</tbody>
</table>

### Perfect foresight scenario

<table>
<thead>
<tr>
<th>Inflation</th>
<th>Short Term Interest Rate</th>
<th>Long Term Interest Rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.4665 0.4570 0.0764)</td>
<td>(0.5287 0.4286 0.0425)</td>
<td>(0.3393 0.4804 0.1801)</td>
<td>(0.6054 0.2562 0.1382)</td>
<td>(0.4702 0.2729 0.2567)</td>
</tr>
<tr>
<td>(0.3023 0.5942 0.1033)</td>
<td>(0.2313 0.6740 0.0945)</td>
<td>(0.6274 0.1618 0.2107)</td>
<td>(0.5960 0.1781 0.2258)</td>
<td>(0.4204 0.0017 0.5777)</td>
</tr>
<tr>
<td>(0.1582 0.3420 0.4996)</td>
<td>(0.1225 0.3328 0.5445)</td>
<td>(0.3828 0.4542 0.1629)</td>
<td>(0.6548 0.2493 0.0958)</td>
<td>(0.2298 0.3734 0.3967)</td>
</tr>
</tbody>
</table>

**Table 2**: Mean misclassification matrices $\hat{\Pi}$ for the 6 months ahead forecasting and the perfect foresight scenario inflation rate, short and long term interest rate, DAX and USD/EUR FX-rate. ACM parameter estimates of the coefficient matrix $G$ given by Equation (7).
5 Robustness Checks

In this section we perform a number of robustness checks to assess the credibility of our results. Please note, that we choose to disregard the perfect foresight scenario, and solely focus on the 6 months ahead forecasting scenario, for two particular reasons. First, as already mentioned at the beginning of the previous section, we do not find general qualitative differences in the effects of the explanatory variables on the degree of misclassification between the perfect forecast and the 6 months ahead forecasting scenarios. Second, the perfect foresight scenario represents the extreme case for which the future is already known today, and is therefore not a feasible case in reality.

First we present a series of tests of the rational expectation hypothesis. According to our previous definition in Section 2, forecasters' expectations are rational in the sense of Gourieroux & Pradel (1986) if \( \Pi_{t+h} = I_k \), but this also means that, as pointed out by Lui, Mitchell & Weale (2011), forecasters' expectations are rational if

\[
\pi_{kk} \geq \pi_{kl}, \quad l \neq k, \quad \forall k.
\]

(8)

In other words, they form their expectations according to the rational expectation hypothesis, if the probability mass lies in the no-misclassification category. Under the assumption that the forecasters make their predictions independently from each other, we can test the following null hypothesis \( \pi_{kk} = \pi_{kl} \) against the alternative \( \pi_{kk} < \pi_{kl} \). According to Lui et al. (2011) this leads to the following test statistics

\[
\sqrt{n_k} \left( \frac{\hat{\pi}_{kk} - \hat{\pi}_{kl}}{\sqrt{2\pi_{kk}}} \right) \to N(0, 1)
\]

where \( n_k \) is the number of forecasters who answered \( k \) at time \( t \).

Figures 3 to 7 plot across time, the estimated misclassification probabilities for each series under the 6 months ahead forecasting scenario. In each graphic, the thick line represents the no-misclassification case (\( \pi_{kk} \)). The red, blue and green lines denote the forecasters that replied go up, stay the same, go down, respectively. A simple visual inspection of these figures shows that condition (8) is not satisfied for most of the months, regardless of the series under scrutiny. Table 3 summarises the results of the test described above. It not only confirms the impression given by the figures, but also shows that the null hypothesis is mostly rejected (more than 50% of the times) when the true state of the world is the going down state (two last columns). In other words, it seems that our forecasters deviate more from the rational expectation hypothesis during pessimistic periods, and need more time to revise their beliefs in such situations, suggesting a kind of anchoring of their future beliefs. Nevertheless,
this does not mean that the predictions of our forecasters are not useful or reliable. As noted by Gourieroux & Pradel (1986), testing rationality as in (8) is based on the fact that forecasters form an optimal prediction given some information set. It does not mean that the forecasters use the available information in an optimal way, in order to come up with an optimal prediction.

Figure 3: Estimates of the misclassification probabilities for the inflation rate over the sample period. The graphics at the top, the middle and the bottom, display the estimated misclassification probabilities given that the true state is the going up state, respectively the stay the same and down state in the future. In each graphic, the thick line represents the no misclassification case. The other lines represent the aggregated estimated misclassification probabilities $\hat{\pi}^h_{t+h}$ for $k \neq l$. 

26
Figure 4: Estimates of the misclassification probabilities for the short term interest rate over the sample period. The graphics at the top, the middle and the bottom, display the estimated misclassification probabilities given that the true state is the going up state, respectively the stay the same and down state in the future. In each graphic, the thick line represents the no misclassification case. The other lines represent the aggregated estimated misclassification probabilities $\hat{\pi}_{t+h}^{kl}$ for $k \neq l$. 

27
Figure 5: Estimates of the misclassification probabilities for the long term interest rate over the sample period. The graphics at the top, the middle and the bottom, display the estimated misclassification probabilities given that the true state is the going up state, respectively the stay the same and down state in the future. In each graphic, the thick line represents the no misclassification case. The other lines represent the aggregated estimated misclassification probabilities $\hat{\pi}^{kl}_{t+h}$ for $k \neq l$. 
Figure 6: Estimates of the misclassification probabilities for the DAX30 over the sample period. The graphics at the top, the middle and the bottom, display the estimated misclassification probabilities given that the true state is the going up state, respectively the stay the same and down state in the future. In each graphic, the thick line represents the no misclassification case. The other line represent the aggregated estimated misclassification probabilities $\hat{\pi}_{t+h}^{kl}$ for $k \neq l$. 
Figure 7: Estimates of the misclassification probabilities for the USD/EUR currency pair over the sample period. The graphics at the top, the middle and the bottom, display the estimated misclassification probabilities given that the true state is the going up state, respectively the stay the same and down state in the future. In each graphic, the thick line represents the no misclassification case. The other lines represent the aggregated estimated misclassification probabilities $\hat{\pi}_{t+h}^{kl}$ for $k \neq l$. 
Second, in order to get more insights into the dynamics of the deviation from the rational expectation hypothesis, we construct a measure of the degree of misclassification computed as being the sum of the elements of the squared differences between the estimated misclassification matrix and the identity matrix at every point in time. The closer the value of this measure is to 0, the more the professional forecasters form their expectations rationally on average.

Figure 8 displays the fluctuations of the misclassification measure over time for all series and provides a first insight into these characteristics. We observe that during the European Currency Crisis (1992-1993), forecasters seem to deviate more from the rational expectation hypothesis for the USD/EUR currency pair (the line is in an upwards move), but less shortly thereafter. The same pattern can be observed during the Mexican Crisis, the Dot-Com Bubble, but not during the Russian Crisis or the Financial Crisis. This pattern is less clear cut for the DAX30. We observe that during crisis periods, forecasters seem to become more rational for the inflation, short term interest rate, and USD/EUR currency pair (the corresponding line exhibits a downward move). This pattern is less clear cut for the two remaining variables. These results support again the results obtained previously.

<table>
<thead>
<tr>
<th></th>
<th>$\pi^{11} = \pi^{12}$</th>
<th>$\pi^{11} = \pi^{13}$</th>
<th>$\pi^{22} = \pi^{21}$</th>
<th>$\pi^{22} = \pi^{23}$</th>
<th>$\pi^{33} = \pi^{31}$</th>
<th>$\pi^{33} = \pi^{32}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation</td>
<td>60.74</td>
<td>34.71</td>
<td>11.98</td>
<td>29.75</td>
<td>64.46</td>
<td>72.31</td>
</tr>
<tr>
<td>Short term</td>
<td>26.03</td>
<td>43.38</td>
<td>72.31</td>
<td>38.42</td>
<td>48.76</td>
<td>78.09</td>
</tr>
<tr>
<td>interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long term</td>
<td>33.47</td>
<td>17.76</td>
<td>89.25</td>
<td>70.66</td>
<td>60.33</td>
<td>97.52</td>
</tr>
<tr>
<td>interest rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAX30</td>
<td>12.39</td>
<td>5.785</td>
<td>21.07</td>
<td>10.74</td>
<td>99.58</td>
<td>4.95</td>
</tr>
<tr>
<td>USD/EUR</td>
<td>45.86</td>
<td>75.20</td>
<td>88.01</td>
<td>49.58</td>
<td>23.55</td>
<td>57.02</td>
</tr>
</tbody>
</table>

**Table 3:** Rejection rate. This table displays the number of times in percentage for which the specific null hypothesis specified in each column is rejected. A 5% significance level was used to perform the tests.
We then compute the correlation between the price changes/returns and the inverse of the misclassification measure for each series under consideration. The correlation coefficients over the complete sample period can be found in Table 4. We observe that the correlation between returns and the inverse of the misclassification measure is significantly negative only for the DAX30, and significantly positive for the macro variables. This means that when the returns go up, the misclassification measure becomes larger, implying that our forecasters deviate more from the rational expectations hypothesis when forming their predictions for the DAX30, but deviate less for the macro variables. This result suggests that our forecasters seem to form their expectations differently depending on the nature of the variable they have to predict.

<table>
<thead>
<tr>
<th>Inflation</th>
<th>Short Term Interest Rate</th>
<th>Long Term Interest Rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0746***</td>
<td>0.4001***</td>
<td>0.0231**</td>
<td>-0.4062***</td>
<td>-0.0907</td>
</tr>
</tbody>
</table>

Table 4: Correlation coefficients between the price changes/returns and the misclassification measure for the complete sample. To test the significance of the correlation coefficients, we performed a rank correlation test. ***, **, and * denote that the coefficient is statistically significantly different from zero at the 1%, 5% and 10% significance level respectively.
Third, we decide to split our sample period into crisis/turmoil periods and non crisis periods, in order to detect a difference in the forecasting behaviour of our professional forecasters. Here we are considering 5 periods of turmoil, namely the European Currency Crisis\(^\text{15}\) from April 1992 to June 1993, which peaked on the 16th September 1992, also called “Black Wednesday”, the Mexican Economic Crisis\(^\text{16}\) in 1994, the Russian financial crisis\(^\text{17}\) in 1998, the burst of the Dot-Com Bubble from March 2000 to October 2002, and the recent financial crisis from August 2007 to February 2009. These periods correspond to those already identified in Figure 8. This set-up will help us to analyse the forecasting behaviour more specifically during these particular periods.

Table 5 shows the mean misclassification matrices for the 6 months ahead forecasting scenario of all variables during crisis and non crisis periods, while Table 6 presents the correlation coefficients between price changes/returns and the inverse of the misclassification measure during the two sub periods. Altogether we consider 90 months of crisis and 152 of calm periods. We observe that the correlations are, when significant, always negative during non-crisis periods, suggesting that during those quieter periods more misclassification is generated, so that forecasters seem to deviate more from the rational expectations hypothesis on average. However, the result is less clear cut during crisis periods. We observe for both financial series a significant negative correlation. This implies that forecasters misclassify more during crisis periods, and that is why they deviate more from the rational expectation hypothesis; a phenomena also observed by examining the average estimated misclassification matrices in Table 5. On the opposite, we observe a significant positive correlation for short term interest rates, suggesting that our forecasters deviate less from the rational expectation hypothesis during crisis periods for that macro variable. This result sheds light again on the fact that our forecasters form their expectations differently depending on the nature of the variable they are asked to predict and the overall economic climate.

Overall the various robustness checks carried out enable us to draw the following conclusions. First of all, our results suggest that the forecasters in our dataset do not rely on a clearly pre-defined rule to form their expectations. They seem to form their expectations differently depending on the variables (macro or financial variables) they

\(^{15}\) It reflects the withdraw of the UK and Italy from the European Monetary System.

\(^{16}\) The Mexican Economic Crisis saw the devaluation of the Mexican Peso, with a bailout mostly funded by the US.

\(^{17}\) It resulted in the Russian government and the Russian Central Bank devaluing the ruble and defaulting on its debt.
are asked to forecast. We noticed, indeed, that they deviate more from the rational expectation hypothesis for the DAX30 and less for the macro variables when the complete sample is considered. Second, we find a different behaviour during crisis and non-crisis periods. We observe, indeed, more misclassification during non-crisis periods regardless of the variable under consideration (the correlations are always negative). Whereas this result holds for the financial variables during crisis periods, it reverses for the short term interest rate. One could argue, that there is more homogenous information available for macro variables, so that during crisis periods forecasters are more focused, interpret this homogenous information in the same way, and therefore generate less mis-predictions. On the opposite, given the easy access to information about financial markets, and given the fact that this information is highly dispersed and heterogenous, forecasters have more difficulties to come up with a consensus during crisis periods, and therefore this situation leads to more misclassification, i.e., they deviate more from the rational expectation hypothesis.
### Table 5: Mean misclassification matrices $\tilde{\Pi}$ for the 6 months ahead forecasting scenario for the inflation rate, short and long term interest rate, DAX and USD/EUR FX-rate during crisis and non-crisis periods.

<table>
<thead>
<tr>
<th></th>
<th>Inflation</th>
<th>Short Term Interest Rate</th>
<th>Long Term Interest Rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crisis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3306 0.4352 0.234)</td>
<td>(0.3072 0.2202 0.4725)</td>
<td>(0.3784 0.3688 0.2527)</td>
<td>(0.6939 0.2233 0.0827)</td>
<td>(0.2429 0.2856 0.4713)</td>
</tr>
<tr>
<td></td>
<td>(0.1492 0.4224 0.428)</td>
<td>(0.6009 0.0491 0.3498)</td>
<td>(0.6004 0.0363 0.3631)</td>
<td>(0.1476 0.5574 0.2948)</td>
<td>(0.8445 0.1025 0.0529)</td>
</tr>
<tr>
<td></td>
<td>(0.2933 0.4296 0.277)</td>
<td>(0.0823 0.7131 0.2044)</td>
<td>(0.2262 0.6889 0.0847)</td>
<td>(0.7790 0.0753 0.1456)</td>
<td>(0.2715 0.4170 0.3114)</td>
</tr>
<tr>
<td><strong>Non-Crisis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2756 0.4507 0.2736)</td>
<td>(0.3489 0.1971 0.4539)</td>
<td>(0.4782 0.3913 0.1303)</td>
<td>(0.5482 0.2795 0.1722)</td>
<td>(0.2530 0.2593 0.4875)</td>
</tr>
<tr>
<td></td>
<td>(0.1376 0.5880 0.2743)</td>
<td>(0.6274 0.0454 0.3271)</td>
<td>(0.6455 0.0408 0.3136)</td>
<td>(0.3274 0.5085 0.1639)</td>
<td>(0.8269 0.1121 0.0608)</td>
</tr>
<tr>
<td></td>
<td>(0.4241 0.4125 0.1632)</td>
<td>(0.0905 0.7876 0.1217)</td>
<td>(0.2971 0.6269 0.0759)</td>
<td>(0.9081 0.0256 0.0651)</td>
<td>(0.2155 0.4352 0.3491)</td>
</tr>
</tbody>
</table>
### crisis periods

<table>
<thead>
<tr>
<th></th>
<th>Inflation</th>
<th>Short Term Interest Rate</th>
<th>Long Term Interest Rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.1784</td>
<td>0.5106***</td>
<td>-0.2717</td>
<td>-0.3961**</td>
<td>-0.0808*</td>
<td></td>
</tr>
</tbody>
</table>

### non crisis periods

|                | -0.5225*** | -0.4203*** | -0.0202*** | -0.0394** | 0.0574  |

Table 6: Correlation coefficients between the returns and the misclassification measure for the two sub-samples. To test the significance of the correlation coefficients, we performed a rank correlation test. 

***, **, and * denote that the coefficient is statistically significantly different from zero at the 1%, 5% and 10% significance level respectively.
6 Concluding Remarks

In this paper, we present a new empirical approach to analyse individual expectation formation processes based on tendency survey data. Using a quantal response model with misclassification we define qualitative mispredictions in terms of deviations from the qualitative rational expectation forecast and relate them to individual and macro factors driving these individual mispredictions. Our approach is consistent with the rather general definition of expectations as subjectively held beliefs by individuals. Since individual expectations are taken from individuals’ qualitative responses to survey questions, no assumption on the individual expectation formation process is necessary. In this sense our approach is double robust. First, it is robust against the critique of classical tests of the rational expectation hypothesis based on aggregate time series data which require distributional and/or functional form assumptions. Second, the approach does not require an aggregation rule by considering individual forecasts at any time period with the equivalent outcome at the macro level.

First of all our results show that surveys are useful to predict macro-variables, as already mentioned in Ang et al. (2007), and represent a timely source of information (Lui et al. (2010)). Second, we show that our results are not sensitive to the choice of the specific forecasting model for the 6 months ahead prediction of the objective macro world up/same/down probabilities. For all series (inflation rate, short and long term interest rate, DAX30, and the USD/EUR exchange rate with respect to Germany) considered, we find a specific learning pattern. We observe a general long term learning effect, reflected by the fact that forecasters who gave a better assessment in the past continue to do so in the future, whereas we do not observe a general short term learning effect. Furthermore, we observe for the DAX30 series that the forecasters do not learn from short term past misprediction if these financial series are in a downward move but rather accelerate their degree of misprediction. For the USD/EUR series and the short term interest rate respectively, we observe a similar but less pronounced effect for upward state, and for the downward state respectively. Moreover, we also observe that forecasters who are more reliable in returning the questionnaire in time do not provide better forecasts, and that managers, experts from insurance and industry firms seem to generate a higher degree of misclassification than bankers, indicating that better informed participants outperform the less informed ones. All these results contribute to the literature on delayed information processing (Mankiw & Reis (2010), Andrade & LeBihan (2013), among others) and also on the literature related to the concept of inattentiveness (Reis (2006a, b)).
The estimation of the misclassification matrix allow us not only to systematically analyse forecasting behavior at the individual level, but also in terms of average forecasts for specific groups of forecasters or the overall group of survey respondents. Our results suggest that our forecasters form their expectations differently depending on the variables they are predicting, deviating more for the DAX30 and less for the macro variables when the whole sample period is considered. Because the underlying time series are reasonably long, estimates of certain subperiods were used for a systematic analysis of individual forecasting behavior in different scenarios (crisis times vs calm times). More misclassification was observed during non-crisis periods regardless of the variable under consideration as well as for the financial variables during crisis periods. However, we found less misclassification for the short term interest rate during crisis periods. This findings suggest that forecasters might have asymmetric, different loss functions, or that substantial groups of forecasters do not have rational expectations, as pointed out by Das, Dominitz & van Soest (1999).

In future research our approach can be used to test specific expectation hypotheses or learning algorithms. Moreover, the degree of homogeneity of the individual misclassification matrices could also be used as a simple measure of the degree of consensus among forecasters. Finally, the estimates can be used to construct subsamples of superior forecasters to improve the overall forecasting performance of survey data.
References


## A Appendix

### GLARMA Model

<table>
<thead>
<tr>
<th>Par.</th>
<th>Est.</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>0.1916</td>
<td>10.8864</td>
</tr>
<tr>
<td>$c_2$</td>
<td>-1.1169</td>
<td>-33.7432</td>
</tr>
<tr>
<td>$c_3$</td>
<td>4.4144</td>
<td>35.8312</td>
</tr>
<tr>
<td>$c_4$</td>
<td>2.0866</td>
<td>36.2563</td>
</tr>
<tr>
<td>$c_5$</td>
<td>-2.7614</td>
<td>-103.8120</td>
</tr>
<tr>
<td>$c_6$</td>
<td>-1.4682</td>
<td>-74.9082</td>
</tr>
<tr>
<td>$c_7$</td>
<td>0.1916</td>
<td>10.8864</td>
</tr>
<tr>
<td>$c_8$</td>
<td>-1.1169</td>
<td>-33.7432</td>
</tr>
<tr>
<td>$c_9$</td>
<td>4.4144</td>
<td>35.8312</td>
</tr>
<tr>
<td>$c_{10}$</td>
<td>2.0866</td>
<td>36.2563</td>
</tr>
<tr>
<td>$c_{11}$</td>
<td>-2.7614</td>
<td>-103.8120</td>
</tr>
<tr>
<td>$c_{12}$</td>
<td>-1.4682</td>
<td>-74.9082</td>
</tr>
</tbody>
</table>

### Covariates

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Est.</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insurance$_i$</td>
<td>0.0529</td>
<td>1.3634</td>
</tr>
<tr>
<td>Industry$_i$</td>
<td>0.3895</td>
<td>8.9540</td>
</tr>
<tr>
<td>Manager$_i$</td>
<td>0.0938</td>
<td>2.4947</td>
</tr>
<tr>
<td>Performance$_it$</td>
<td>-0.2689</td>
<td>-19.6277</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Est.</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insurance$_i$</td>
<td>0.0529</td>
<td>1.3634</td>
</tr>
<tr>
<td>Industry$_i$</td>
<td>0.3895</td>
<td>8.9540</td>
</tr>
<tr>
<td>Manager$_i$</td>
<td>0.0938</td>
<td>2.4947</td>
</tr>
<tr>
<td>Performance$_it$</td>
<td>-0.2689</td>
<td>-19.6277</td>
</tr>
</tbody>
</table>

### Mean ln $L$

| Mean ln $L$ | -283.259 | -247.965 | -290.743 | -257.910 | -300.036 |

| Mean ln $L$ | -283.259 | -247.965 | -290.743 | -257.910 | -300.036 |

### Table 7: Estimation results for the perfect forecast scenario: inflation rate (column 1), short term interest rate (column 2), short term interest rate (column 3), DAX (column 4), USD/EUR TX-rate (column 5).
<table>
<thead>
<tr>
<th>Par.</th>
<th>inflation</th>
<th>short term interest rate</th>
<th>long term interest rate</th>
<th>DAX</th>
<th>USD/EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>t-stat</td>
<td>Est.</td>
<td>t-stat</td>
<td>Est.</td>
</tr>
<tr>
<td>c1</td>
<td>2.8633</td>
<td>78.0191</td>
<td>2.9608</td>
<td>34.1409</td>
<td>2.6906</td>
</tr>
<tr>
<td></td>
<td>-2.2461</td>
<td>-50.1302</td>
<td>-4.7365</td>
<td>-162.2098</td>
<td>-5.4109</td>
</tr>
<tr>
<td>c3</td>
<td>-0.1956</td>
<td>-6.9362</td>
<td>-4.6996</td>
<td>-106.5669</td>
<td>-1.4709</td>
</tr>
<tr>
<td>c4</td>
<td>-2.2107</td>
<td>-48.8996</td>
<td>-3.4183</td>
<td>-91.3994</td>
<td>-1.1911</td>
</tr>
<tr>
<td>a13</td>
<td>9.3718</td>
<td>1.2477</td>
<td>-4.1347</td>
<td>-12.6560</td>
<td>7.6473</td>
</tr>
<tr>
<td>a14</td>
<td>21.2628</td>
<td>2.8512</td>
<td>15.7637</td>
<td>93.6643</td>
<td>2.4513</td>
</tr>
<tr>
<td>a22</td>
<td>-2.1322</td>
<td>-3.7526</td>
<td>36.6687</td>
<td>73.9736</td>
<td>-3.6931</td>
</tr>
<tr>
<td>a31</td>
<td>6.0001</td>
<td>83.9612</td>
<td>29.5327</td>
<td>120.4267</td>
<td>8.5399</td>
</tr>
<tr>
<td>a32</td>
<td>8.9718</td>
<td>102.6522</td>
<td>31.3908</td>
<td>134.0914</td>
<td>11.0552</td>
</tr>
<tr>
<td>a33</td>
<td>16.3218</td>
<td>11.4899</td>
<td>0.1667</td>
<td>0.0030</td>
<td>14.2125</td>
</tr>
<tr>
<td>a34</td>
<td>21.2459</td>
<td>14.9157</td>
<td>24.9244</td>
<td>34.2886</td>
<td>11.3041</td>
</tr>
<tr>
<td>a42</td>
<td>3.8339</td>
<td>58.0894</td>
<td>26.4144</td>
<td>112.6414</td>
<td>4.2386</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance_i</td>
<td>0.0130</td>
<td>-0.1929</td>
<td>0.1082</td>
<td>-0.5971</td>
<td>0.0017</td>
</tr>
<tr>
<td>Industry_i</td>
<td>0.0518</td>
<td>0.6174</td>
<td>0.1987</td>
<td>0.5807</td>
<td>0.5528</td>
</tr>
<tr>
<td>Reliability_i</td>
<td>0.0769</td>
<td>-1.1212</td>
<td>0.1261</td>
<td>1.5626</td>
<td>0.0619</td>
</tr>
<tr>
<td>Performance_i</td>
<td>0.0225</td>
<td>0.1623</td>
<td>0.4594</td>
<td>2.7248</td>
<td>-0.2700</td>
</tr>
</tbody>
</table>

**Table 8:** Estimation results for the 6 months ahead forecasting scenario: inflation rate (column 1), short term interest rate (column 2), short term interest rate (column 3), DAX (column 4), USD/EUR FX-rate (column 5).