What Determines Forecasters' Forecasting Errors?

Ingmar Nolte*
Lancaster University
CoFE

Sandra Nolte†
Lancaster University

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

This version: January 30, 2014

*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 comment we like to thank Karim Abadir and Angelo de Melino.

‡Department of Economics, Box 124, University of Konstanz, 78457 Konstanz, Germany. email: Winfried.Pohlmeier@uni-konstanz.de.

Konstanzer Online-Publikations-System (KOPS)
URL: http://nbn-resolving.de/urn:nbn:de:bsz:352-0-385063

Electronic copy available at: http://ssrn.com/abstract=1541799

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 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. This enables a detailed analysis of the individual expectation formation process.

JEL classification: $C23, C25, D84, E27$

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

Electronic copy available at: http://ssrn.com/abstract=1541799

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 the 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)) and delayed (only a share of individuals receives up-to-date information, c.f. Calvo (1983), Mankiw & Reis (2002) and Lorenzoni (2010)) information have been developed which provide more realistic mechanisms for information processing. A recent 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 based and believed based models of aggregate stock market behavior.[1] In their capital asset pricing model Barberis et al. (2013) assume besides traditional rational traders who extrapolate their believes 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 (2013)) than traditional rational expectations based models..

This paper tries to contribute 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 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. While pure time series approaches require some functional form

[1] Barberis, Greenwood, Jin & Shleifer (2013), Table 1 surveys a range of models of aggregate stock market behavior.

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 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. In principle, heterogeneity in beliefs and information sets can 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 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 going 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 of 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.

Twenty seven 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.[2] 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

[2]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.

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. 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)). The 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 ('the ugly state') in qualitative survey data, which has no exact mathematical counterpart in the rational expectations framework. In business survey data the no-change category simply serves 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.[3] 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 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.

[3] Experimental 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.

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 non-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, while Section 5 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 $\{\mathcal{S}_k, k = 1, \ldots, K\}$ and $\{\mathcal{G}_l, l = 1, \ldots, L\}$ as two given partitions of the outcome space of Y_t^*. Let Y_t be the discrete counterpart of Y_t^* within a threshold crossing ordinal response model taking on the form of the kth unit-vector for $Y_t^* \in \mathcal{S}_k$, i.e., if $Y_t^* \in \mathcal{S}_k$ the kth 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 lth unit-vector of dimension L indicating that $Y_t^* \in \mathcal{G}_l$. 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 $\mathcal{S}_k$ and let $\mathcal{F}_t$ be the information set given at time t, then for Y_{t+h} there exists a K dimensional vector $\mathbf{P}_{t+h}$ with the typical element

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

that given the information set $\mathcal{F}_t$ the outcome Y_{t+h}^* at time $t + h$ lies in $\mathcal{S}_k$. Since $\mathrm{E}[Y_{k,t+h}|\mathcal{F}_t] = \Pr[Y_{k,t+h} = 1|\mathcal{F}_t]$ the k-th element of $\mathbf{P}_{t+h}$ 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}. $\mathbf{P}_{i,t+h}$ denotes the L dimensional vector of probabilities

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

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.[4]

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

$$\mathbf{P}_{i,t+h} = \Pi'_{i,t+h} \mathbf{P}_{t+h}, \tag{1}$$

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

The individual misclassification matrix, $\Pi'_{i,t+h}$, defined in (1) relates the objective expectation formation process at the macro level[5] to the individual (possibly subjective) expectation formation process at the micro level. If $\mathbf{P}_{t+h}$ 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 $\mathcal{S}_k = \mathcal{G}_k, K = L$, the forecaster's expectations are rational in the sense of Muth (1961), if the misclassification matrix is the identity matrix, $\Pi_{i,t+h} = \mathrm{I}_K$. 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 $\mathbf{P}_{i,t+h}$ 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

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

[5]We refer to objective expectation and objective probabilities as those implied by the (theoretical) model on the macro level, see Muth (1961).

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. evolvement 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 $\bar{\mathbf{P}}_{t+h}$ be the average expectation over the state vector Y_{t+h} defined as

$$\bar{\mathbf{P}}_{t+h} = \sum_{i=1}^{N_t} w_{it}\mathbf{P}_{i,t+h} = \sum_{i=1}^{N_t} w_{it}\mathrm{E}[Y_{t+h}|\mathcal{F}_{it}] \tag{2}$$

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

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

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

$$u_{i,t+h} = \big(\mathrm{I}_K - \Pi'_{i,t+h}\big)\mathbf{P}_{t+h} + u_{t+h}.$$

Aggregating over the number of forecasters in survey wave t gives the average expectation error $\bar{u}_{t+h} = \sum_{i=1}^{N_t} w_{it}u_{i,t+h}$

$$\bar{u}_{t+h} = \big(\mathrm{I}_K - \bar{\Pi}'_{t+h}\big)\mathbf{P}_{t+h} + u_{t+h}.$$

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 $\mathrm{I}_K = \bar{\Pi}'_{t+h}$, so that average expectations equal rational expectations.[6] Using the average misclassification matrix $\bar{\Pi}'_{t+h}$ we are able to test the

[6] q.m. stands for convergence in quadratic mean.

expectation formation in terms of a 'consensus' or market concept. Note that $\Pi'_{i,t+h}$ 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 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.

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.

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

$$\ln \mathcal{L} = \sum_{t=1}^{T} \sum_{i=1}^{N_t} \sum_{l=1}^{L} \mathbb{1}_{\{\tilde{Y}_{l,t+h|\mathcal{F}_{it}}=1\}} \ln \mathbf{P}_{it+h,l}, \tag{4}$$

where $\mathbf{P}_{it+h,l} = \sum_{k=1}^{K} \pi_{it+h}^{kl} \cdot \mathbf{P}_{t+h,k}$. In order to facilitate the computation of the likelihood $\mathbf{P}_{t+h,k}$ 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 Π_{it+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 (τ_{up}) and down (τ_{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 $\mathbf{P}_{t+h}$ at time t. We assume the following time series specification for Y^*_{t+h}:

$$Y^*_{t+h} = \mu_{t+h} + \varepsilon_{t+h},$$

where $\mu_{t+h} = \mathrm{E}[Y^*_{t+h}|\mathcal{F}_t]$ 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 μ_{t+h} appropriately. The kth component of $\mathbf{P}_{t+h}$ can then be obtained by

$$\mathbf{P}_{t+h,k} = F_{t+h}\big(\tau^k_{t+h} - \mu_{t+h}\big) - F_{t+h}\big(\tau^{k-1}_{t+h} - \mu_{t+h}\big), \text{ for } k = 1, \ldots, K,$$

where F_{t+h} denotes the conditional cumulative density function (c.d.f.) of ε_{t+h} given $\mathcal{F}_t$ and τ^k_{t+h} the kth threshold, where we set for convenience $\tau^K_{t+h} = \infty$ and $\tau^0_{t+h} = -\infty$ for all t. Assuming an ordered probit model we obtain:

$$\mathbf{P}_{t+h,k} = \Phi\left(\frac{\tau^k_{t+h} - \mu_{t+h}}{\sigma_{t+h}}\right) - \Phi\left(\frac{\tau^{k-1}_{t+h} - \mu_{t+h}}{\sigma_{t+h}}\right), \text{ for } k = 1, \ldots, K,$$

where Φ denotes the standard normal c.d.f. and $\sigma_{t+h} = \mathrm{E}[\sigma_{t+h}|\mathcal{F}_t]$, with σ_t the standard deviation of ε_t.

h-Step Ahead Forecast In our empirical study we obtain the forecasts $\hat{\mathbf{P}}_{t+h}$ by estimating σ_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 σ_{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 σ_{t+h} and Y^*_{t+h} already at time t. Hence, we obtain $\hat{\mathbf{P}}_{t+h}$ 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 $\hat{\mathbf{P}}_{t+h}$. An alternative strategy consist of deriving the probability vector $\hat{\mathbf{P}}_{t+h}$ by dynamic quantile regression based forecasts of Y^*_{t+h}, in which the respective quantiles, which then yield $\hat{\mathbf{P}}_{t+h}$, 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 $\hat{\mathbf{P}}_{t+h}$ directly with the discrete information using for example an Autoregressive Conditional Multinomial (ACM) model of Russell & Engle (2002).

Misclassification Matrix

Recall the misclassification matrix Π_{it+h}, whose components are given by

$$\pi_{i,t+h}^{kl} = \Pr\left[\tilde{Y}_{l,t+h|\mathcal{F}_{i,t}} = 1 \middle| Y_{k,t+h|\mathcal{F}_t} = 1\right]. \tag{5}$$

Thus $\pi_{i,t+h}^{kl}$ 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 $\pi_{i,t+h}^{kk}$ and the sum over l of $\sum_l \pi_{i,t+h}^{kl}$ 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

$$\pi_{i,t+h}^{kl} = \frac{\exp\{\Lambda_{i,t+h}^{kl}\}}{\sum_{l=1}^{3} \exp\{\Lambda_{i,t+h}^{kl}\}}, \qquad k, l = 1, 2, 3, \tag{6}$$

where the log-odds ratio $\Lambda_{i,t+h}^{kl}$ will be specified below. As normalization constraint, we use as the reference category the corresponding correct macro outcome category, i.e. $\Lambda_{i,t+h}^{kk} = 0$ 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

$$\begin{aligned}
\Lambda_{i,t+h}^{k} &= \left(\Lambda_{i,t+h}^{kl}, \Lambda_{i,t+h}^{kl'}\right)' \\
&= \left(\ln\left[\frac{\pi_{i,t+h}^{kl}}{\pi_{i,t+h}^{kk}}\right], \ln\left[\frac{\pi_{i,t+h}^{kl'}}{\pi_{i,t+h}^{kk}}\right]\right)', \text{for } k, l, l' = 1, 2, 3, \text{ and } k \neq l \neq l',
\end{aligned}$$

is specified as a multivariate ARMA type process, including explanatory variables statically

$$\begin{aligned}
\Lambda_{i,t+h}^{k} &= G^k Z_{i,t+h} + \alpha_{i,t+h}^{k} \\
\alpha_{i,t+h}^{k} &= c^k + \sum_{j=1}^{q} A_j^k \xi_{i,t+h-j} \quad \forall k = 1, 2, 3,
\end{aligned} \tag{7}$$

with $\{A_j^k, j : 1 \to q\}$ being a matrix of dimension (2×3) with the elements $\{a_{ll'}^{k(j)}\}$, for $l, l' = 1, 2, 3$ with $l \neq k$. c^k denotes the (2×1) vector of constants $c^k = (c_l^k, c_{l'}^k)'$,

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

$$\begin{aligned}
\xi_{i,t+h} &= (\xi^1_{i,t+h}, \xi^2_{i,t+h}, \xi^3_{i,t+h})' \\
&= \begin{pmatrix} \left(\mathbb{1}_{\{\tilde{Y}_{2,t+h|\mathcal{F}_{it}}=1\}} + \mathbb{1}_{\{\tilde{Y}_{3,t+h|\mathcal{F}_{it}}=1\}}\right) \cdot \mathbb{1}_{\{Y_{1,t+h}=1\}} \\ \left(\mathbb{1}_{\{\tilde{Y}_{1,t+h|\mathcal{F}_{it}}=1\}} + \mathbb{1}_{\{\tilde{Y}_{3,t+h|\mathcal{F}_{it}}=1\}}\right) \cdot \mathbb{1}_{\{Y_{2,t+h}=1\}} \\ \left(\mathbb{1}_{\{\tilde{Y}_{1,t+h|\mathcal{F}_{it}}=1\}} + \mathbb{1}_{\{\tilde{Y}_{2,t+h|\mathcal{F}_{it}}=1\}}\right) \cdot \mathbb{1}_{\{Y_{3,t+h}=1\}} \end{pmatrix},
\end{aligned}$$

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^k_j implies lower values in the corresponding log-odds ratio vector $\Lambda^k_{i,t+h}$ and thus a specific learning effect. Consider for example the case for A^1_j which is given by[7]

$$A^1_j = \begin{pmatrix} a^1_{21} & a^1_{22} & a^1_{23} \\ a^1_{31} & a^1_{32} & a^1_{33} \end{pmatrix},$$

and measures the influence of $\xi_{i,t+h}$ on $\Lambda^1_{i,t+h}$, which is given by

$$\Lambda^1_{i,t+h} = \left(\ln \left[\frac{\pi^{12}_{i,t+h}}{\pi^{11}_{i,t+h}} \right], \ln \left[\frac{\pi^{13}_{i,t+h}}{\pi^{11}_{i,t+h}} \right] \right)',$$

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^1_j 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 $\Lambda^1_{i,t+h}$ (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^1_j represents a learning effect for an up assessment with regard to a misprediction for a previous no change (down) movement. For A^2_j and A^3_j we obtain the corresponding interpretations. In the case that a column in A^k_j is positive we observe no learning but a more pronounced and accelerated degree of misprediction. In the case that we observe alternating signs

[7] We suppress j.

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 three classes: purely time dependent, purely individual specific and both, time and individual specific variables. We consider only one time specific variable, namely Volatility_t, which is at time t the historical 6 months realized volatility from $t-6$ to t and represents a general level of uncertainty at the date at which the survey participants make their assessments.

We specify four individual specific explanatory variables: Insurance_i, Industry_i, Manager_i and Reliability_i. 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_i and Industry_i to be equal to 1 if individual i 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_i is equal to 1 if individual i 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_i is defined as the share of the number of questionnaires returned in time over the overall returned number of questionnaires for each individual i. 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_{it} is a time and individual specific variable defined as the share of correct predictions over the last 12 months of individual i 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 $\xi^k_{i,t+h}$ 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.

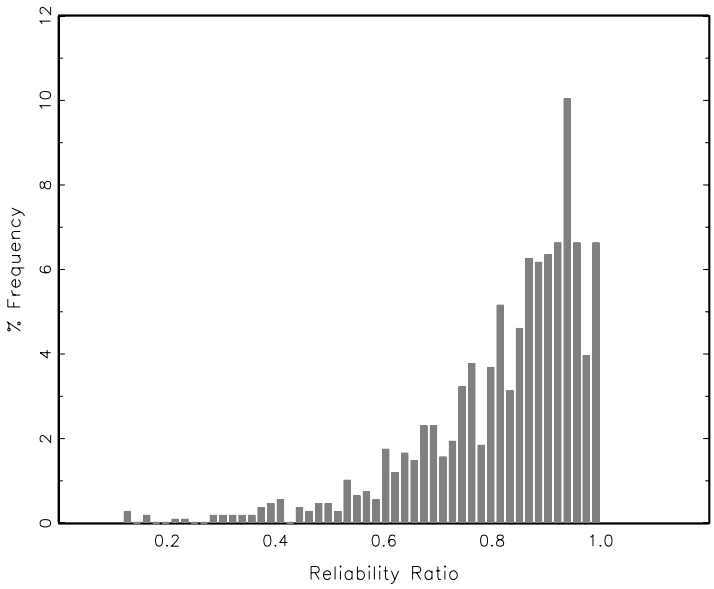

Figure 1: Distribution of the response rate across experts

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, the objective probabilities from the 6 months ahead forecasting scenario do anticipate their general behavior also very well but to a less amplified extent. This is of course due to the the specific forecasting setup for probabilities that adds additional forecasting uncertainty and thus contaminates 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 $t + 6$.

Simply from a pure inspection of the (up/same/down) sample probabilities 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 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 are going to analyze within the proposed modeling framework.

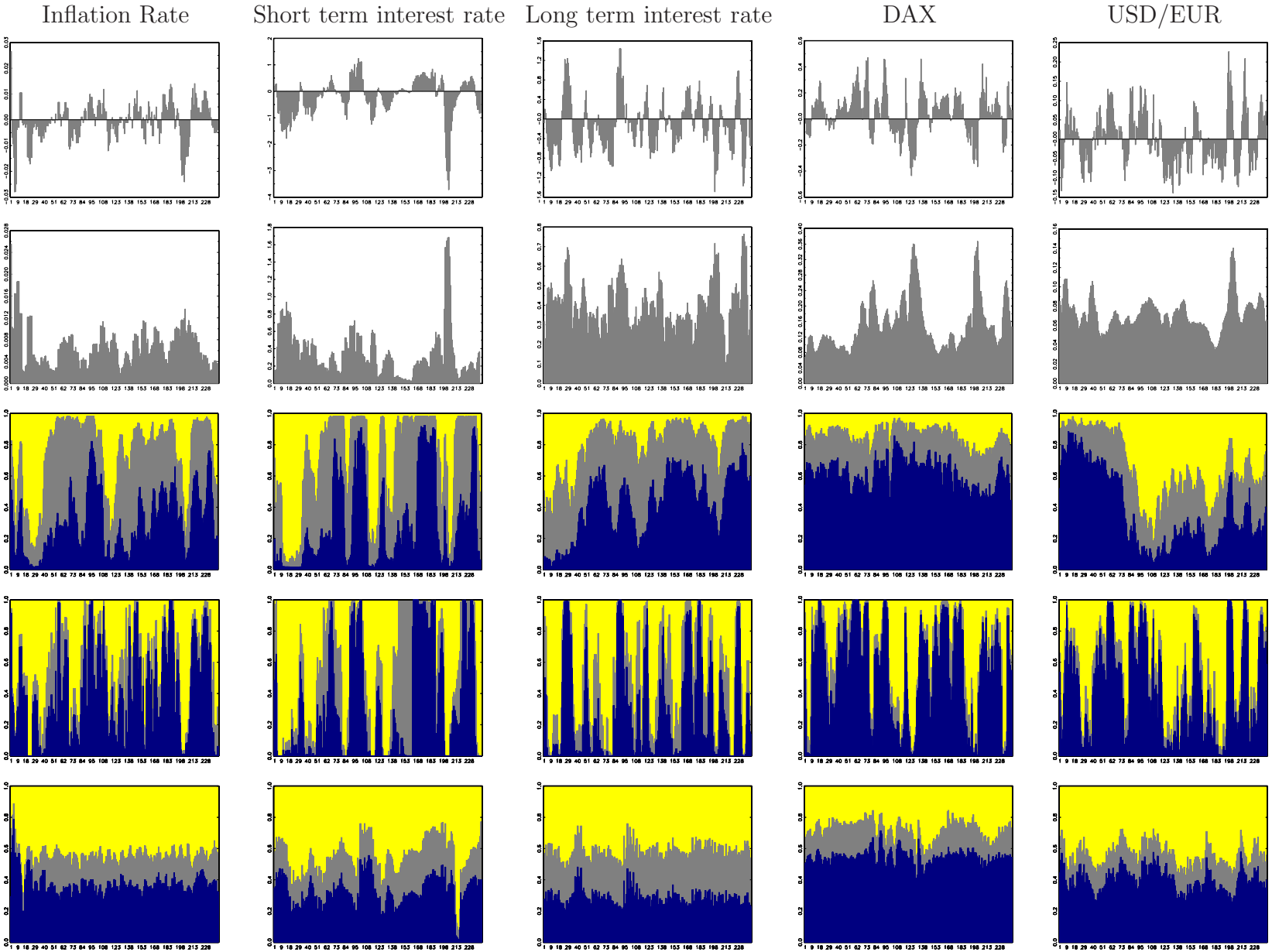

Figure 2: The first row displays the graphs of the 6 months changes (inflation and interest rates) or the 6 months returns (DAX 30 and USD/EUR) of the actual series, second row: 6 months standard deviation estimated with a realized volatility estimator, third row: shares for up/same/down answered by 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.

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 DAX that the A_1^3 matrices have significant positive coefficients, implying that forecasters do not learn from short term past mis-prediction if 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. On the contrary, their degree of mis-prediction is accelerated. This can either be related to bubbles observed for these series or attributed to an optimistic bias.[8] If we take a closer look at the graphic in the third row

[8]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.

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, respectively the short term interest rate, we observe a similar but less pronounced effect for upward state matrices A_1^1, respectively the for the downward state matrices A_3^1. On the contrary, we see that for the inflation the A_1^1 matrix has significant negative coefficients, implying that forecasters do learn from short term past mis-prediction if this financial series is in an upward move. These 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 (2008) 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 consistent with the information processing mechanism put forward in macroeconomic models on delayed information processing, see Calvo (1983), Mankiw & Reis (2002) and Lorenzoni (2010).

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 provide better forecasts. For the DAX 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. Nevertheless, 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.

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 (except for one case) which might also reflect a less time and effort effect. The literature suggests 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 the forecast 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 on the intrinsic value of an asset always leads to higher returns when trading on 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 volatility variable is if significantly different from zero always negative in the perfect forecast scenario for all 5 series, and for the financial series (DAX30 and USD/EUR) in the 6 months ahead forecasting scenario, except for two macroeconomic series (short term and long term interest rates) also in the 6 months ahead forecasting scenario. This might reflect temporary down or upward tendencies which yield to a higher degree of predictability. This result is in contradiction with the findings in the literature. Research in judgemental forecasting indicates that volatility of past price series might have a significant impact on investors' forecasting behaviors.[9] Du & Budescu (2007) investigate the influence of past volatility on forecasts and subjective beliefs regarding future stock prices. They test the impact of volatility on the

[9]See Lawrence, Goodwin, O'Connor & Önkal (2006) for a review of judgmental forecasting.

forecasts of future prices, and their accuracy. They show that participants' assessment about future prices are more accurate for low volatility stocks than for high volatility stocks.

The detailed estimation outputs (Table 3 and Table 4) can be found in Appendix A. Table 2 contains the mean misclassification matrices $\tilde{\Pi}$, and we observe that all matrices differ from the identity matrix[10], which clearly shows that forecasters reveal a prediction behavior which is not consistent with average rational expectations.

[10] 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, $\Pi_{i,t+h} = \mathrm{I}_K$.

6 months ahead forecasting scenario

	Inflation		Short Term Interest Rate		Long Term Interest Rate		DAX		USD/EUR	
Variable	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Volatility_t	-0.1468	-0.1411	0.9911	51.3523	2.4838	25.6591	-3.2513	-16.7939	-11.9419	-35.0511
Insurance_i	-0.0130	-0.1929	-0.0382	-0.5512	-0.0017	-0.0162	0.2473	3.4395	0.3018	4.0894
Industry_i	0.0518	0.6174	-0.1987	-1.5807	0.5528	4.4048	0.9237	10.1283	-0.0374	-0.4207
Manager_i	-0.0769	-1.1210	0.1261	1.5626	0.0619	0.7651	0.1549	1.9657	-0.1706	-2.7296
Reliability_i	0.0225	0.1623	0.4594	2.7248	-0.2700	-1.3595	1.1797	6.9640	-0.2141	-1.4012
Performance_{it}	-1.5250	-61.9919	-0.5405	-32.1726	-1.5109	-25.4360	-3.4863	-50.8207	-1.9912	-46.8518

perfect forecast scenario

	Inflation		Short Term Interest Rate		Long Term Interest Rate		DAX		USD/EUR	
Variable	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Volatility_t	-71.0181	-111.4534	-3.5510	-279.6063	-0.8480	-33.2549	-0.5303	-5.5182	-7.9877	-35.2658
Insurance_i	0.0529	1.3634	0.0211	0.6394	-0.1311	-2.9795	0.1655	4.2219	0.0858	2.5385
Industry_i	0.3895	8.9540	0.0454	0.9458	0.3768	5.5493	0.4560	9.2683	0.0706	1.4379
Manager_i	0.0938	2.4947	0.1272	3.4102	-0.0804	-1.8483	0.0408	0.9358	0.0035	0.0895
Reliability_i	0.3019	4.1470	0.0005	0.0076	0.3562	3.2799	0.7057	7.2528	0.1576	1.9267
Performance_{it}	-0.2689	-19.6277	-0.0870	-10.6098	1.2670	64.6429	-1.4607	-93.0382	-0.3850	-27.1127

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

Inflation	Short Term Interest Rate	Long Term Interest Rate	DAX	USD/EUR
$\begin{pmatrix} 0.2961 & 0.4449 & 0.2589 \\ 0.1419 & 0.5264 & 0.3316 \\ 0.3754 & 0.4189 & 0.2055 \end{pmatrix}$	$\begin{pmatrix} 0.3334 & 0.2056 & 0.4608 \\ 0.6175 & 0.0468 & 0.3355 \\ 0.0874 & 0.7599 & 0.1525 \end{pmatrix}$	$\begin{pmatrix} 0.4411 & 0.3829 & 0.1758 \\ 0.6287 & 0.0391 & 0.3320 \\ 0.2707 & 0.6500 & 0.0792 \end{pmatrix}$	$\begin{pmatrix} 0.6024 & 0.2586 & 0.1389 \\ 0.2605 & 0.5267 & 0.2126 \\ 0.8607 & 0.0441 & 0.0951 \end{pmatrix}$	$\begin{pmatrix} 0.2493 & 0.2691 & 0.4815 \\ 0.8334 & 0.1085 & 0.0579 \\ 0.2363 & 0.4284 & 0.3351 \end{pmatrix}$

perfect foresight scenario

Inflation	Short Term Interest Rate	Long Term Interest Rate	DAX	USD/EUR
$\begin{pmatrix} 0.4665 & 0.4570 & 0.0764 \\ 0.3023 & 0.5942 & 0.1033 \\ 0.1582 & 0.3420 & 0.4996 \end{pmatrix}$	$\begin{pmatrix} 0.5287 & 0.4286 & 0.0425 \\ 0.2313 & 0.6740 & 0.0945 \\ 0.1225 & 0.3328 & 0.5445 \end{pmatrix}$	$\begin{pmatrix} 0.3393 & 0.4804 & 0.1801 \\ 0.6274 & 0.1618 & 0.2107 \\ 0.3828 & 0.4542 & 0.1629 \end{pmatrix}$	$\begin{pmatrix} 0.6054 & 0.2562 & 0.1382 \\ 0.5960 & 0.1781 & 0.2258 \\ 0.6548 & 0.2493 & 0.0958 \end{pmatrix}$	$\begin{pmatrix} 0.4702 & 0.2729 & 0.2567 \\ 0.4204 & 0.0017 & 0.5777 \\ 0.2298 & 0.3734 & 0.3967 \end{pmatrix}$

Table 2: Mean misclassification matrices $\tilde{\Pi}$ 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 Concluding Remarks

In this paper, we present a new empirical approach to analyze expectation formation based on tendency survey data on individual expectations. 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.

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 five financial 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 in the sense that 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, respectively the short term interest rate, we observe a similar but less pronounced effect for upward state, respectively for the downward state. 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.

The estimation of the misclassification matrix allow us to systematically analyze 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. The degree of homogeneity of the individual misclassification matrices can be used as a simple measure of the degree of consensus among forecasters. In future research our approach

can be used to test specific expectation hypotheses or learning algorithms. Because the underlying time series are reasonably long, estimates of certain subperiods may be used for a systematic analysis of individual forecasting behavior in different scenarios (crisis times vs calm times). Finally, the estimates can be used to construct subsamples of superior forecasters to improve the overall forecasting performance of survey data.

References

Ackert, L., B. K. Church, & P. Zhang (2002): "Market Behavior in the Presence of Divergent and Imperfect Private Information: Experimental Evidence from Canada, China, and the United States," *Journal of Economic Behavior and Organization*, 47, 435–450.

Anderson, O. (1951): "Konjunkturtest und Statistik, Möglichkeiten und Grenzen einer Quantifizierung von Testergebnissen," *Allgemeines Statistisches Archiv*, 35, 209–220.

Barberis, N., R. Greenwood, L. Jin, & A. Shleifer (2013): "X-CAPM: An Extrapolative Capital Asset Pricing Model," Working Paper 19189, National Bureau of Economic Research.

Barrick, M. & M. Mount (1991): "The Big Five Personality Dimensions and Job Performance: a Meta-Analysis," *Personnel Psychology*, 44, 1–26.

Binder, M. & M. H. Pesaran (1998): "Decision Making in the Presence of Heterogeneous Information and Social Interactions," *International Economic Review*, 39, 1027–1052.

Calvo, G. A. (1983): "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12 (3), 383–398.

Carlson, J. A. & M. Parkin (1975): "Inflation Expectations," *Economica*, 42, 123–138.

Coibion, O. & Y. Gorodnichenko (2012): "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy*, 120, 116–159.

Du, N. & D. Budescu (2007): "Does Past Volatility Affect Investors' Price Forecasts and Confidence Judgements?" *International Journal of Forecasting*, 23, 497–511.

Dustmann, C. & A. van Soest (2004): "An Analysis of Speaking Fluency of Immigrants Using Ordered Response Models with Classification Errors," *Journal of Business and Economic Statistics*, 22, 312–321.

Fischhoff, B. (1975): "Hindsight $\neq$ Foresight: the Effect of Outcome Knowledge on Judgment under Uncertainty," *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288–299.

Gourieroux, C. & J. Pradel (1986): "Direct Test of the Rational Expectation Hypothesis," *European Economic Review*, 30, 265–284.

Greenwood, R. & A. Shleifer (2013): "Expectations of Returns and Expected Returns," Working Paper 18686, National Bureau of Economic Research.

Harvey, N., F. Bolger, & A. McClelland (1994): "On the Nature of Expectations," *British Journal of Psychology*, 85, 203–229.

Hausman, J. (2001): "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left," *The Journal of Economic Perspectives*, 15, 57–67.

Hausman, J., J. Abreyava, & F. M. Scott-Morton (1998): "Misclassification of the Depedent Variable in a Disctrete Response Setting," *Journal of Econometrics*, 87, 239–269.

Huber, J., M. Kirchler, & M. Sutter (2008): "Is More Information Always Better? Experimental Financial Markets with Cumulative Information," *Journal of Economic Behavior and Organization*, 65, 86–104.

König, H., M. Nerlove, & G. Oudiz (1981): "On the Formation of Price Expectations: An Analysis of Business Test Data by Log-Linear Probability Models," *European Economic Review*, 16, 103–138.

Lawrence, M., P. Goodwin, M. O'Connor, & D. Önkal (2006): "Judgemental Forecasting: A Review of Progress over the Last 25 Years," *International Journal of Forecasting*, 22, 493–518.

Lorenzoni, G. (2010): "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," *Review of Economic Studies*, 77 (1), 305–338.

Mankiw, N. G. & R. Reis (2002): "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 117 (4), 1295–1328.

—— (2010): "Chapter 5 - Imperfect Information and Aggregate Supply," *in:* Friedman, B. M. & M. Woodford, eds., "Handbook of Monetary Economics," vol. 3, Elsevier, pp. 183–229.

Meyvis, T., R. Ratner, & J. Levav (2008): "Why We Don't Learn to Accurately Forecast Our Feelings: How Misremembering Our Predictions Blinds Us to Our Past Forecasting Errors," Tech. rep., Working Paper, http://ssrn.com/abstract=1135805.

Mitchell, J., R. Smith, & M. R. Weale (2013): "Efficient Aggregation of Panel Qualitative Survey Data," *Journal of Applied Econometrics*, 28, 580–603.

Muth, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29 (3), 315–335.

Nerlove, M. (1983): "Expectations, Plans, and Realizations in Theory and Practice," *Econometrica*, 51 (5), 1251–1279.

Pesaran, M. & M. Weale (2006): "Survey Expectations," *in:* Elliott, G., C. Granger, & A. Timmermann, eds., "Handbook of Economic Forecasting," vol. 1, chap. 14, Elsevier, pp. 715–776.

Pesaran, M. H. (1984): "Expectations Formations and Macro-Econometric Modelling," *in:* Malgrange, P. & P.-A. Muet, eds., "Contemporary Macroeconomic Modelling," Basil Blackwell, Oxford, pp. 27–55.

—— (1985): "Formation of Inflation Expectations in British Manufacturing Industries," *The Economic Journal*, 95, 948–975.

—— (1987): *The Limits to Rational Expectations*, Basil Blackwell, Oxford.

Robertson, I. & A. Kinder (1993): "Personality and Job Competencies: the Criterion-Related Validity of Some Personality Variables," *in:* Cooper, C. & D. Bosseau, eds., "Trends in Organisational Behaviour," vol. 1, Wiley, pp. 75–89.

Russell, J. R. & R. F. Engle (2002): "Econometric Analysis of Discrete-Valued Irregularly-Spaced Financial Transactions Data," Revised Version of Discussion Paper 98-10, University of California, San Diego.

Stickel, S. E. (1992): "Reputation and Performance Among Security Analysts," *The Journal of Finance*, 47, 1811–1836.

Weinstein, N. (1980): "Unrealistic optimism about future life events," *Journal of Personality and Social Psychology*, 39, 806–820.

Wilson, T. D., J. Meyers, & D. T. Gilbert (2001): "Lessons From the Past: Do People Learn From Experience That Emotional Reactions Are Short-Lived?" *Personality and Social Psychology Bulletin*, 27, 1648–1661.

Woodford, M. (2002): "Imperfect Common Knowledge and the Effects of Monetary Policy," *in:* Aghion, P., R. Frydman, J. Stiglitz, & M. Woodford,

eds., “Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps,” Princeton University Press, pp. 25–58.

A Appendix

Par.	inflation Est.	inflation t-stat	short term interest rate Est.	short term interest rate t-stat	long term interest rate Est.	long term interest rate t-stat	DAX Est.	DAX t-stat	USD/EUR Est.	USD/EUR t-stat
	GLARMA Model									
c_2^1	0.1916	10.8864	-1.5506	-119.2769	0.0691	2.4504	-0.8988	-11.1652	-2.2111	-101.4266
c_3^1	-1.1169	-33.7432	-4.8696	-125.1825	-2.1717	-72.1495	-1.3781	-19.3553	-3.3579	-170.4518
c_1^2	4.4144	35.8312	-1.9272	-87.6000	-4.3394	-30.2819	2.8853	9.1918	5.1416	0.4347
c_3^2	2.0086	36.2563	0.2907	20.0483	4.1956	24.0436	-1.3275	-4.3269	-0.0728	-0.0032
c_1^3	-2.7614	-103.8120	-1.6294	-95.2865	0.3389	21.0497	1.2349	29.1250	-3.2590	-152.2897
c_2^3	-1.4682	-74.9082	-0.5235	-48.0275	0.6595	25.1718	0.4150	6.1209	0.0650	2.7778
a_{21}^1	-1.5333	-58.5229	0.6013	43.2590	0.6043	13.4888	1.2266	6.4626	3.0495	65.4399
a_{31}^1	-2.4668	-39.0935	-0.4384	-3.0110	3.2440	94.5773	0.8304	5.1228	5.4364	188.1107
a_{22}^1	-2.6326	-77.4294	5.5720	170.3976	-0.5167	-9.1290	1.0057	9.3380	2.1208	60.4217
a_{32}^1	-6.6444	-57.6271	9.3487	134.1277	0.7952	12.7029	0.7804	7.5183	3.9075	174.4420
a_{23}^1	-0.9652	-45.1028	8.2725	188.8699	0.2267	6.2110	0.5060	5.4176	2.8647	93.3127
a_{33}^1	-1.1866	-22.4310	7.3165	45.6425	1.5819	38.8673	0.3734	4.2097	3.7412	157.1933
a_{11}^2	-7.0258	-38.6245	1.1316	38.2297	21.6073	9.9573	-6.2898	-6.7163	12.4827	0.1151
a_{31}^2	-19.5392	-3.6328	-6.8375	-137.2992	-12.7806	-3.7167	5.8418	7.7932	-11.8902	-0.0159
a_{12}^2	-20.5203	-53.6619	1.8007	53.4332	6.9608	29.6709	0.7750	0.6654	-8.3229	-0.3199
a_{32}^2	-30.3038	-5.7227	-10.4772	-74.8906	-8.9927	-27.2671	4.6943	4.3166	8.3811	0.2190
a_{13}^2	-44.1393	-18.2183	3.0537	57.9450	5.9565	35.4132	-0.4187	-1.0002	-14.5468	-0.1952
a_{33}^2	-6.4510	-51.2798	-3.9391	-46.5615	-4.9208	-20.4692	-0.7220	-0.4670	20.7472	0.2480
a_{11}^3	2.1871	58.0133	1.7246	94.2404	-0.1916	-12.6053	4.1867	27.1687	2.0869	59.9684
a_{21}^3	1.0521	41.0977	1.1294	110.7255	-0.2828	-8.5438	3.9334	16.3009	-0.3077	-9.2402
a_{12}^3	0.0650	1.7016	-0.3412	-12.3177	1.2224	55.3122	3.7208	10.8605	4.3030	135.3145
a_{22}^3	-0.4380	-15.8123	-0.1078	-7.8116	0.2704	6.2884	3.6899	9.4954	-0.0810	-1.2796
a_{13}^3	1.9052	62.2614	-0.5497	-8.6160	0.5439	22.9494	0.8153	15.1825	5.0720	173.1058
a_{23}^3	1.2343	62.3384	0.5249	30.6959	0.6947	17.9974	0.4917	5.6323	-0.4631	-13.9488
	Covariates									
Volatility$_t$	-71.0181	-111.4534	-3.5510	-279.6063	-0.8480	-33.2549	-0.5303	-5.5182	-7.9877	-35.2658
Insurance$_i$	0.0529	1.3634	0.0211	0.6394	-0.1311	-2.9795	0.1655	4.2219	0.0858	2.5385
Industry$_i$	0.3895	8.9540	0.0454	0.9458	0.3768	5.5493	0.4560	9.2683	0.0706	1.4379
Manager$_i$	0.0938	2.4947	0.1272	3.4102	-0.0804	-1.8483	0.0408	0.9358	0.0035	0.0895
Reliability$_i$	0.3019	4.1470	0.0005	0.0076	0.3562	3.2799	0.7057	7.2528	0.1576	1.9267
Performance$_{it}$	-0.2689	-19.6277	-0.0870	-10.6098	1.2670	64.6429	-1.4607	-93.0382	-0.3850	-27.1127
Mean $\ln \mathcal{L}$	-283.259		-247.965		-290.743		-257.910		-300.036	

Table 3: 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).

Par.	inflation Est.	inflation t-stat	short term interest rate Est.	short term interest rate t-stat	long term interest rate Est.	long term interest rate t-stat	DAX Est.	DAX t-stat	USD/EUR Est.	USD/EUR t-stat
	GLARMA Model									
c_2^1	2.8633	78.0191	2.9608	34.1499	2.6968	20.3840	0.0164	0.1116	-8.4136	-75.0544
c_3^1	-2.2461	-50.1362	-4.7365	-162.2089	-5.4109	-55.7825	-1.6734	-12.6294	-4.8502	-84.4983
c_1^2	-5.2781	-18.9519	2.2650	16.6667	-4.4995	-9.5897	3.4532	10.3266	4.8241	2.0869
c_3^2	4.1260	60.5874	3.9490	23.5480	4.1796	9.1337	2.7987	8.5326	3.3623	1.4090
c_1^3	-0.1956	-6.9362	-4.6996	-106.5669	-1.4709	-15.2425	1.5402	7.8823	-5.2013	-74.3043
c_2^3	-2.1907	-48.8996	-3.4183	-91.3984	-1.1911	-12.1417	-1.2074	-1.4186	2.1779	48.2905
a_{21}^1	-4.0768	-63.6006	-30.9707	-56.4850	-15.1632	-8.5215	1.8376	5.7859	-13.1783	-0.0061
a_{31}^1	3.9774	97.2469	5.5559	209.6566	9.3596	65.5893	5.5448	19.6903	9.1772	93.9324
a_{22}^1	9.3178	1.2477	-4.1347	-12.6560	7.6473	1.0135	2.0809	9.0395	-0.6472	-0.0001
a_{32}^1	21.2628	2.8512	15.7637	93.6643	2.4513	0.0025	3.3947	16.4074	11.2790	113.1294
a_{23}^1	-3.4514	-64.7542	-9.7899	-11.4878	-4.9675	-26.0898	-1.1440	-6.4306	45.4784	6.4151
a_{33}^1	-2.1322	-3.7526	36.6687	73.9736	-3.6931	-0.0566	-1.9122	-4.3838	40.3182	5.6841
a_{11}^2	-0.8027	-0.0471	31.0830	1.3828	22.0604	14.5729	-7.8284	-8.8727	25.6128	0.1631
a_{31}^2	-7.2316	-55.1609	-31.6840	-1.3185	-10.1338	-6.5731	-8.3559	-7.8644	-8.6322	-0.0402
a_{12}^2	12.6179	20.4869	-0.8653	-4.8072	10.2979	13.1435	0.2465	0.4235	-17.8566	-0.1926
a_{32}^2	-27.2339	-11.2124	-18.2751	-8.2865	-3.3662	-4.3057	-1.4596	-2.5589	-6.9552	-2.6100
a_{13}^2	-3.5376	-0.0187	-25.2171	-0.9636	7.9175	14.4744	-9.5520	-1.9901	4.8755	0.2754
a_{33}^2	-5.6926	-45.2872	15.1311	0.6766	-2.4354	-4.4136	-2.9733	-7.1166	0.9516	0.0519
a_{11}^3	6.7001	83.9612	29.5527	120.4267	8.5309	20.1533	10.9493	16.3033	0.2206	0.0556
a_{21}^3	8.9718	102.6522	31.3908	134.0914	11.0552	24.7708	0.8613	0.0616	-0.9720	-18.1682
a_{12}^3	16.3218	11.4869	0.1667	0.0030	14.2125	8.8568	10.0431	1.1838	10.7679	91.2534
a_{22}^3	21.2459	14.9157	24.9244	34.2886	11.3041	7.1283	1.3968	0.0079	-0.1777	-1.5629
a_{13}^3	-1.5127	-20.6936	26.4336	109.1846	-3.2266	-3.2533	0.9551	4.2734	9.0057	102.9223
a_{23}^3	3.8339	58.0894	26.4144	112.6414	4.2386	24.2900	2.0427	2.1347	-2.8484	-33.4712
	Covariates									
Volatility$_t$	-0.1468	-0.1411	0.9911	51.3523	2.4838	25.6591	-3.2513	-16.7939	-11.9419	-35.0511
Insurance$_i$	-0.0130	-0.1929	-0.0382	-0.5512	-0.0017	-0.0162	0.2473	3.4395	0.3018	4.0894
Industry$_i$	0.0518	0.6174	-0.1987	-1.5807	0.5528	4.4048	0.9237	10.1283	-0.0374	-0.4207
Manager$_i$	-0.0769	-1.1210	0.1261	1.5626	0.0619	0.7651	0.1549	1.9657	-0.1706	-2.7296
Reliability$_i$	0.0225	0.1623	0.4594	2.7248	-0.2700	-1.3595	1.1797	6.9640	-0.2141	-1.4012
Performance$_{it}$	-1.5250	-61.9919	-0.5405	-32.1726	-1.5109	-25.4360	-3.4863	-50.8207	-1.9912	-46.8518
Mean $\ln \mathcal{L}$	-292.948		-254.543		-283.246		-283.274		-284.488	

Table 4: 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).