Ebc European Banking Center

Discussion paper

# UNCERTAINTY AND DISAGREEMENT IN FORECASTING INFLATION: EVIDENCE FROM THE LABORATORY

By Damjan Pfajfar, Blaž Žakelj

April 27, 2011

European Banking Center Discussion Paper
No. 2011-014

This is also a CentER Discussion Paper
No. 2011-053

ISSN 0924-7815

CentER

TILBURG UNIVERSITY

# Uncertainty and Disagreement in Forecasting Inflation: Evidence from the Laboratory*

Damjan Pfajfar[†]
CentER, EBC, University of Tilburg

Blaž Žakelj[‡]
European University Institute

May 2, 2011

Abstract. We establish several stylized facts about the behavior of individual uncertainty and disagreement between individuals when forecasting inflation in the laboratory. Subjects correctly perceive the underlying inflation uncertainty in only 60% of cases, which can be interpreted as the overconfidence bias. Determinants of individual uncertainty, disagreement among forecasters and properties of aggregate distribution are analyzed. We find that the interquartile range of the aggregate distribution has the highest correlation with inflation variability; however the average confidence interval performs best in a forecasting exercise. Allowing subjects to insert asymmetric confidence intervals results in wider upper intervals than lower intervals on average, thus perceiving higher uncertainty with respect to inflation increases. In different treatments we study the influence of different monetary policy designs on the formation of confidence bounds. Inflation targeting produces lower uncertainty and higher accuracy of intervals than inflation forecast targeting.

**JEL:** C91, C92, E37, D80

**Keywords:** Laboratory Experiments, Confidence Bounds, New Keynesian Model, Inflation Expectations

*We would like to thank Seppo Honkapohja, Rosemarie Nagel, Cars Hommes, Chryssi Giannitsarou, Karl Schlag and participants at the 2010 North American ESA Conference in Tucson, 2010 Money, Macro and Finance Annual Conference in Limassol, 2010 Computing in Economics and Finance conference in London, and 2010 EEA Annual congress in Glasgow for their comments and suggestions. Blaž Žakelj gratefully acknowledges financial support of the Max Weber Fellowship. Damjan Pfajfar gratefully acknowledges funding from a Marie Curie project PIEF-GA-2009-254956 – EXPMAC. All remaining errors are ours.

[†]Department of Economics, Faculty of Economics and Business Administration, P.O. Box 90153, NL-5000 LE Tilburg, Netherlands. *E-mail*: D.Pfajfar@uvt.nl. *Web*: https://sites.google.com/site/dpfajfar/.

[‡]Address: Villa La Fonte, Via delle Fontanelle 10, 50014 San Domenico di Fiesole, Italy. *Email*: blaz.zakelj@eui.eu.

## 1. Introduction

This paper discusses an experimental study on the expectations formation process and associated uncertainty within a macroeconomic framework. The importance of inflation uncertainty has been recognized at least since Friedman's Nobel Lecture (Friedman, 1977). He argues that higher rates of inflation are associated with higher inflation variability, which in turn causes a reduction in the efficiency of the price system and leads into the reduction of output due to institutional rigidities. Levi and Makin (1980) and Mullineaux (1980) have found empirical support for the hypothesis that higher inflation uncertainty is associated with lower output. This represents a clear rationale why central banks should care about inflation uncertainty. Inflation targeting central banks in particular have recently increased their interest in the distribution of inflation expectations. They trust their communication strategies to play an important role in the shaping of inflationary expectations. Both individual uncertainty and disagreement (interpersonal uncertainty) can therefore be viewed as a measure of the effectiveness of central banks' communication strategies. With some central banks these strategies also include publishing their probabilistic forecasts of inflation in terms of fan charts. More generally, credibility of inflation targets can be assessed using both the point forecast and agents' perceived uncertainty. As Giordani and Söderlind (2003) demonstrate this is particularly relevant when there is a regime switch.[1] In his speech about Federal Reserve communications, Mishkin (2008) stresses that the cost of inflation should be viewed both in terms of levels and of its uncertainty. This claim is actually consistent with the standard New Keynesian dynamic stochastic general equilibrium (DSGE) model, where it has been shown that in order to maximize consumer welfare the central bank should minimize variation of inflation (see e.g. Woodford, 2003).[2]

In our experiment subjects are introduced to fictitious economy described with series of inflation, interest rates and output gap. They are asked to forecast inflation and to provide 95% confidence intervals around their point forecasts. These forecasts are then fed into the simplified version of the New Keynesian model which gives us the realized values of inflation, output gap and interest rates. These values are displayed to the subject and the process is iterated. This allows us to study the individual uncertainty about forecasts as well as the disagreement on the point forecasts.[3] We compare our results to those in the survey data literature aimed at distinguishing between uncertainty and disagreement and evaluating their relation to inflation variability.

We study the determinants of different measures of inflation uncertainty proposed in the literatures and evaluate which measure should be used as a proxy of inflation variance. We also focus on the relationship between monetary policy and inflation uncertainty and examine whether some environments are better then others at stabilizing both inflation and its uncer-

[1] See also Evans and Wachtel (1993).

[2] Recognizing the importance of different aspects of expectations distribution Lorenzoni (2010) shows that monetary policy affects agents (with different pieces of information) differently, arguing that there is a trade-off between aggregate and cross-sectional efficiency.

[3] Our companion paper (Pfajfar and Žakelj, 2011) focuses on the inflation expectation formation mechanism and its relation to monetary policy, i.e. how should the monetary policy be designed in order to be robust for the potential presence of heterogeneous expectations.

tainty. We study two different monetary policy rules: inflation targeting and inflation forecast targeting. For the latter we use three different specifications of the coefficient that describes the reaction of interest rates to the deviations of inflation forecasts from the inflation target. We find that the design significantly affects both the width and the accuracy of forecast intervals. In particular, we find that the instrumental rule that reacts to current inflation reduces overall uncertainty and increases subjects' forecast accuracy compared to the rules that react to expected inflation. Most of these differences can be attributed to the fact that the contemporaneous rule (inflation targeting) produces lower variability of actual inflation. However there are some treatment effects that go beyond the interest rate channel. Similar evidence is also observed for the treatment where the central bank reacts more strongly to the deviations of inflation expectations from inflation target.

Results on the analysis of the behavior of individual confidence intervals suggest that the width of confidence interval is highly inertial and it increases when inflation is below the target level. This contrasts the results of the survey data literature, where it is a high inflation that usually leads to an increase in uncertainty. However, our results show little evidence of different degrees of uncertainty in different phases of the business cycle.

Which representation of inflation expectations is most relevant for monetary authority? The forecast ability of different measures has been mostly examined using the survey data of professional forecasters.[4] Three measures have been predominantly used in the survey data literature: standard deviation of point forecasts, average individual forecast error variance, and the variance of the aggregate distribution. They are complementary in terms of informative value. The first measure describes disagreement but says little about uncertainty and the second captures uncertainty but disregards the disagreement. Zarnowitz and Lambros (1987) show that there can be substantial differences between the variation of disagreement and the variation of uncertainty. Variance of the aggregate distribution of forecasts gives information about both, uncertainty and disagreement; however it is difficult to separate the two effects. In our setup we can compare different measures obtained from the individual responses and the aggregate distribution and study their ability to forecast inflation variability. We find that average confidence intervals perform best in the forecasting exercise, although simple correlation analysis shows that the interquartile range of the aggregate distribution ($IQR$) is a measure that has the highest correlations with the variability of inflation.

Several dynamic panel data regressions are designed to identify the determinants of three measures discussed above. Disagreement among subjects measured with standard deviation of point forecasts increases when average group forecast error increases and when inflation is below the target level. Similar explanatory variables also affect the individual uncertainty although disagreement is arguably less inertial. All factors that enter significantly to the specification for uncertainty and disagreement are by definition also important for the interquartile range.

[4] See Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003) for example. One disadvantage of survey data is that panel members do not always provide their forecast and also the panel pool changes continously. See Engelberg, Manski, and Williams (2009) for some other methodological issues involved. A laboratory environment presents a potential solution to this problem.

Indeed, inflation, mean forecast error and lagged interquartile range exert significant effects.

When looking at individual responses we also find that forecasters usually tend to underestimate the underlying uncertainty when forecasting inflation as only 60% of realization falls within the specified 95% interval. Giordani and Söderlind (2003) reach similar conclusions when analyzing the survey data of professional forecasters. An observation that subjects tend to report narrower confidence intervals than the one asked for is a well-known fact, labelled as the "overconfidence effect." This issue has been extensively debated in the experimental psychology literature. A common approach in the latter literature is to frame the experiment in the context of stock market forecasting exercises.[5]

D'Amico and Orphanides (2008), Giordani and Söderlind (2003) and Rich and Tracy (2003) all argue that the observed confidence intervals of forecasters in the survey data are usually symmetric. Also studies in psychological literature usually assume symmetric confidence intervals (see O'Connor, Remus, and Griggs, 2001 for a discussion). Symmetric intervals are easier to handle in empirical analysis when we want to construct the aggregate distribution of expectations, because we can simply assume that individual's distribution is normally distributed. Furthermore there are no theoretical reasons in our model why confidence intervals should not be symmetric, as the underlying model and the distribution of shocks do not exhibit any asymmetries. In the field data, this might not necessarily be the case as there are several documented potential asymmetries, especially asymmetric monetary policy effects over the business cycle. We have decided to perform treatment A with the restriction to symmetric confidence intervals while in treatment B we test this assumption and allow subjects to have potentially asymmetric intervals.[6] However, only 12.5% of confidence intervals are symmetric when we allow subjects to report asymmetric confidence intervals. Du and Budescu (2007) and O'Connor, Remus, and Griggs (2001) also point out that confidence intervals tend to be asymmetric. Du and Budescu (2007) explain the use of asymmetry with the hedging effect, where subjects tend to provide slightly more optimistic point forecasts and hedge for this risk by inserting skewed confidence intervals. They also find a negative relationship between asymmetric confidence intervals and the volatility of the underlying series. Our results suggest that there is less asymmetry when there is an upward path of output gap (expansion) and when inflation is below the target level.

Experimental economic research on forecasting uncertainty has been less abundant than survey based research. Fehr and Tyran (2008) ask the subjects to provide descriptive measures of their confidence level (but do not perform any analysis on it), while we ask subjects to provide numerical responses. Similarly Bottazzi and Devetag (2005) ask subjects to provide 95% confidence intervals in an asset pricing experiment, with the aim (almost exclusively) of defining average forecast and not of studying the behavior of uncertainty or disagreement.[7]

[5]For surveys, see Hoffrage (2004) or Lichtenstein, Fischhoff, and Phillips (1982) (see also e.g. Oskamp, 1965, Lawrence and O'Connor, 1992, Muradoglu and Onkal, 1994, Gilovich, Griffin, and Kahneman, 2002). These studies usually do not provide payment for the accuracy or the width of the confidence intervals, only for the accuracy of the point forecasts.

[6]Engelberg, Manski, and Williams (2009) documents another potential asymmetry in the forecasting process (on which we do not focus), i.e. asymmetry between central tendencies of subjective distributions and point forecasts.

[7]Bottazzi, Devetag, and Pancotto (2010) argue that asking for the confidence intervals instead of point predic-

Our focus is also quite different to that of psychology experiments. These usually limit their attention to independent event forecasts, while the present study concentrates on a series of (dependant) forecasts. This allows us to perform time-series analysis of confidence bounds. We also provide subjects with other relevant information (besides the past history of prices) that might influence confidence. In this way we are able to examine whether confidence intervals are affected by stages of the business cycle.

This paper is organized as follows: Section 2 describes the model and the experimental design; in Section 3 we focus on the analysis of individual responses while in Section 4 we analyze disagreement and properties of aggregate distribution; Section 5 discusses and assesses the forecasting ability of different measures, while Section 6 concludes.

## 2. Experimental Design

We design an experiment where subjects participate in a fictitious economy and are asked to provide inflation forecasts and a measure of uncertainty about their forecasts. The mean of point forecast is then used by the data generating process for calculating inflation, interest rate, and output gap. These variables are available to subjects before the next period forecast. These, so-called "learning to forecast" experiments have been conducted before within a simple macroeconomic setup (e.g. Williams, 1987; Marimon, Spear, and Sunder, 1993; Evans, Honkapohja, and Marimon, 2001; Arifovic and Sargent, 2003; Adam, 2007) and also within the asset pricing framework (see Hommes, Sonnemans, Tuinstra, and van de Velden, 2005 and Anufriev and Hommes, 2007). Closest to our framework, but with a different focus, are experiments by Adam (2007) and Assenza, Heemeijer, Hommes, and Massaro (2011). In this paper we decided to focus on the reduced form of the New Keynesian (NK) model, where we can clearly observe forecasts and study their relationship with monetary policy. Of course, there is a trade-off between using the model from "first principles" and employing a reduced form. The former has the advantage of setting the objectives (payoff function) exactly in line with microfoundations, however forecasts are difficult to elicit in this environment where subjects act as producers and consumers and interact on labor and final product markets and do not explicitly provide their forecasts (for the latter approach, see Noussair, Pfajfar, and Zsiros, 2011). Therefore, an appropriate framework for the question that we address in this paper is the "learning to forecast" design where incentives are set in order to induce the most accurate forecasts as possible.[8] We first present the model and then focus on the design.

Data generating process is forward-looking sticky price NK monetary model with different monetary policy reaction functions.[9] The baseline framework in the NK approach is a dy-

tions in asset pricing framework has the effect of reducing price fluctuations and increasing subjects' coordination on a common prediction strategy.

[8] In this framework, thus, we do not assign subjects a particular role in the economy, rather they act as "professional" forecasters. One way to think about the relation between "professional forecasters" and consumers/firms is that these economic subjects employ professional forecasters to provide them with forecasts of inflation.

[9] The advantage of this small-scale NK model is that it reproduces relatively well several stylized facts about major economies and is the simplest model that is widely used for policy analysis by central banks and governments. However, there are two implicit complications for participants. First, it requires forecasting two periods ahead. It would definitely be easier for participants to produce a one period ahead forecast (sometimes called "nowcasting"). The second drawback is that the standard forward-looking NK models assume that agents have

namic stochastic general equilibrium model with money, nominal price rigidities, and rational expectations. The model consists of a forward-looking Phillips curve (PC), an IS curve, and a monetary policy reaction function.[10] The information set at the time of forecasting consists of macro variables at the time $t-1$, although the forecasts are made in period $t$ for period $t+1$. Mathematically we denote this as $E_t\pi_{t+1}$. Strictly speaking, it should be denoted as $E_t\left(\pi_{t+1}|\mathcal{I}_{t-1}\right)$. In the experiment we replace $E_t\pi_{t+1}$ by $\frac{1}{K}\sum^k \pi^k_{t+1|t}$ , where $\pi^k_{t+1|t}$ is $k$-subject' point forecast of inflation ($K$ is total number of subjects in an economy).

The IS curve is specified as follows:

$$y_t = -\varphi\left(i_t - E_t\pi_{t+1}\right) + y_{t-1} + g_t, \tag{1}$$

where interest rate is $i_t$, $\pi_t$ denotes inflation, $y_t$ is output gap, and $g_t$ is an exogenous shock. The parameter $\varphi$ is the intertemporal elasticity of substitution in demand. We calibrate it to 0.164.[11] We can observe that we do not have expectations of output gap in the specification. Instead, we have lagged output gap.[12] Compared to purely forward-looking specifications, our model might display more persistence in the output gap. This is the most significant departure from the otherwise standard macroeconomic model.

Aggregating across the price setting decisions of individual firms yields the linear relationship in the equation (2). Thus, the supply side of the economy is summarized in the following PC:

$$\pi_t = \lambda y_t + \beta E_t\pi_{t+1} + u_t. \tag{2}$$

The longer prices are fixed on average, i.e. the smaller is $\lambda$, the less sensitive inflation is to the current output gap. The parameter $\beta$ is the subjective discount rate. According to McCallum and Nelson (2004) calibration, we set $\lambda = 0.3$ and $\beta = 0.99$. The shocks $g_t$ and $u_t$ are uncorrelated and unobservable to subjects and follow an AR(1) process:

$$\begin{aligned} g_t &= \kappa g_{t-1} + \widetilde{g}_t; \\ u_t &= \nu u_{t-1} + \widetilde{u}_t, \end{aligned}$$

where $0 < |\kappa| < 1$ and $0 < |\nu| < 1$. $\widetilde{g}_t$ and $\widetilde{u}_t$ are independent white noise, $\widetilde{g}_t \backsim N\left(0, \sigma_g^2\right)$ and $\widetilde{u}_t \backsim N\left(0, \sigma_u^2\right)$. In the NK literature it is standard to assume AR(1) shocks. $g_t$ could be justified as government spending shock or taste shocks and standard interpretation of $u_t$ is the technology shock. Empirical literature finds these shocks to be quite persistent (see e.g.

to forecast both inflation and output gap. We decided to simplify this experiment by asking only for expectations of inflation.

[10] Detailed derivations are in, for example, Woodford (1996), or textbooks such as Walsh (2003) or Woodford (2003).

[11] We implement McCallum and Nelson (2004) calibration, which is one of the standard calibrations for NK models.

[12] In principle, one could argue that this specification of IS equation corresponds to the case when subjects have naive expectations on output gap or it is assumed extreme case of habit persistence. We do not ask subjects to forecast both inflation and output gap. We were afraid that this would represent a too difficult task for them (also as they have to forecast for two periods ahead). We leave the fully forward-looking NK model for future work.

Cooley and Prescott, 1995 or Ireland, 2004). In particular, $\kappa$ and $\nu$ are calibrated to 0.6, while their standard deviations are 0.08. The treatments are fully comparable as we have exactly the same shocks in all treatments. The introduction of shocks as an exogenous unobservable component in the law of motion for observable variables is an important source of uncertainty in our experiment. It helps to avoid outcomes where all agents would coordinate on the forecasts identical to the inflation target and maintains the focus on the process of learning to forecast.

To close the model, the interest rate rule has to be specified. We use two alternative monetary policy reaction functions in different experimental treatments: (i) inflation forecast targeting where the interest rate is set in response to inflation expectations; we study three different parametrizations of this rule, and (ii) instrumental rule where interest rates are set in response to current inflation – inflation targeting.

We start with the following inflation forecast targeting rule:

$$i_t = \gamma \left( E_t \pi_{t+1} - \overline{\pi} \right) + \overline{\pi}. \tag{3}$$

Here the central bank responds to deviations of inflation from the target, $\overline{\pi}$. We vary $\gamma$ in different treatments. In order to have inflation in positive numbers for most of the periods we set the inflation target to $\overline{\pi} = 3$.

The alternative that we study is inflation targeting. Here contemporaneous rather than forecasted inflation is encompassed in the monetary policy rule:

$$i_t = \gamma \left( \pi_t - \overline{\pi} \right) + \overline{\pi}. \tag{4}$$

There are two distinct treatments, A and B, to analyze the formation of confidence intervals. In treatment A we restrict ourselves to symmetric confidence intervals. Subjects insert the difference from their point forecast which is roughly equivalent to 1.96 standard errors of their expectation, assuming it is represented by a normal distribution. This is relaxed in treatment B, where subjects have to report the upper and the lower bound of their forecast together with their mean forecast, so that we do not require individuals to report symmetric confidence intervals (in both treatments we ask them to report 95% confidence intervals). As explained above, there are also four treatments which use different specifications of the monetary policy reaction function. The summary is provided in the Table 1.

The first three treatments, deal with the parametrization of the inflation forecast targeting given in equation (3). In this setup, the slope coefficient $\gamma$ determines the central bank's aggressiveness in response to deviations of inflation (or inflation expectations) from its target. Higher $\gamma$ implies stronger stabilizing effect of the Taylor-type monetary policy rule. The majority of empirical findings in the literature agree that the magnitude of the slope coefficient is around 1.5. Our key interest is in seeing how subjects react to more and less aggressive interest rate policies and how polices influence the uncertainty of their forecasts (see Section 3). Detailed discussion on the treatment selection regarding monetary policy can be found in our companion

| Subtreatments | | Treatment A Symmetric confidence interval | Treatment B Asymmetric confidence bounds |
|---|---|---|---|
| Taylor rule (equation) | Parameters | Groups | Groups |
| 1 – Forward looking (3) | $\gamma = 1.5$ | 1-4 | 5-6 |
| 2 – Forward looking (3) | $\gamma = 1.35$ | 7-10 | 11-12 |
| 3 – Forward looking (3) | $\gamma = 4$ | 13-16 | 17-18 |
| 4 – Contemporaneous (4) | $\gamma = 1.5$ | 19-22 | 23-24 |

Table 1: Treatments

paper (Pfajfar and Žakelj, 2011) where we focus on questions regarding the performance of different monetary policy rules and their impact on expectations (point forecasts).

**2.1. Experimental procedures.** Experimental subjects participate in a simulated economy of 9 agents.[13] Each session of a treatment has 2 independent groups ("economies"), therefore 18 subjects participate in each session. Participants are enlisted through a recruitment program for undergraduate students at the Universitat Pompeu Fabra and the University of Tilburg. The experiment consists of 12 sessions each containing 2 independent groups, thus 24 groups in total. Participants remain in the same group throughout the experiment. They earn on average around €15, depending on the treatment and individual performance. Participants receive detailed instructions (here attached in Appendix C), a quiz questionnaire and play 5 practice rounds before the start of the experiment to make sure they fully understand their task. The program is written in Z-Tree experimental software (Fischbacher, 2007).

Subjects are presented with a simple fictitious economy setup. As shown above, the economy is described with three macroeconomic variables: inflation, output gap and interest rate. Participants observe time series of these variables and their past forecasts, up to the period $t-1$. They do not observe the forecasts of other individuals and their performance. 10 initial values are generated by the computer under the assumption of rational expectations. The subjects' task is to provide inflation forecasts for the period $t+1$ and 95% confidence interval. The underlying model of the economy is qualitatively described to them. We explain the meaning and relevance between the main macroeconomic variables and inform them that their decisions have an impact on the realized output, inflation and interest rate in time $t$. This is a predominant strategy in the learning to forecast experiments (see Duffy, 2008, and Hommes, 2011).[14] Each

[13] The common view among the experimental economists is that we do not need many subjects in the micro-founded experiments. Most of the learning to forecast experiments are conducted with 5-6 subjects, e.g. Hommes, Sonnemans, Tuinstra, and van de Velden (2005), Adam (2007), Fehr and Tyran (2008).

[14] In learning to forecast experiments it is not possible to achieve REE (Rational Expectations Equilibrium) simply by introspection. This holds even if we provide subjects with the data generating process as there exists uncertainty how other participants forecast, so subjects have to engage in a number of trial and error exercises or in other words adaptive learning. It has been analytically proven in Marcet and Sargent (1989) and further formalized in a series of papers by Evans and Honkapohja (see their book: Evans and Honkapohja, 2001) that it is enough that agents observe all relevant variables in the economy (as in our case, where they are specifically instructed that all of them might be relevant) and update their forecasts according to the adaptive learning algorithm (their errors) they will end up in the REE. This has been acknowledged also in Duffy (2008) and Hommes (2011).

session consists of 70 periods.

After each period subjects receive information about realized inflation in that period, their prediction of it, and the payoff they have gained. The payoff function is a sum of two convex components. The first component depends on their forecast errors, while the second depends on the width of their confidence interval.

$$
\begin{aligned}
W &= W_1 + W_2, \\
W_1 &= \max\left\{\frac{1000}{1+f} - 200,\ 0\right\}, \\
W_2 &= \max\left\{\frac{1000x}{1+CI} - 200,\ 0\right\}, \\
x &= \begin{cases} 1 & \text{if} \quad CI \geq f \\ 0 & \text{if} \quad otherwise \end{cases}, \\
f &= \left|\pi_t - \pi^k_{t+1|t}\right|.
\end{aligned}
$$

The first component, $W_1$ is designed to encourage subjects to give accurate point forecasts. It depends on their forecast errors and is designed to encourage subjects to give accurate predictions. It gives subjects a payoff if their forecast errors, $f$, are smaller than 4. The second component, $W_2$, depends on the width of their confidence interval and intends to motivate subjects to think about the variance of actual inflation since it is more rewarding when it is narrower. It exhibits a trade-off between the width of this interval and its accuracy. A similar functional form of the payoff function is used in Adam (2007). $CI$ is either equal to their point estimate of confidence interval or half of the difference between upper and lower bound. Subjects receive a reward if their confidence intervals, $CI$, are not larger than $\pm 4$ percentage points, conditional on the fact that actual inflation falls in the given interval: $CI \geq \left|\pi_t - \pi^k_{t+1|t}\right|$. With this setup we restrict to positive payoffs.

We performed several simulations regarding the incentive compatibility of the part of the payoff function that addresses confidence bounds. Desirable payoff functions have to exhibit a trade-off between the width of the interval and the accuracy of the interval, which makes it difficult to specify and calibrate incentive compatible payoff function. Assuming all agents are rational (and they know that all others are rational) then the function chosen gives a maximum payoff when 96.5% confidence intervals are taken into account. When not all subjects are rational there are two effects on their confidence intervals: (i) forecast accuracy decreases and the required 95% confidence interval widens; (ii) the payoff function is maximized with narrower confidence intervals than 96.5%.[15] Maximizing the objective function under nonrational agents requires several assumptions regarding the perceived law of motion of both point forecasts and confidence intervals since optimal confidence intervals are not necessarily constant as in the case of rational agents. Therefore the only natural benchmark is rational expectations and we

[15] This is also supported with previous evidence in the literature. Pfajfar and Žakelj (2011) show that in this experiment nonrational forecast results in more variability of inflation. Du and Budescu (2007) demonstrate that higher variability of the underlying series is associated with greater overconfidence (narrower intervals).

decided to formulate the question in terms of 95% confidence intervals.

## 3. Individual Uncertainty

While the distribution of means across subjects captures only interpersonal variation, individual confidence bounds help us to approximate individual uncertainty of future inflation. Zarnowitz and Lambros (1987) claim that interpersonal variation is an appropriate measure of disagreement among forecasters while uncertainty can be described as intrapersonal variation. Their study shows that there can be substantial differences between the variation of disagreement and variation of uncertainty, therefore both might not be appropriate measures for forecasting variability of inflation. Our experimental design allows us to analyze both features of the distribution of responses. The current section concentrates on individual uncertainty, while the next section investigates aggregate distribution of forecasts and disagreement.

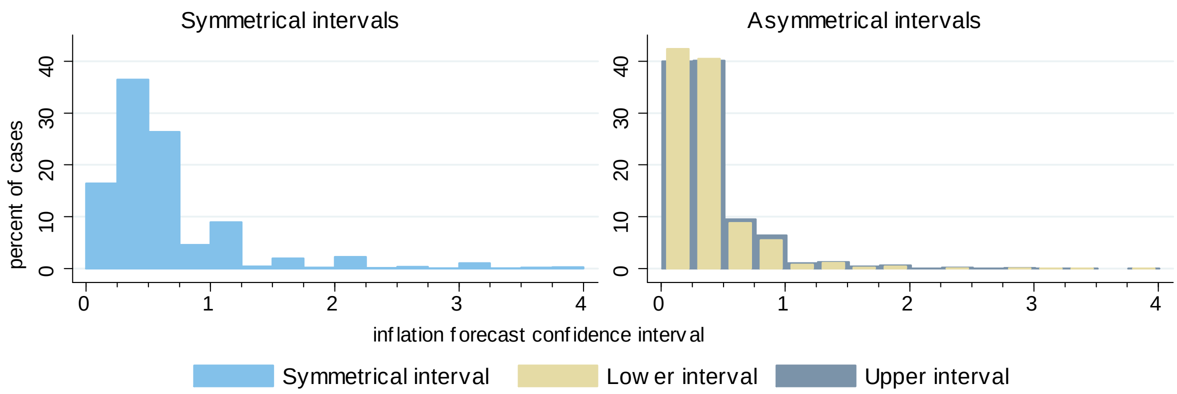


Figure 1: Histogram of confidence intervals for all treatments, subjects and periods

Figure 1 displays the distribution of all confidence interval forecasts. The range of responses for confidence intervals is between 0 and 8.3, although it should be noted that responses larger than 4 do not result in any payoff.[16] The average symmetrical confidence interval is 0.61 with an average standard deviation of 0.28. Introducing asymmetrical confidence bounds across all treatments gives us an average lower confidence interval of 0.37 with an average standard deviation of 0.19 while the average upper confidence interval equals 0.41 with an average standard deviation of 0.28. There are considerable differences across treatments as the lowest symmetrical (asymmetrical lower, upper) average interval in treatment A (treatment B) is 0.41 (0.24, 0.27) and the highest is 0.91 (0.47, 0.53). Evidence of rounding is present in responses 0.5, 1, 1.5, 2, and 3 as they have significantly higher frequencies than other responses. Overall, 13% of responses are integers, while the majority are to one decimal point accuracy, 77%. The remaining responses are to 2 decimal point accuracy. Rounding of the inputs for confidence intervals (probabilistic forecasts) has been previously documented by D'Amico and Orphanides (2008) and Engelberg, Manski, and Williams (2009).

[16] The overall share of responses greater than 4 is 0.98%.

| Average confidence interval | All | Treat. A (symmetric) | Treat. B (asymmetric) |
|---|---|---|---|
| 1 – Forward looking (3), $\gamma = 1.5$ | 0.564 | 0.669 | 0.352 |
| 2 – Forward looking (3), $\gamma = 1.35$ | 0.776 | 0.914 | 0.500 |
| 3 – Forward looking (3), $\gamma = 4$ | 0.395 | 0.466 | 0.254 |
| 4 – Contemporaneous (4), $\gamma = 1.5$ | 0.430 | 0.410 | 0.471 |

Table 2: Width of confidence intervals across treatments. Note: The width of asymmetric confidence interval is calculated as (Upper b. - Lower b.)/2.

The average confidence intervals in each treatment are listed in Table 2, while per-group summary is presented in Table 3. In general, confidence intervals are narrower in treatment B than in treatment A at 1% significance using nonparametric tests (Wilcoxon/ Mann-Whitney). In Section 3.1 we show that treatments A and B also differ in forecast accuracy of subjects' interval predictions. The factors that determine the differences in confidence intervals are discussed in Section 3.2.

We also have the opportunity to compare the results with the underlying uncertainty that we have embedded in our set-up. Under the assumption that all agents use rational expectations in all periods, a rational agent would set her confidence interval of 0.2046 for treatments 1-3 and 0.2081 for treatment 4.[17] Of course, as soon as one subject departs from rationality, the confidence interval of a rational agent should immediately become larger as she has to account for the uncertainty of other subjects' expectations. Under rational expectations in treatments 1-3 the uncertainty should not be affected by the $\gamma$, while in treatment 4 it depends on $\gamma$; higher $\gamma$ leads to lower uncertainty.

As outlined above, uncertainty should be slightly lower when the central bank is pursuing inflation forecast targeting compared to inflation targeting. Contrary to that, we find that the average confidence interval is narrower in treatment 4 compared to the other treatments. The average confidence interval is indeed narrower in treatment 4 compared to other treatments. This difference is statistically significant with standard parametric (t-test) and nonparametric tests (Wilcoxon/Mann-Whitney). However, if we compare treatment 4 separately to all other treatments, we can observe that while it is significantly narrower than treatments 1 and 2 it is wider than treatment 3. The theory suggests that in treatments 1-3 all confidence intervals should have the same width; however this is strongly rejected with our experimental data. We can conclude that the monetary policy significantly affects the width of the confidence interval. Inflation targeting results in a narrower confidence interval than inflation forecast targeting. Furthermore, in the case of inflation forecast targeting, the width of the confidence interval also depends on how strongly the monetary policy is reacting to deviations of inflation from its

[17] Unconditional variances of the residuals following AR(1) process are $vr_g = \frac{\sigma_g^2}{1-\kappa^2}$ and $vr_u = \frac{\sigma_u^2}{1-\nu^2}$. Associated confidence interval for treatments 1-3 is therefore $1.96 \cdot \sqrt{vr_g + \lambda^2 vr_u}$. For treatment 4 the value is $1.96 \cdot \sqrt{\left(\frac{1}{\lambda\gamma\varphi+1}\right)^2 vr_g + \left(\frac{\lambda}{\lambda\gamma\varphi+1}\right)^2 vr_u} = 0.2081$.

target.[18]

| | | Inflation | | Confidence bound | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Symmetric | | Lower | | Upper | |
| Treat. | Group | mean | stdev | mean | stdev | mean | stdev | mean | stdev |
| 1-A | 1 | 2.85 | 5.87 | 0.97 | 0.71 | - | - | - | - |
| 1-A | 2 | 2.88 | 2.91 | 0.65 | 0.40 | - | - | - | - |
| 1-A | 3 | 2.92 | 1.97 | 0.70 | 0.35 | - | - | - | - |
| 1-A | 4 | 3.00 | 0.76 | 0.34 | 0.16 | - | - | - | - |
| 1-B | 5 | 3.13 | 1.10 | - | - | 0.36 | 0.19 | 0.41 | 0.24 |
| 1-B | 6 | 3.12 | 0.90 | - | - | 0.29 | 0.14 | 0.35 | 0.41 |
| 2-A | 7 | 3.12 | 0.76 | 1.09 | 0.30 | - | - | - | - |
| 2-A | 8 | 3.09 | 1.82 | 1.15 | 0.63 | - | - | - | - |
| 2-A | 9 | 3.13 | 0.51 | 0.38 | 0.21 | - | - | - | - |
| 2-A | 10 | 3.02 | 5.53 | 1.02 | 0.56 | - | - | - | - |
| 2-B | 11 | 2.52 | 3.58 | - | - | 0.61 | 0.45 | 0.72 | 0.43 |
| 2-B | 12 | 3.03 | 0.88 | - | - | 0.33 | 0.12 | 0.33 | 0.14 |
| 3-A | 13 | 3.01 | 0.52 | 0.53 | 0.13 | - | - | - | - |
| 3-A | 14 | 3.02 | 0.94 | 0.65 | 0.32 | - | - | - | - |
| 3-A | 15 | 2.99 | 0.24 | 0.35 | 0.09 | - | - | - | - |
| 3-A | 16 | 3.00 | 0.26 | 0.33 | 0.10 | - | - | - | - |
| 3-B | 17 | 2.99 | 0.31 | - | - | 0.28 | 0.09 | 0.28 | 0.10 |
| 3-B | 18 | 3.01 | 0.24 | - | - | 0.20 | 0.08 | 0.25 | 0.35 |
| 4-A | 19 | 3.09 | 0.39 | 0.36 | 0.13 | - | - | - | - |
| 4-A | 20 | 3.23 | 0.81 | 0.56 | 0.20 | - | - | - | - |
| 4-A | 21 | 3.05 | 0.48 | 0.38 | 0.09 | - | - | - | - |
| 4-A | 22 | 3.05 | 0.38 | 0.34 | 0.10 | - | - | - | - |
| 4-B | 23 | 3.09 | 0.52 | - | - | 0.31 | 0.12 | 0.31 | 0.15 |
| 4-B | 24 | 3.11 | 1.29 | - | - | 0.60 | 0.28 | 0.65 | 0.37 |
| All-A | | 3.03 | 1.51 | 0.61 | 0.28 | - | - | - | - |
| All-B | | 3.00 | 1.10 | - | - | 0.37 | 0.18 | 0.41 | 0.28 |

Table 3: Confidence bounds, summary statistics.

Our results might not be directly comparable to those based on surveys. Probabilistic forecasts in surveys are usually collected in terms of histograms where intervals are predefined and fixed for all participants. Another difference between our experiment and surveys concerns the risk attitude. With professional forecasters it can be claimed that their probability and point forecasts are correlated because they interact and influence each other.[19] Zarnowitz and Lambros (1987) argue that risk averse forecasters tend to make their forecasts as close to the relevant value as possible, and this holds for point forecasts and probabilistic forecasts. In our experiment, subjects could neither exchange information about each other's expectations, nor is the average aggregate prediction directly observable.

[18] This relationship is further analyzed in Section 3.2.

[19] Scharfstein and Stein (1990), Banerjee (1992) and Zwiebel (1995) argue that forecasters are occasionally afraid to deviate from the majority or the consensus opinion. Pons-Novell (2003) provides empirical evidence on this.

**3.1. Forecasting Accuracy.** In this section we first establish some stylized facts about forecasting performance and then we focus on establishing which factors affect the probability that the actual inflation falls within the specified bounds.

It is interesting to see how accurate experimental subjects are in determining the confidence bounds. Thaler (2000) suggests that when people are asked "for their 90% confidence limits ... the correct answers will lie within the limits less than 70% of the time" (p. 133). Giordani and Söderlind (2003) get a very similar result (72%).[20] Our results confirm the overconfidence effect in an even stronger manner than survey data results. Only 60.5% of the times subjects manage to set confidence bounds that include actual inflation in the next period.[21] This proportion is higher in treatment A where 64.3% correctly specify confidence intervals while in treatment B the proportion is only 52.8%. It is interesting to note that the actual inflation is lower than their confidence intervals in 19% of cases while it is higher in 20.5%. If we compare this among treatments we find that in treatment A (B) actual inflation is lower than their confidence intervals in 17.1% (22.9%) of cases while it is higher in 18.5% (24.4%). As we mentioned in the introduction this overconfidence effect has attracted a lot of attention in psychology literature. Some studies even document that the success rate of these forecasts is less than 50% when people are asked for $90-99\%$ confidence intervals (e.g. Lichtenstein, Fischhoff, and Phillips, 1982).[22] The most striking example of this bias has been recently documented by Ben-David, Graham, and Harvey (2010) who assembled a panel of forecasts by top financial executives. They show that the realized market returns are only 33% of the time within 80% confidence bounds. They put forward two possible explanations for these results: (i) CEOs overestimate their ability to predict the future or (ii) they underestimate the volatility of random events. Moreover, Biais, Hilton, Mazurier, and Pouget (2005) argue that traders who underestimate risk are prone to the winner's curse.[23]

The accuracy of confidence intervals differs also across different monetary policies. We find that in treatment 3 and 4, subjects are more accurate (62.9% and 69.4% accuracy respectively) than in the benchmark treatment 1 (51.7% accuracy). Differences are significantly at a 10% level with Wilcoxon/ Mann-Whitney test.

As confidence intervals forecast the distribution of the expected forecast errors we can actually dig deeper and analyze each individual separately. We find that only 11.1% of the subjects on average overestimate risk in treatment A and 2.8% (1.4%) of the subjects in treatment B for lower (upper) bound. Closer inspection allows us to conclude that only about 9.0% of the subjects in treatment A and 1.4% (8.4%) of the subjects in treatment B for lower (upper) bound on average report the confidence bounds that are not significantly different than 95% confidence intervals based on actual forecast errors. The rest of subjects on average forecast confidence bounds that are significantly lower than the actual forecast errors. Per-group statistics are

[20] See also Giordani and Söderlind (2006).

[21] Moreover, our instructions required subjects to introduce their prediction with 95 confidence bounds.

[22] Onkal and Bolger (2004) and Du and Budescu (2007) document that the overconfidece effect weakens when subjects are asked for 70 or 50% confidence intervals.

[23] Yaniv and Foster (1995) argue that overconfidence can be explained by the fact that the subjects are worried that inserting too narrow confidence intervals will reduce the informativeness of their inputs.

reported in Table A1 in Appendix A.

We check how volatility of inflation, the width of confidence bounds, and macroeconomic variables affect the likelihood of inflation falling within the specified confidence bound.[24] We estimate the following regression:

$$\begin{aligned} x_t^k &= \alpha^k + \beta sip_{t|t-1}^k + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} + \\ &+ \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \zeta i_{t-1} + \delta sd_{t-1}^j + u_t^{em}, \end{aligned} \quad (5)$$

where $x_t^k$ takes the value 1 when inflation falls within the provided bounds and 0 otherwise, $sip_{t|t-1}^k$ is subject $k$'s interval prediction for period $t$ (for treatment B it is (Upper b. - Lower b.)/2), $y_t$ is output gap, $\pi_t$ is actual inflation and $i_t$ is the interest rate. $D_1, \ldots, D_3$ are dummy variables. $D_1$ equals 1 when $y_{t-1} > 0.1$ and $\Delta y_{t-1} > 0$ and is 0 otherwise; $D_2$ equals 1 when $y_{t-1} < 0.1$ and $\Delta y_{t-1} < 0$ and is 0 otherwise; while $D_3$ equals 1 when $D_1 = 0$ and $D_2 = 0$ jointly and is 0 otherwise. $sd_{t-1}^j$ is standard deviation of inflation up to period $t-1$ for group $j$. $D_L$ equals 1 when inflation is below the target and 0 otherwise, while $D_H$ equals 1 when inflation is above its target and 0 otherwise.

| $x_t^k$ : | *all* | *treat.A* | *treat.B* |
|---|---|---|---|
| $sip_{t\|t-1}^k$ | 2.3985*** | 2.3578*** | 2.9678*** |
| | (0.2340) | (0.3620) | (0.4567) |
| $D_1 y_{t-1}$ | -0.8720*** | -1.1590*** | -0.7103*** |
| | (0.2117) | (0.4328) | (0.2137) |
| $D_2 y_{t-1}$ | 1.3565*** | 1.9346*** | 1.4602*** |
| | (0.2304) | (0.5309) | (0.2439) |
| $D_3 y_{t-1}$ | 0.3092* | 0.3000 | 0.2717 |
| | (0.1684) | (0.3153) | (0.2023) |
| $D_L\|\pi_{t-1}\|$ | 0.2179** | 0.0933 | 0.3218* |
| | (0.0948) | (0.5938) | (0.1856) |
| $D_H\|\pi_{t-1}\|$ | 0.5955*** | 1.2236** | 0.5659*** |
| | (0.1344) | (0.4821) | (0.1497) |
| $i_{t-1}$ | -0.1529** | -0.3655 | -0.0960 |
| | (0.0758) | (0.3859) | (0.0817) |
| $sd_{t-1}^j$ | -1.4642*** | -0.8690* | -1.8730*** |
| | (0.2722) | (0.4525) | (0.4803) |
| $N$ | 14628 | 4968 | 9660 |
| Wald $\chi^2_{(8)}$ | 168.4 | 230.0 | 122.9 |

Table 4: Forecasiting accuracy and confidence intervals. Note: Coefficients are based on fixed effects logit estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

The results for fixed effects logit estimation are reported in Table 4 while for Poisson fixed

[24] Frequencies of forecast errors depending on inflation cycle can be found in the Table A2 in Appendix A.

effects and random effects are reported in Tables B1-B3 in Appendix B. As one would expect, when there is higher volatility of inflation there are more outcomes outside the interval, especially in treatment B. This is well documented in psychology literature as greater volatility leads to overconfidence (e.g. Lawrence and Makridakis, 1989, Lawrence and O'Connor, 1992),[25] however some studies find also that there is no such effect (Du and Budescu, 2007). In both treatments wider confidence intervals result in a higher probability of correctly specifying the confidence interval. Interestingly, we can observe that there exists some pattern across business cycles. There are more outcomes outside the interval, when the output gap is positive and has a clear upward trend of inflation, while in the opposite situation there is lower probability of misperceiving inflation uncertainty. Inflation also has a significant positive impact on the likelihood of the forecast falling within the interval, especially when inflation is above the target value.[26]

**3.2. Determinants of Individual Uncertainty.** Below we analyze the determinants of confidence bounds using panel data. All regressions below are estimated using system GMM estimator of Blundell and Bond (1998) for dynamic panel data. They are replicated for the whole sample (*all*), treatment A (*treat.A*), and separately for part of the interval below the point forecast ($treat.B-L$) and above the point forecast ($treat.B-U$) in treatment B. In order to transform the asymmetric confidence intervals into a measure comparable to the symmetric ones, we compute the average of the upper and lower interval.

We begin by detailing the relationship between confidence interval and standard deviation of inflation. We estimate the following regression:

$$sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma sd^j_{t-1} + u^{em}_t, \tag{6}$$

where the individual $k$'s current perceived uncertainty in period $t$, is measured by her confidence interval, $sip^k_{t+1|t}$. Results are reported in Table 5.

We find that confidence intervals are highly inertial. This has been previously documented in Bruine de Bruin, Manski, Topa, and van der Klaauw (2011), and Giordani and Söderlind (2003). Higher standard deviation of inflation leads to wider confidence intervals, however with a smaller effect in treatment A. Du and Budescu (2007) find no relationship between these variables. Positive correlation between the self-reported range of responses and underlying uncertainty is also found for survey data in Bruine de Bruin, Manski, Topa, and van der Klaauw (2011).[27]

A second feature of the confidence intervals that we want to study is the subjects' responses to inflation falling outside the confidence interval. To discriminate between the effects

[25] Psychologists argue that this overconfidence is due to hard-easy effects, i.e. miscalibration (reported narrower confidence intervals) is higher in hard tasks and attenuated or even eliminated in easy tasks (e.g. Keren, 1991).

[26] In Table B7 in Appendix B we also report the results of the relationship between the individual $k$'s forecast error $r^k_{t+1} = \pi^k_{t+1|t} - \pi_{t+1}$, and the confidence interval as a measure of uncertainty.

[27] We also study the relationship between confidence intervals and inflation forecasts as in Rich and Tracy (2010). They find mixed evidence of the existence of this relationship, while we do not find any evidence in favor of this relationship. These results are available upon request from the authors.

| $sip^k_{t+1\|t}$ : | $all$ | $treat.A$ | $treat.B-L$ | $treat.B-U$ |
|---|---|---|---|---|
| $sip^k_{t\|t-1}$ | 0.4390*** | 0.5445*** | 0.4407*** | 0.0925 |
| | (0.1114) | (0.0921) | (0.0485) | (0.0982) |
| $sd^j_{t-1}$ | 0.1167*** | 0.0955** | 0.1357*** | 0.2643*** |
| | (0.0450) | (0.0401) | (0.0220) | (0.0561) |
| $\alpha$ | 0.2143*** | 0.2039*** | 0.1142*** | 0.1884*** |
| | (0.0283) | (0.0285) | (0.0187) | (0.0323) |
| $N$ | 14904 | 9936 | 4968 | 4968 |
| Wald $\chi^2_{(3)}$ | 140.9 | 259.1 | 346.1 | 34.6 |

Table 5: Confidence intervals and standard deviation of inflation. Note: Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

of overshooting and undershooting we introduce two dummy variables. $D^k_4$ takes the value 1 if $\left(\left|r^k_{t-1}\right| > sip^k_{t-1}\right) \wedge \left(r^k_{t-1} \geq 0\right)$, and 0 otherwise. Note that $r^k_{t-1} = \pi_{t-1} - \pi^k_{t-1|t-2}$ is subject $k$'s last observed forecast error. $D^k_5$ equals 1 if $\left(\left|r^k_{t-1}\right| > sip^k_{t-1}\right) \wedge \left(r^k_{t-1} \leq 0\right)$, and 0 otherwise, while $D^k_6$ is 1 when $\left|r^k_{t-1}\right| < sip^k_{t-1}$, and 0 otherwise. Therefore $D^k_4 = 1$ when subject $k$ is underestimating inflation; while $D^k_5 = 1$ when subject $k$ is overestimating inflation. We run the following regression:

$$sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma D^k_4 r^k_{t-1} + \delta D^k_5 r^k_{t-1} + \epsilon D^k_6 r^k_{t-1} + u^{em}_t. \quad (7)$$

| $sip^k_{t+1\|t}$ : | $all$ | $treat.A$ | $treat.B-L$ | $treat.B-U$ |
|---|---|---|---|---|
| $sip^k_{t\|t-1}$ | 0.4430*** | 0.5496*** | 0.4641*** | 0.1068 |
| | (0.1080) | (0.0865) | (0.0491) | (0.1059) |
| $D_4 r^k_{t-1}$ | 0.0363** | 0.0292** | 0.0023 | 0.0669* |
| | (0.0153) | (0.0147) | (0.0228) | (0.0343) |
| $D_5 r^k_{t-1}$ | -0.0760*** | -0.0647*** | -0.0955*** | -0.0668** |
| | (0.0193) | (0.0190) | (0.0094) | (0.0269) |
| $D_6 r^k_{t-1}$ | 0.0015 | 0.0025 | -0.0191* | 0.0506 |
| | (0.0201) | (0.0204) | (0.0107) | (0.0309) |
| $\alpha$ | 0.2799*** | 0.2568*** | 0.1882*** | 0.3504*** |
| | (0.0396) | (0.0416) | (0.0251) | (0.0406) |
| $N$ | 14688 | 9792 | 4896 | 4896 |
| Wald $\chi^2_{(5)}$ | 203.6 | 248.5 | 1048.1 | 19.7 |

Table 6: Confidence intervals and phases of economic cycle. Note: $treat.B-L$ ($treat.B-U$) only includes part of the interval beneath (above) the point forecast. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

Results shown in Table 6 suggest that subjects increased their confidence intervals after the last observed inflation was outside the interval.[28] This holds for both "undershooting" and "overshooting." In the latter $r^k_{t-1}$ is negative, so negative coefficient $\delta$ implies that confidence intervals are widened after $\left|r^k_{t-1}\right| > sip^k_{t-1}$. Positive or negative errors do not result in any significant change of confidence intervals in the next period when inflation falls within the interval. It is also interesting to note that confidence intervals in the treatment B exhibit less inertia, especially the upper bound, compared to the treatment A. Moreover, the interval above point forecast widens with both overshooting and undershooting while the interval below is more stable and responds only to undershooting. This also represents the first potential source of observed asymmetries. Ben-David, Graham, and Harvey (2010) also note that there is a difference regarding the formation of the upper and the lower bound of confidence intervals. They argue that lower forecast bounds are significantly affected by the past return while upper are not.

Several studies have established that there are significant variations in uncertainty over the business cycles; in particular, uncertainty is found to be countercyclical. Bloom (2009) and Bloom, Floetotto, and Jaimovich (2010) build theoretical models where uncertainty shocks exert a key role in the business cycle fluctuations. We estimate equation (8), where we control for the path of output gap. In addition, specification (8) also allows for the possibility that subjects change their interval forecasts on the basis of their last point forecast errors:

$$\begin{aligned} sip^k_{t+1|t} &= \alpha + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} \\ &\quad + \zeta i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi \left|r^k_{t-1}\right| + \vartheta T2 + \iota T3 + \kappa T4 + u^{em}_t, \end{aligned} \tag{8}$$

where $y_t$ is output gap, $i_t$ is the interest rate, and $D_1, \ldots, D_3$ are dummy variables as identified in equation (14). Estimation results are in the Table 7. $T2$, $T3$ and $T4$ are treatment dummies.[29]

Friedman (1968) points that there is a positive link between inflation and inflation uncertainty. While Liu and Lahiri (2006) and D'Amico and Orphanides (2008) find empirical support for this conjecture, we cannot confirm it in our experiment. Regressing equation (8) with inflation ($\pi_{t-1}$) instead of $D_L \left|\pi_{t-1}\right|$ and $D_H \left|\pi_{t-1}\right|$ would result in inflation having a negative impact on the width of confidence interval. The empirical studies that find a positive correlation between inflation and uncertainty are based on the US economy where, especially in the 70s, there was mostly an upward risk for inflation. In our experiment, inflation fluctuates around the inflation target, so also decreases of inflation below the inflation target increase uncertainty. With specification (8) we concentrate on the absolute deviations of inflation from inflation target, while controlling for high and low inflation levels. We indeed observe that downside risk has an even more important impact on the uncertainty than the upside risk. Moreover, being above the target inflation only the upper part of the confidence interval will be widened, whereas being

[28] Table B6 in Appendix B reports regression with dummies without interaction with actual forecast errors.

[29] Treatment dummies are included only into regression *all* as in other specifications due to too few observations within one treatment we would have to abolish clustering of standard errors if we were to include treatment dummies.

| $sip^k_{t+1\|t}$ : | *all* | *treat.A* | *treat.B − L* | *treat.B − U* |
|---|---|---|---|---|
| $sip^k_{t\|t-1}$ | 0.3976*** | 0.5333*** | 0.4305*** | 0.0900 |
| | (0.1034) | (0.0990) | (0.0398) | (0.0997) |
| $D_1 y_{t-1}$ | 0.0067 | 0.0198 | 0.0202 | -0.0560 |
| | (0.0219) | (0.0258) | (0.0252) | (0.0465) |
| $D_2 y_{t-1}$ | -0.0188 | -0.0118 | -0.0144 | -0.0650** |
| | (0.0225) | (0.0217) | (0.0304) | (0.0262) |
| $D_3 y_{t-1}$ | 0.0067 | 0.0183 | 0.0051 | -0.1142*** |
| | (0.0296) | (0.0271) | (0.0183) | (0.0413) |
| $i_{t-1}$ | 0.0110 | 0.0070 | -0.0066 | 0.0025 |
| | (0.0076) | (0.0073) | (0.0059) | (0.0157) |
| $D_L\|\pi_{t-1}\|$ | 0.0294** | 0.0241** | 0.0247** | 0.0782*** |
| | (0.0115) | (0.0105) | (0.0108) | (0.0234) |
| $D_H\|\pi_{t-1}\|$ | 0.0180 | 0.0173 | 0.0668*** | 0.0248 |
| | (0.0167) | (0.0135) | (0.0139) | (0.0310) |
| $\|r^k_{t-1}\|$ | 0.0552*** | 0.0473*** | 0.0474** | 0.0749*** |
| | (0.0154) | (0.0159) | (0.0203) | (0.0250) |
| $T2$ | 1.0505* | | | |
| | (0.5519) | | | |
| $T3$ | -0.6098 | | | |
| | (0.5743) | | | |
| $T4$ | -0.6351 | | | |
| | (0.5913) | | | |
| $\alpha$ | 0.2790 | 0.2062*** | 0.1694*** | 0.2898*** |
| | (0.2893) | (0.0415) | (0.0266) | (0.0541) |
| $N$ | 14688 | 9792 | 4896 | 4896 |
| Wald $\chi^2_{(12)}$ | 393.8 | 715.8 | 865.0 | 145.0 |

Table 7: Confidence intervals and macroeconomic variables. Note: *treat.B − L* (*treat.B − U*) only includes part of the interval beneath (above) the point forecast. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

below the target inflation both sides of the confidence interval will be widened.

Interest rates are positively related to the individual confidence intervals in the regressions above, however their effects are not significant. Zarnowitz and Lambros (1987) point out that uncertainty about inflation and interest rates can be either positively or negatively related in the field, however for their sample they find a negative relationship. Giordani and Söderlind (2003) additionally argue that the forecast uncertainty is positively related to the forecast errors. In Table 7 we also demonstrate that confidence intervals depend on the last observed absolute forecast error.[30]

[30] Bruine de Bruin, Manski, Topa, and van der Klaauw (2011) also point out that individual uncertainty is positively related to point forecasts.

Above regressions confirm the asymmetries between the upper and lower confidence bound demonstrated in Table 2. We can argue that the upper bound is more sensitive to the stage of the business cycle than the lower bound. In addition, different monetary policy rules also have an effect on the width of the confidence interval. The confidence intervals are wider for example in treatment 2 compared to the other treatments. One reason behind this is that uncertainty is related to the variability of inflation which in turn depends on $\gamma$ and more generally on the monetary policy (see Pfajfar and Žakelj, 2011).[31] However, there also exist other treatment effects as can be observed when controlling for a standard deviation of inflation. Table B5 in Appendix B demonstrates that if we include treatment dummies to regression (6) we find that dummy variable for treatment 2 is significant.

**3.3. Which factors affect the choice of (a)symmetric confidence intervals?.** So far we have found several asymmetries between the formation of the upper and the lower confidence bounds. In this section we analyze the choice of asymmetric confidence interval by using data from treatment B. Let us first analyze a proportion of subjects that systematically chose either a wider interval above the point forecast as compared to the one below point forecast or vice-versa. It is clear from Table 8 that when subjects are given an option to insert an asymmetric confidence interval they often do so, especially in treatments 1 and 2. Moreover, among more than 40% of subjects who systematically insert asymmetric intervals, there are fewer than 6% that perceive higher uncertainty on the left hand side of their point forecast. We can also observe that the proportion of subjects inserting symmetric intervals is the highest in treatment 4. Table 8 shows that the behavior of subjects in the inflation targeting treatment is more in line with theory than in treatments with inflation forecast targeting.

| Lower vs. upper (% of subjects) | $C_L < C_U$ | $C_L \approx C_U$ | $C_L > C_U$ |
|---|---|---|---|
| 1 – Forward looking (3), $\gamma = 1.5$ | 44.4 | 50.0 | 5.6 |
| 2 – Forward looking (3), $\gamma = 1.35$ | 50.0 | 44.4 | 5.6 |
| 3 – Forward looking (3), $\gamma = 4$ | 33.3 | 66.7 | 0.0 |
| 4 – Contemporaneous (4), $\gamma = 1.5$ | 16.7 | 72.2 | 11.1 |
| All | 36.1 | 58.3 | 5.6 |

Table 8: Proportions of subjects from Treatment B, depending on the difference between their upper ($C_U$) and lower ($C_L$) confidence intervals. When $C_L < C_U$, subject inserted on average a smaller lower interval than upper interval. Based on pairwise t-test with 5% signifficance level.

Now we turn our attention to the factors that determine the probability of the asymmetric interval. We first define $D_7 = 1$ if the upper interval has exactly the same width as the lower one and 0 otherwise. There are only about 12.5% of these cases. We observe, however, that 84% of the subjects inserted their responses with one or two decimal points accuracy. It is therefore reasonable to define symmetry also as $|C_L - C_U| \leq 0.1$; in this case we set $D_8 = 1$.[32] According

[31] Due to the presence of heterogeneous expectations this relationship is not monotonic. It is found that the relationship between $\gamma$ and variability of inflation is U-shaped (see Pfajfar and Žakelj, 2011 for further details).

[32] Alternatively, we also tried $D_9 = 1$ if $0.9 \leq \left| \frac{ConfIntH_{n-1}}{ConfIntL_{n-1}} \right| \leq 1.1$. Results can be found Tables B7-9 in

to this definition 47.2% of our responses in treatment B are approximately symmetric. We estimate the following regressions:

$$\begin{aligned} D_z &= \alpha + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} + \zeta i_{t-1} \\ &\quad + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi sd^j_{t-1} + u^{em}_t; \quad z \in \{7, 8\}. \end{aligned} \tag{9}$$

Results for logit fixed effects estimator are reported in the first two columns of Table 9 while Poisson and logit random effects estimations can be found in Tables B8 and B9 in Appendix B. While the above regressions inform us about the likelihood that subjects would insert symmetric intervals, they are not suitable for measuring the magnitude of the asymmetry of the individual forecast distributions and its direction. For that purpose it is convenient to introduce a new variable, skewness, similar to that used in Du and Budescu (2007). We define the skewness variable, $skw^k_t$ by subtracting the point forecast from the midpoint of the confidence interval. If $skw^k_t$ is smaller (greater) than 0, then the interval is left (right) skewed, and the confidence interval below point forecast is wider (narrower) than the one above. If $skw^k_t = 0$ then the interval is symmetric. Factors affecting skewness are analyzed on the right side of Table 9 using the Blundell-Bond system GMM estimator.

$$\begin{aligned} skw^k_t &= \alpha + \eta skw^k_{t-1} + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} \\ &\quad + \zeta i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi sd^j_{t-1} + u^{em}_t. \end{aligned} \tag{10}$$

Regressions for $D_7$ and $D_8$ demonstrate that some indicators of cycle are significant. In particular for $D_7$ when the output gap is negative and downward sloping it is less likely to observe symmetric intervals while for $D_8$ it is more likely to observe symmetrical intervals in the opposite stages of the business cycle. For both regressions interest rate exerts a significantly positive impact and absolute inflation above the target a significantly negative impact, i.e. there is less symmetry when inflation is low.

Skewness measure, on the other hand, gives us also an indication of direction of the asymmetry. We find that this measure is inertial and tends to decrease (left skewness) when the previous confidence interval was larger. The measure also varies across the business cycles; it is lower when $D_3 = 1$. Du and Budescu (2007) find a negative relationship between standard deviation of inflation and skewness of confidence distribution, while we find this relationship only for the case of $D_7$.

## 4. Disagreement and aggregate expectation distribution

Different measures can be used to proxy inflation variability. We first analyze the features of the standard deviation of point forecasts. Second we take account of individual uncertainty as well. We define probability density functions of individual distributions, add them up and analyze the features of aggregate distribution.

Variance of a point forecasts is a "natural" measure of disagreement. It is often used in

Appendix B.

| | Symmetry | | Skewness |
|---|---|---|---|
| | $D_7$ | $D_8$ | $skw_t^k$ |
| $skw_{t-1}^k$ | - | - | 0.2861*** |
| | | | (0.0576) |
| $sip_{t\|t-1}^k$ | 0.2498 | -0.7167 | -0.2375*** |
| | (0.1643) | (0.5136) | (0.0852) |
| $D_1 y_{t-1}$ | 0.3345 | 0.4867** | -0.0496 |
| | (0.4229) | (0.2048) | (0.0415) |
| $D_2 y_{t-1}$ | -0.4418** | 0.1259 | -0.0447 |
| | (0.2093) | (0.2308) | (0.0337) |
| $D_3 y_{t-1}$ | -0.3388 | 0.1152*** | -0.0776*** |
| | (0.2504) | (0.0420) | (0.0240) |
| $i_{t-1}$ | 0.2547* | 0.2111*** | 0.0004 |
| | (0.1306) | (0.0684) | (0.0150) |
| $D_L\|\pi_{t-1}\|$ | 0.1828 | 0.1802 | 0.0273 |
| | (0.2757) | (0.1174) | (0.0275) |
| $D_H\|\pi_{t-1}\|$ | -0.4488* | -0.2613** | -0.0126 |
| | (0.2550) | (0.1177) | (0.0232) |
| $sd_{t-1}^k$ | -0.3237** | -0.5066 | -0.0050 |
| | (0.1510) | (0.3272) | (0.0498) |
| $\alpha$ | - | - | 0.1037** |
| | | | (0.0519) |
| $N$ | 4968 | 4968 | 4968 |
| Wald $\chi^2_{(8,9)}$ | 79.3 | 58.3 | 156.3 |

Table 9: Determinants of symmetric and skewed intervals. Note: Coefficients for the symmetry tests are based on fixed effects logit estimations, while coefficients for skewness are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

the empirical literature since the data on point forecasts are more frequently available than the data on individual distributions. It is studied for example in Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003). We investigate the relation of standard deviation of point forecasts to the phases of economic cycle, interest rate, inflation and the mean forecast error:

$$\begin{aligned} sdv_{t+1|t}^j &= \alpha + \beta sdv_{t|t-1}^j + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} \\ &\quad + \zeta i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi mr_{t-1}^j + u_t^{em}, \end{aligned} \tag{11}$$

where $sdv_{t+1|t}^j$ is a cross-sectional standard deviation of point forecasts in a group $j$ at period $t$, while mean absolute forecast error in a group $j$ at period $t-1$ is $mr_{t-1}^j$.

Regressions based on (11) are displayed on the left side of Table 10. Standard deviation of point forecasts exhibits sensitivity to inflation, mean absolute forecast error and to some degree business cycles. However it tends to be less sensitive to these variables in the treatment with asymmetric confidence intervals, where we can observe only inertia and sensitivity to business cycles. Disagreement increases when output gap is below the steady state and falling. We observe higher disagreement when absolute inflation is below the target. Rich and Tracy (2010) and D'Amico and Orphanides (2008) find that there is a positive relationship between inflation and disagreement. Our results conversely point out that also low inflation can generate higher uncertainty.

There are some treatment differences regarding the determination of standard deviation of point forecast (*sdv*). In particular, treatment 3 seems to produce lower *sdv* compared to treatment 1. However, we are not able to introduce treatment dummies to the regressions for the *sdv* and $IQR$ as then we would not be able to compute clustered standard errors across treatments. The results in this paragraph are from estimations of eq. (11) with treatment dummies using robust standard errors.

**4.1. Dispersion of aggregate distribution.** Several central banks have started to put the data from distribution of inflation expectations on the agenda for policy meetings. This is partly a product of advances in Bayesian estimation methods for monetary models and also the adoption of new communication strategies by many central banks. Thus, it is often desirable to aggregate individual distributions and analyze them, rather than calculating averages from the individual moments. Frequently, only aggregate distributions are available from survey data, assuming that different samples of forecasters have similar aggregate properties to the whole population.

We derive the distribution from the asymmetric confidence bounds using a triangles approach similar to Engelberg, Manski, and Williams (2009). The mode is set to be equal to point forecast, while 95% of the derived triangular distribution is set to be between the lower and the upper confidence bound. This way we generate probability density functions for each forecast of an individual. Distributions are then aggregated (cross sectionally) across individuals in a group.

We choose interquartile range (IQR)[33] as it is less sensitive to small variations in tails of the estimated density compared to cross sectional standard deviation of the aggregate distribution as an appropriate measure.[34] Nevertheless, it is useful to show that the variance of aggregate distribution is related to the two measures that we study above. Boero, Smith, and Wallis (2008) show explicitly that the variance of the aggregate distribution can be decomposed into the average individual uncertainty and disagreement of point forecasts. To discover the properties of aggregate distribution, we run the following regression:

$$
\begin{aligned}
IQR_t^j &= \alpha + \zeta IQR_{t-1}^j + \beta D_1 y_{t-1} + \gamma D_2 y_{t-1} + \delta D_3 y_{t-1} \\
&\quad + \epsilon i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \eta m r_{t-1}^j + u_t^{em},
\end{aligned}
\tag{12}
$$

[33] The interquartile range is a range between the $25^{th}$ and $75^{th}$ percentile.

[34] Giordani and Söderlind (2003) use a similar measure to ours. In the literature also other measures were proposed. Boero, Smith, and Wallis (2008) use standard deviation of the aggregate distribution, while Batchelor and Dua (1996) suggest root mean subjective variance.

| | $sdv^j_{t+1\|t}$ | | | $IQR^j_{t+1\|t}$ | | |
|---|---|---|---|---|---|---|
| | *all* | *treat.A* | *treat.B* | *all* | *treat.A* | *treat.B* |
| $sdv^j_{t\|t-1}$ | 0.1463 | 0.1265 | 0.4970*** | | | |
| | (0.1409) | (0.1046) | (0.0247) | | | |
| $IQR^j_{t\|t-1}$ | | | | 0.4982*** | 0.4738*** | 0.6280*** |
| | | | | (0.0896) | (0.0787) | (0.0670) |
| $D_1 y_{t-1}$ | -0.0171 | 0.0122 | 0.0157 | -0.0298 | -0.0282 | 0.0385 |
| | (0.0154) | (0.0168) | (0.0376) | (0.0622) | (0.0854) | (0.0582) |
| $D_2 y_{t-1}$ | 0.0026 | 0.0136 | -0.1275*** | -0.0809 | -0.0713 | -0.1122** |
| | (0.0311) | (0.0250) | (0.0164) | (0.0492) | (0.0572) | (0.0475) |
| $D_3 y_{t-1}$ | 0.0392 | 0.0520 | -0.0249 | 0.0848 | 0.1073 | -0.0538 |
| | (0.0593) | (0.0749) | (0.0369) | (0.0841) | (0.1000) | (0.0402) |
| $i_{t-1}$ | 0.0279 | 0.0249 | -0.0002 | 0.0083 | 0.0076 | 0.0109 |
| | (0.0330) | (0.0288) | (0.0389) | (0.0131) | (0.0210) | (0.0246) |
| $D_L\|\pi_{t-1}\|$ | 0.1430*** | 0.1533*** | 0.0773 | 0.0758** | 0.0789*** | 0.0497 |
| | (0.0507) | (0.0353) | (0.0534) | (0.0294) | (0.0286) | (0.0345) |
| $D_H\|\pi_{t-1}\|$ | 0.0787 | 0.0901 | 0.0794 | 0.0438 | 0.0492 | 0.0022 |
| | (0.0717) | (0.0701) | (0.0675) | (0.0530) | (0.0615) | (0.0401) |
| $mr^j_{t-1}$ | 0.2211*** | 0.2447*** | 0.0704 | 0.2174*** | 0.2438*** | 0.0790 |
| | (0.0297) | (0.0169) | (0.0779) | (0.0348) | (0.0184) | (0.0959) |
| $\alpha$ | -0.0218 | -0.0252 | 0.0332 | 0.0739** | 0.0836 | 0.0668 |
| | (0.0911) | (0.0726) | (0.1211) | (0.0308) | (0.0610) | (0.0750) |
| $N$ | 1632 | 1088 | 544 | 1632 | 1088 | 544 |
| Wald $\chi^2_{(8)}$ | 3763.3 | 12747.1 | 5495.4 | 19215.1 | 15228.4 | 3032.3 |

Table 10: Analysis of Disagreement: Interquartile Range (left) and Standard Deviation of Point Forecasts (right). Note: Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in treatments. */**/*** denotes significance at 10/5/1 percent level.

where $IQR^j = Q_3 - Q_1$ is interquartile range, $y_t$ is output gap, $i_t$ is the interest rate, and $D_1, \ldots, D_3$ are dummy variables as identified above.

Equation (12) considers sources of divergences in expectations, such as output gap, interest rate or previous value of the interquartile range. As above we introduce a dummy variable for each of the phases of the cycle. Several studies observe considerable inertia in disagreement of expectations (see Giordani and Söderlind, 2003), therefore we also include previous period interquartile range among independent variables and find them highly significant. Results on the right part of Table 10 show that there is some influence of the cyclical phase and inflation on the interquartile range. For a negative and decreasing output gap there is more disagreement. This is similar to the results in survey data, where it is common to observe countercyclical behavior of variance of inflation expectations.[35] We observe that interquartile range is positively correlated

[35] Pfajfar and Santoro (2010) also study kurtosis and skewness of the distribution of forecasts and find that both exhibit procyclical behavior.

with the absolute level inflation when inflation is below the target level. In treatment A, mean absolute forecast error also significantly affects the IQR. It is worth noting that regressions for the treatments with symmetric and asymmetric confidence intervals show very similar results. Regression results yield no significant differences between the different monetary policy rules employed.

## 5. Discussion

The aim of this section is to compare different measures of individual uncertainty and disagreement among forecasters and to assess their ability to forecast inflation variability. Different studies argue that disagreement measured as standard deviation of point forecasts lacks a theoretical basis and therefore is not a suitable proxy for uncertainty and consequently also for inflation variability as is implicit in Zarnowitz and Lambros (1987). However as we pointed out above, Boero, Smith, and Wallis (2008) question this statement and show that disagreement is a component of the variance of aggregate distribution.

There are several advantages and disadvantages to each measure proposed. The choice of the measure should therefore be oriented to the purpose for which it is intended. Several survey data articles point out that the advantage of the measure for disagreement among forecasters (*sdv*) is that it is available in any survey, whereas only a limited number of surveys asks for measures of individual uncertainty. A proxy for the uncertainty may be average individual forecast error variance. Measure of the variance of the aggregate distribution of forecasts gives information about both, uncertainty and disagreement. Figures A1 and A2 in Appendix A display a timewise comparison between the average confidence interval, standard deviation of point forecasts and interquartile range for each group.

We compute pairwise correlation coefficients between different measures of uncertainty and disagreement as in D'Amico and Orphanides (2008) to compare these different measures and to make a preliminary assessment of their forecast ability which is further scrutinized below using dynamic panel regression analysis.

| | $\lvert r_t^k \rvert$ | $asip_t^j$ | $sdv_t^j$ | $IQR_t^j$ | $\pi_{t+1}$ | $i_{t+1}$ | $y_{t+1}$ |
|---|---|---|---|---|---|---|---|
| $\lvert r_t^k \rvert$ | 1 | | | | | | |
| $asip_t^j$ | 0.577*** | 1 | | | | | |
| $sdv_t^j$ | 0.822*** | 0.532*** | 1 | | | | |
| $IQR_t^j$ | 0.827*** | 0.689*** | 0.777*** | 1 | | | |
| $\pi_{t+1}$ | -0.080** | -0.030 | -0.063* | -0.131*** | 1 | | |
| $i_{t+1}$ | 0.169*** | 0.196*** | 0.226*** | 0.198*** | 0.845*** | 1 | |
| $y_{t+1}$ | -0.321*** | -0.246*** | -0.259*** | -0.267*** | -0.016 | -0.177*** | 1 |
| $sd_{t+1}^j$ | 0.818*** | 0.690*** | 0.722*** | 0.877*** | -0.143*** | 0.185*** | -0.280*** |

Table 11: Pairwise correlation coeficients. Note: */**/*** denotes significance at 10/5/1 percent level.

All three measures that we compare in this section are significantly positively correlated between each other and with the standard deviation of inflation. However, some of the correlation coefficients are not very high. As we can observe in Table 11 there is a significant correlation coefficient (about 0.5) between the average width of the confidence interval and the standard deviation of point forecasts.[36] Rich and Tracy (2010) and Boero, Smith, and Wallis (2008) find little evidence that this relationship exists in the survey data, while D'Amico and Orphanides (2008) find a correlation coefficient of 0.4. Present analysis suggests that uncertainty and disagreement are modestly correlated.

A positive correlation between interquartile range and individual uncertainty can be observed. Correlation coefficient (around 0.7) is higher than the one studied in the previous paragraph. As shown in the statistical analysis by Boero, Smith, and Wallis (2008) there exists a "structural" relationship between these two variables so a positive relationship is expected. Due to similar reasons there is also a correlation between the disagreement and the interquartile range. The latter correlation is of similar magnitude to the former one. Therefore one could argue that interquartile range is in our experiment at least as much, if not more, a measure of disagreement as average individual uncertainty. Bomberger (1996) argues that standard deviation of point forecasts is a useful proxy for uncertainty and that disagreement tracks uncertainty better than the GARCH model; however, the latter is questioned by Rich and Butler (1998).[37]

Policymakers are interested in inflation uncertainty and to obtain its proxies. Therefore the question that needs to be addressed is which proxy or combination of proxies best forecasts inflation uncertainty. As we can observe in Table 11 the highest correlation is between the interquartile range ($IQR$) and the standard deviation of inflation ($sd$). It reaches almost 0.9, while somehow surprisingly disagreement is a slightly better proxy of inflation uncertainty than average perceived uncertainty of subjects. In order to further assess the forecasting performance of these measures we estimate the following regression:

$$\begin{aligned} sd_t^j &= \alpha + \beta sd_{t-1}^j + \gamma asip_{t-1}^j + \epsilon sdv_{t-1}^j + \delta IQR_{t-1}^j \\ &\quad + \zeta i_{t-1} + \eta \pi_{t-1} + \phi y_{t-1} + u_t^{em}, \end{aligned} \tag{13}$$

where $asip_{t-1}^j$ is the average confidence interval in period $t-1$ for the group $k$. Table 12 reports the results. We estimate three different specifications which are the subset of the above equation. In the variant ($a$) we include all three measures, while in the variant ($b$) we include only measures of individual uncertainty and disagreement. Variant ($c$) embeds only $IQR$ as it is a measure of both individual uncertainty and disagreement and, as pointed out above, it is the measure that has the highest correlation with standard deviation of inflation.

Regressions confirm that the average individual uncertainty and the standard deviation of point forecasts have a positive effect on inflation variance. It comes as a surprise however that interquartile range has a marginally significant negative effect. This may be due to a

[36] Table B4 in Appendix B depicts the relationship between confidence bounds and dispersion of point forecasts in more details. We find no evidence of this relationship for symmetric intervals while for asymmetric there is a positive relationship.

[37] Lahiri and Sheng (2010) point out that disagreement is useful in stable periods for forecasting but not in periods of high volatility.

| $sd_t^j$ : | $(a)$ | $(b)$ | $(c)$ |
|---|---|---|---|
| $sd_{t-1}^j$ | 0.9913*** | 0.9843*** | 1.0036*** |
| | (0.0124) | (0.0116) | (0.0137) |
| $asip_{t-1}^j$ | 0.0837*** | 0.0708** | - |
| | (0.0298) | (0.0289) | |
| $sdv_{t-1}^j$ | 0.0106 | 0.0073 | - |
| | (0.0179) | (0.0152) | |
| $IQR_{t-1}^j$ | -0.0170* | - | -0.0018 |
| | (0.0100) | | (0.0114) |
| $i_{t-1}$ | 0.0108 | 0.0108 | 0.0129 |
| | (0.0084) | (0.0082) | (0.0096) |
| $\pi_{t-1}$ | -0.0135 | -0.0136 | -0.0148 |
| | (0.0118) | (0.0115) | (0.0137) |
| $y_{t-1}$ | -0.0094* | -0.0092* | -0.0125*** |
| | (0.0052) | (0.0049) | (0.0037) |
| $\alpha$ | -0.0109 | -0.0071 | 0.0169 |
| | (0.0227) | (0.0218) | (0.0178) |
| $N$ | 1656 | 1656 | 1656 |
| Wald $\chi^2_{(7,6,5)}$ | 54840.3 | 50525.4 | 22529.2 |

Table 12: Factors affecting the standard deviation of inflation. Note: Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in treatments. */**/*** denotes significance at 10/5/1 percent level.

degree of multicollinearity between $IQR$ and standard deviation of point forecasts and/or mean confidence intervals. In specification $(c)$ the effect of $IQR$ is insignificant, while in specifications $(a)$ and $(b)$ we observe that only average individual confidence interval has a positive and highly significant effect on $sd$. Therefore we can conclude that for forecasting inflation it is most important to know the average individual confidence interval, which is still rarely the case in surveys of inflation opinions. These regressions confirm the results from survey data literature, as we reach similar conclusions to those of Zarnowitz and Lambros (1987), Boero, Smith, and Wallis (2008), and Giordani and Söderlind (2003) who argue that average individual uncertainty is the proxy of inflation uncertainty that central banks should monitor.

Inflation affects standard deviation of inflation negatively, which might also be surprising. However, it is likely that if we separated the positive and negative developments of inflation we would find similar effects as in the above regressions for $IQR$ and $sdv$, i.e. both terms would enter significantly positively with negative development having a more profound effect. Output gap exerts a negative effect on $sdv$.

## 6. CONCLUSION

In this paper we design a macroeconomic experiment where subjects are asked to forecast inflation and its uncertainty. The underlying model of the economy is a simple NK model which is commonly used for the analysis of monetary policy. The focus is on the analysis of the confidence bounds reported by subjects as a perceived measure of the uncertainty in the economy. It has been shown that uncertainty has implications for both inflation outcomes and for unemployment and is an increasingly important indicator for monetary policy making. Similarly to inflation expectations, the formation of confidence bounds is also found to be heterogeneous. In different treatments we focus on various modifications of the original Taylor rule and study the influence of different monetary policy designs to the formation of confidence bounds. We find that inflation targeting produces lower uncertainty and higher accuracy of intervals than inflation forecast targeting. Also the treatment that reacts strongly to deviations of inflation expectations from the inflation target produces similar effects as stated above, compared to treatments that do not react as strongly to deviations of inflation forecasts. This effect does not only channel through the variability of inflation, but there is evidence that there are additional effects.[38]

Subjects on average underestimate risk. This is a standard result in the psychology literature and is known as overconfidence bias. We find that subjects only in 60.5% of cases correctly estimate risk. In particular, less than 10% of subjects on average report confidence bounds that represent approximately the 95% confidence intervals consistent with actual realizations; around 10% overestimate risk, while all others underestimate risk. We observe more cases of inflation falling outside the confidence interval when volatility of inflation is higher and when confidence intervals are narrower. Outcomes outside the interval are also more frequent when output gap is lower and has a downward trend, while in the opposite situation there is lower probability of misperceiving inflation uncertainty.

We also analyze measures of individual uncertainty, disagreement among forecasters and properties of aggregate distribution. All these measures are related as argued in Boero, Smith, and Wallis (2008), however they have very different features. Interquartile range is a measure of both uncertainty and disagreement. We first analyze the formation of confidence intervals. We find that confidence intervals are positively related to inflation variability, that they are highly inertial and that widen after an "error." It is also interesting to observe the relation between inflation and confidence intervals. In the survey data literature it has been established that these two variables are positively related, i.e. higher inflation is causing wider confidence intervals. Below target inflation also causes the interval to increase and that absolute deviations from inflation target is an appropriate variable to take into account. Furthermore, we can establish some facts about the differences between the formation of lower and upper bounds. In particular, we find that the upper bound is more sensitive to the stage of the business cycle while the lower bound exhibit significantly more inertia.

More generally, we also study the determinants of the choice of asymmetric interval. In our treatment B subjects have the possibility to insert an asymmetric confidence interval, while in treatment A they are restricted to symmetric intervals. We find that there are only about 12.5% of cases when subjects insert symmetric intervals when they have the possibility of inserting

[38] Pfajfar and Žakelj (2011) show that variability of inflation in this experiment depends on monetary policy rules.

an asymmetric interval. Moreover, in treatment B more than 35% of subjects report higher upper bounds than the lower ones, while there are only about 5% of subjects with the opposite pattern. It is more likely to observe symmetric intervals when the interest rate is high and less likely when inflation is below the target. Symmetric intervals are also more common when the output gap is positive and rising compared to the opposite stage of the business cycle.

What determines the evolution of standard deviation of point forecasts and the interquartile range of the aggregate distribution? We document that $IQR$ is more inertial than $sdv$ while they both increase when inflation is below the target level. We also compare forecasting performance of these measures and observe that the interquartile range of the aggregate distribution is the one that has the highest correlation with the actual uncertainty. Nevertheless, regression analysis suggests that the average individual confidence interval is the only measure that consistently enters significantly in our forecasting specifications. Therefore, we confirm previous results from the survey data literature that more central banks should design their surveys in order that each individual provides their whole distribution of forecasts or at least some measure of uncertainty of their forecasts. In this sense it might be enough if they are asked for their confidence intervals as in our treatment A. Generally, this would greatly enhance the informativeness of these surveys as central banks would also receive a proxy for forecasting inflation uncertainty.

## REFERENCES

ADAM, K. (2007): "Experimental Evidence on the Persistence of Output and Inflation," *Economic Journal*, 117(520), 603–636.

ANUFRIEV, M., AND C. HOMMES (2007): "Evolution of Market Heuristics," CeNDEF Working Paper 06, University of Amsterdam.

ARIFOVIC, J., AND T. J. SARGENT (2003): "Laboratory Experiments with an Expectational Phillips Curve," in *Evolution and Procedures in Central Banking*, ed. by D. E. Altig, and B. D. Smith. Cambridge University Press, Cambridge.

ASSENZA, T., P. HEEMEIJER, C. HOMMES, AND D. MASSARO (2011): "Individual Expectations and Aggregate Macro Behavior," CeNDEF Working Papers 11-01, Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance.

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107(3), 797–817.

BATCHELOR, R., AND P. DUA (1996): "Empirical Measures of Inflation Uncertainty: A Cautionary Note," *Applied Economics*, 28(3), 333–41.

BEN-DAVID, I., J. R. GRAHAM, AND C. R. HARVEY (2010): "Managerial Miscalibration," NBER Working Papers 16215, National Bureau of Economic Research, Inc.

BIAIS, B., D. HILTON, K. MAZURIER, AND S. POUGET (2005): "Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market," *Review of Economic Studies*, 72(2), 287–312.

BLOOM, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.

BLOOM, N., M. FLOETOTTO, AND N. JAIMOVICH (2010): "Really Uncertain Business Cycles," Mimeo, Stanford University.

BLUNDELL, R., AND S. BOND (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115–143.

BOERO, G., J. SMITH, AND K. F. WALLIS (2008): "Uncertainty and disagreement in economic prediction: the Bank of England Survey of External Forecasters," *Economic Journal*, 118(530), 1107–27.

BOMBERGER, W. A. (1996): "Disagreement as a Measure of Uncertainty," *Journal of Money, Credit and Banking*, 28(3), 381–92.

BOTTAZZI, G., AND G. DEVETAG (2005): "Expectations structure in asset pricing experiments," in *Nonlinear Dynamics and Heterogeneous Interacting Agents*, ed. by T. Lux, S. Reitz, and E. Samanidou, Lecture Notes in Economics and Mathematical Systems. Springer Verlag, Berlin, 550 edn.

Bottazzi, G., G. Devetag, and F. Pancotto (2010): "Does volatility matter? Expectations of price return and variability in an asset pricing experiment," *Journal of Economic Behavior & Organization*, In Press, Corrected Proof, –.

Bruine de Bruin, W., C. F. Manski, G. Topa, and W. van der Klaauw (2011): "Measuring consumer uncertainty about future inflation," *Journal of Applied Econometrics*, forthcoming.

Cooley, T. F., and E. C. Prescott (1995): "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley. Princeton University Press, Princeton.

D'Amico, S., and A. Orphanides (2008): "Uncertainty and disagreement in economic forecasting," Finance and Economics Discussion Series 2008-56, Board of Governors of the Federal Reserve System (U.S.).

Du, N., and D. V. Budescu (2007): "Does past volatility affect investors' price forecasts and confidence judgements?," *International Journal of Forecasting*, 23(3), 497–511.

Duffy, J. (2008): "Macroeconomics: A Survey of Laboratory Research," Working Papers 334, University of Pittsburgh, Department of Economics.

Engelberg, J., C. F. Manski, and J. Williams (2009): "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters," *Journal of Business & Economic Statistics*, 27, 30–41.

Evans, G. W., and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press.

Evans, G. W., S. Honkapohja, and R. Marimon (2001): "Convergence in monetary inflation models with heterogeneous learning rules," *Macroeconomic Dynamics*, 5(1), 1–31.

Evans, M., and P. Wachtel (1993): "Inflation Regimes and the Sources of Inflation Uncertainty," *Journal of Money, Credit and Banking*, 25(3), 475–511.

Fehr, E., and J.-R. Tyran (2008): "Limited Rationality and Strategic Interaction: The Impact of the Strategic Environment on Nominal Inertia," *Econometrica*, 76(2), 353–394.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178.

Friedman, M. (1968): "The role of monetary policy," *American Economic Review*, 58(1), 1–17.

——— (1977): "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy*, 85(3), 451–72.

Gilovich, T., D. Griffin, and D. Kahneman (2002): *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press, Cambridge.

Giordani, P., and P. Söderlind (2003): "Inflation forecast uncertainty," *European Economic Review*, 47(6), 1037–1059.

——— (2006): "Is there evidence of pessimism and doubt in subjective distributions? Implications for the equity premium puzzle," *Journal of Economic Dynamics and Control*, 30(6), 1027–1043.

Hoffrage, U. (2004): "Overconfidence," in *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*, ed. by R. F. Pohl, pp. 235–254. Psychology Press.

Hommes, C. (2011): "The heterogeneous expectations hypothesis: Some evidence from the lab," *Journal of Economic Dynamics and Control*, 35(1), 1–24.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005): "Coordination of Expectations in Asset Pricing Experiments," *Review of Financial Studies*, 18(3), 955–980.

Ireland, P. N. (2004): "Technology Shocks in the New Keynesian Model," *The Review of Economics and Statistics*, 86(4), 923–936.

Keren, G. (1991): "Calibration and probability judgements: Conceptual and methodological issues," *Acta Psychologica*, 77(3), 217 – 273.

Lahiri, K., and X. Sheng (2010): "Measuring forecast uncertainty by disagreement: The missing link," *Journal of Applied Econometrics*, 25(4), 514–538.

Lawrence, M., and S. Makridakis (1989): "Factors affecting judgmental forecasts and confidence intervals," *Organizational Behavior and Human Decision Processes*, 43(2), 172–187.

Lawrence, M., and M. O'Connor (1992): "Exploring judgemental forecasting," *International Journal of Forecasting*, 8(1), 15–26.

Levi, M. D., and J. H. Makin (1980): "Inflation Uncertainty and the Phillips Curve: Some Empirical Evidence," *American Economic Review*, 70(5), 1022–27.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips (1982): "Calibration of probabilities: the state of the art to 1980," in *Judgment under uncertainty: heuristics and biases*, ed. by D. Kahneman, P. Slovic, and A. Tversky, NBER Chapters, pp. 306–334. Cambridge University Press, Cambridge, UK.

Liu, F., and K. Lahiri (2006): "Modelling multi-period inflation uncertainty using a panel of density forecasts," *Journal of Applied Econometrics*, 21(8), 1199–1219.

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

Marcet, A., and T. J. Sargent (1989): “Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information,” *Journal of Political Economy*, 97(6), 1306–22.

Marimon, R., S. E. Spear, and S. Sunder (1993): “Expectationally Driven Market Volatility: An Experimental Study,” *Journal of Economic Theory*, 61(1), 74–103.

McCallum, B. T., and E. Nelson (2004): “Timeless Perspective vs. Discretionary Monetary Policy in Forward-Looking Models,” *Federal Reserve Bank of St. Louis Review*, 86(2), 43–56.

Mishkin, F. S. (2008): “Whither Federal Reserve Communications,” *Speech at the Peterson Institute for International Economics*, Washington, D.C., July 28, 2008.

Mullineaux, D. J. (1980): “Unemployment, Industrial Production, and Inflation Uncertainty in the United States,” *The Review of Economics and Statistics*, 62(2), 163–69.

Muradoglu, G., and D. Onkal (1994): “An exploratory analysis of portfolio managers' probabilistic forecasts of stock prices,” *Journal of Forecasting*, 13(7), 565–578.

Noussair, C. N., D. Pfajfar, and J. Zsiros (2011): “Frictions, persistence, and central bank policy in an experimental dynamic stochastic general equilibrium economy,” Discussion Paper 2011-030, Tilburg University, Center for Economic Research.

O'Connor, M., W. Remus, and K. Griggs (2001): “The asymmetry of judgemental confidence intervals in time series forecasting,” *International Journal of Forecasting*, 17(4), 623–633.

Onkal, D., and F. Bolger (2004): “Provider-user differences in perceived usefulness of forecasting formats,” *Omega*, 32(1), 31–39.

Oskamp, S. (1965): “Overconfidence in case-study judgements,” *The Journal of Consulting Psychology*, 29(3), 261–265.

Pfajfar, D., and E. Santoro (2010): “Heterogeneity, learning and information stickiness in inflation expectations,” *Journal of Economic Behavior & Organization*, 75(3), 426–444.

Pfajfar, D., and B. Žakelj (2011): “Inflation Expectations and Monetary Policy Design: Evidence from the Laboratory,” Mimeo, Tilburg University, Center for Economic Research.

Pons-Novell, J. (2003): “Strategic bias, herding behaviour and economic forecasts,” *Journal of Forecasting*, 22(1), 67–77.

Rich, R., and J. Tracy (2003): “Modeling uncertainty: predictive accuracy as a proxy for predictive confidence,” Staff Reports 161, Federal Reserve Bank of New York.

——— (2010): “The Relationships among Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts,” *The Review of Economics and Statistics*, 92(1), 200–207.

Rich, R. W., and J. S. Butler (1998): "Disagreement as a Measure of Uncertainty: A Comment on Bomberg," *Journal of Money, Credit and Banking*, 30(3), pp. 411–419.

Scharfstein, D. S., and J. C. Stein (1990): "Herd Behavior and Investment," *American Economic Review*, 80(3), 465–79.

Thaler, R. H. (2000): "From Homo Economicus to Homo Sapiens," *Journal of Economic Perspectives*, 14(1), 133–141.

Walsh, C. E. (2003): *Monetary Theory and Policy*. The MIT Press, second edn.

Williams, A. W. (1987): "The Formation of Price Forecasts in Experimental Markets," *Journal of Money, Credit and Banking*, 19(1), 1–18.

Woodford, M. (1996): "Loan Commitments and Optimal Monetary Policy," *Journal of Monetary Economics*, 37(3), 573–605.

——— (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Yaniv, I., and D. P. Foster (1995): "Graininess of Judgment Under Uncertainty: An Accuracy-Informativeness Trade-Off," *Journal of Experimental Psychology: General*, 124(4), 424 – 432.

Zarnowitz, V., and L. A. Lambros (1987): "Consensus and Uncertainty in Economic Prediction," *Journal of Political Economy*, 95(3), 591–621.

Zwiebel, J. (1995): "Corporate Conservatism and Relative Compensation," *Journal of Political Economy*, 103(1), 1–25.

## A. Appendix: Additional Tables and Figures

| | | Confidence bound | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Symmetric | | | Lower | | | Upper | | |
| Treat | Group | < | ≈ | > | < | ≈ | > | < | ≈ | > |
| 1-A | 1 | 100 | 0 | 0 | - | - | - | - | - | - |
| 1-A | 2 | 78 | 11 | 11 | - | - | - | - | - | - |
| 1-A | 3 | 89 | 11 | 0 | - | - | - | - | - | - |
| 1-A | 4 | 78 | 22 | 0 | - | - | - | - | - | - |
| 1-B | 5 | - | - | - | 89 | 11 | 0 | 0 | 22 | 78 |
| 1-B | 6 | - | - | - | 100 | 0 | 0 | 0 | 11 | 89 |
| 2-A | 7 | 44 | 11 | 44 | - | - | - | - | - | - |
| 2-A | 8 | 78 | 11 | 11 | - | - | - | - | - | - |
| 2-A | 9 | 100 | 0 | 0 | - | - | - | - | - | - |
| 2-A | 10 | 100 | 0 | 0 | - | - | - | - | - | - |
| 2-B | 11 | - | - | - | 100 | 0 | 0 | 0 | 0 | 100 |
| 2-B | 12 | - | - | - | 100 | 0 | 0 | 0 | 0 | 100 |
| 3-A | 13 | 56 | 22 | 22 | - | - | - | - | - | - |
| 3-A | 14 | 89 | 11 | 0 | - | - | - | - | - | - |
| 3-A | 15 | 56 | 11 | 33 | - | - | - | - | - | - |
| 3-A | 16 | 100 | 0 | 0 | - | - | - | - | - | - |
| 3-B | 17 | - | - | - | 100 | 0 | 0 | 0 | 0 | 100 |
| 3-B | 18 | - | - | - | 100 | 0 | 0 | 0 | 11 | 89 |
| 4-A | 19 | 78 | 11 | 11 | - | - | - | - | - | - |
| 4-A | 20 | 89 | 11 | 0 | - | - | - | - | - | - |
| 4-A | 21 | 67 | 0 | 33 | - | - | - | - | - | - |
| 4-A | 22 | 78 | 11 | 11 | - | - | - | - | - | - |
| 4-B | 23 | - | - | - | 78 | 0 | 22 | 11 | 11 | 78 |
| 4-B | 24 | - | - | - | 100 | 0 | 0 | 0 | 11 | 89 |
| All | | 80 | 9 | 11 | 96 | 1 | 3 | 1 | 8 | 90 |

Table A1: Percentage of subjects by group with underprediction/overprediction of confidence interval. Note: The benchmark confidence level is $1.96 * sd_{t-1}^{k}$. $<$ ($>$) identifies frequencies of subjects whose inputs are signifficantly lower (higher) than the benchmark value. $\approx$ identifies subjects whose input is not siggnificantly different from the benchmark. Based on t-tests.

| | All | | | Treatment A | | | Treatment B | | |
|---|---|---|---|---|---|---|---|---|---|
| Inflation | ↑ | ↓ | ∼ | ↑ | ↓ | ∼ | ↑ | ↓ | ∼ |
| Underprediction | 34.63 | 3.98 | 17.83 | 30.93 | 4.12 | 16.79 | 41.39 | 3.69 | 20.02 |
| Inside interval | 60.65 | 58.41 | 63.95 | 64.81 | 62.75 | 66.43 | 53.03 | 49.15 | 58.76 |
| Overprediction | 4.72 | 37.6 | 18.22 | 4.25 | 33.13 | 16.79 | 5.58 | 47.17 | 21.22 |

Table A2: Interval correctness depending on the phase of the inflation cycle (% of decisions). $\uparrow$ denotes cases when inflation increases for at least 2 last periods, and $\downarrow$ denotes cases when it decreases for at least 2 last periods. $\sim$ represents all other cases. Subjects "underpredict" when the actual inflation is larger than subject's predicted upper confidence bound; and "overpredict" when the actual inflation is lower than subject's predicted lower confidence bound.

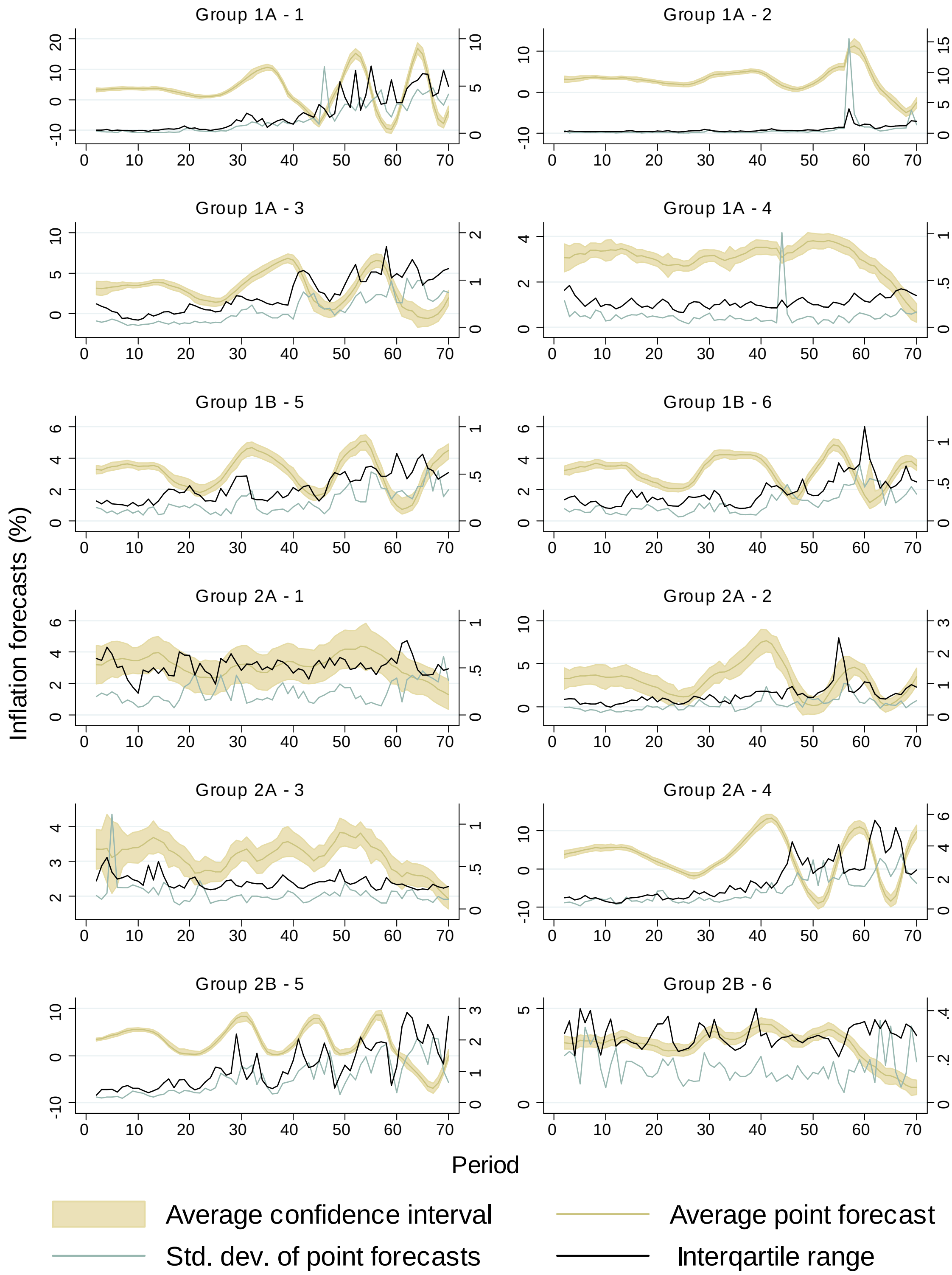


Figure A1: Average inflation forecasts and average confidence intervals (left axis) and disagreement and uncertainty measures (right axis) per group. Interquartile range is calculated from the aggregate expectation distribution as described in Section 4.1.

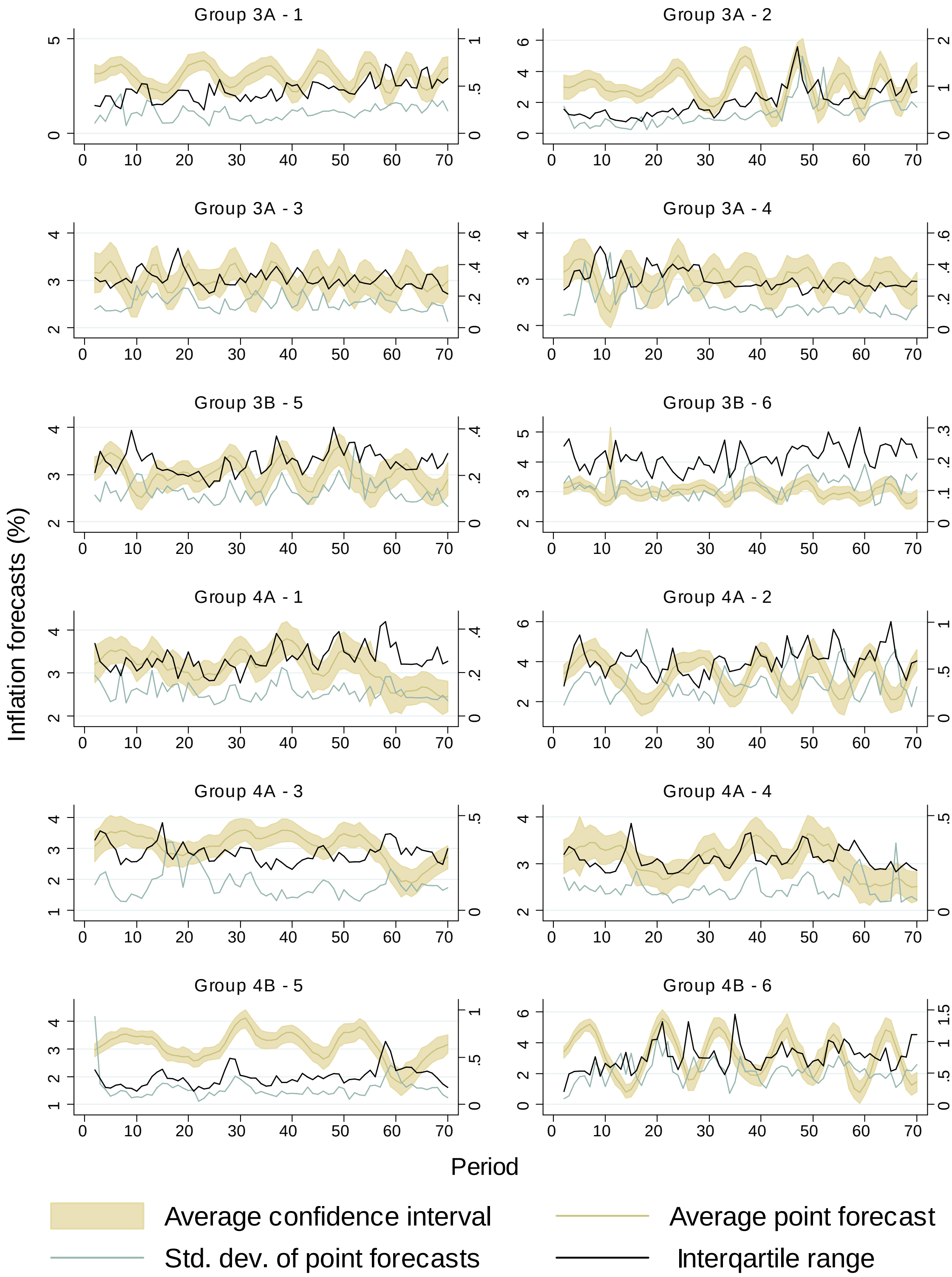


Figure A2: Average inflation forecasts and average confidence intervals (left axis) and disagreement and uncertainty measures (right axis) per group. Interquartile range is calculated from the aggregate expectation distribution as described in Section 4.1.

## B. APPENDIX: ADDITIONAL REGRESSIONS

In Table B7 we estimate the following regression using the system GMM estimator of Blundell and Bond (1998) for dynamic panel data:

$$r_{t+1}^k = \alpha + \beta r_t^k + \gamma sip_{t+1|t}^k + u_t^{em}. \quad (14)$$

Tables B8 and B9 report results of the following regressions:

$$\begin{aligned} D_z &= \alpha + \beta sip_{t|t-1}^k + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} + \zeta i_{t-1} \\ &\quad + \eta D_L |\pi_{t-1}| + \theta D_H |\pi_{t-1}| + \phi sd_{t-1}^j + u_t^{em}; \quad z \in \{7, 8, 9\}, \end{aligned}$$

where $D_7 = 1$ if the upper interval $(C_U)$ has exactly the same width as the lower interval $(C_L)$ one and 0 otherwise, $D_8 = 1$ when $|C_L - C_U| \leq 0.1$, and $D_9 = 1$ when $0.9 \leqslant \left|\frac{ConfIntH_{n-1}}{ConfIntL_{n-1}}\right| \leqslant 1.1$. Table B8 displays the results of logit estimations, while Table B9 presents the results of Poisson estimations.

| $x_t^k$ : | all | treat.A | treat.B |
|---|---|---|---|
| $sip_{t\|t-1}^k$ | 0.2989*** | 0.2260 | 0.3717*** |
| | (0.0723) | (0.1839) | (0.1011) |
| $D_1 y_{t-1}$ | -0.3103*** | -0.4341** | -0.2719*** |
| | (0.0759) | (0.2061) | (0.0870) |
| $D_2 y_{t-1}$ | 0.6037*** | 0.8818*** | 0.5563*** |
| | (0.0826) | (0.2076) | (0.0980) |
| $D_3 y_{t-1}$ | 0.0557 | 0.0553 | 0.0551 |
| | (0.0494) | (0.1065) | (0.0560) |
| $D_L\|\pi_{t-1}\|$ | 0.1164*** | 0.0893 | 0.1184*** |
| | (0.0256) | (0.1862) | (0.0369) |
| $D_H\|\pi_{t-1}\|$ | 0.2353*** | 0.4681** | 0.2035*** |
| | (0.0457) | (0.1872) | (0.0485) |
| $i_{t-1}$ | -0.0411* | -0.1070 | -0.0308 |
| | (0.0236) | (0.1307) | (0.0232) |
| $sd_{t-1}^j$ | -0.5094*** | -0.2835 | -0.5434*** |
| | (0.0778) | (0.1782) | (0.0980) |
| $N$ | 14904 | 4968 | 9936 |
| Wald $\chi^2_{(8)}$ | 180.7 | 110.1 | 106.7 |

Table B1: Forecasting accuracy and confidence intervals. Note: Table is based on equation (5). Coefficients are based on fixed effects Poisson estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| $x_t^k$ : | all | treat.A | treat.B |
|---|---|---|---|
| $sip_{t\|t-1}^k$ | 2.9756*** | 3.0827** | 2.9731*** |
| | (0.6019) | (1.2263) | (0.7409) |
| $D_1 y_{t-1}$ | -0.8511*** | -1.0369** | -0.7651*** |
| | (0.1643) | (0.4555) | (0.1931) |
| $D_2 y_{t-1}$ | 1.5409*** | 1.8617*** | 1.5010*** |
| | (0.2244) | (0.4738) | (0.2397) |
| $D_3 y_{t-1}$ | 0.2866* | 0.3315 | 0.2762 |
| | (0.1612) | (0.2886) | (0.1932) |
| $D_L \|\pi_{t-1}\|$ | 0.2753** | 0.1300 | 0.3216* |
| | (0.1113) | (0.5544) | (0.1784) |
| $D_H \|\pi_{t-1}\|$ | 0.6275*** | 1.0751** | 0.5522*** |
| | (0.1179) | (0.4374) | (0.1343) |
| $i_{t-1}$ | -0.1421** | -0.3303 | -0.0927 |
| | (0.0699) | (0.3517) | (0.0718) |
| $sd_{t-1}^k$ | -1.8189*** | -1.3832** | -1.9333*** |
| | (0.3840) | (0.6491) | (0.5259) |
| $\alpha$ | 0.6642*** | 0.6659 | 0.6702** |
| | (0.2467) | (1.1317) | (0.2871) |
| $\ln(\sigma_u^2)$ | -0.7610 | -0.7702 | -0.6338 |
| | (0.2130) | (0.4134) | (0.2611) |
| $\sigma_u$ | 0.6835 | 0.6804 | 0.7284 |
| | (0.0728) | (0.1406) | (0.0951) |
| $\rho^*$ | 0.1244 | 0.1234 | 0.1389 |
| | (0.0232) | (0.0447) | (0.0312) |
| $N$ | 14904 | 4968 | 9936 |
| Wald $\chi^2_{(8)}$ | 215.3 | 164.1 | 145.1 |

Table B2: Forecasting accuracy and confidence intervals. Note: Table is based on equation (5). Coefficients are based on random effects logit estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| $x_t^k$ : | all | treat.A | treat.B |
|---|---|---|---|
| $sip_{t\|t-1}^k$ | 0.2712*** | 0.2414 | 0.2646** |
| | (0.0649) | (0.1704) | (0.1103) |
| $D_1 y_{t-1}$ | -0.3258*** | -0.4259** | -0.3091*** |
| | (0.0738) | (0.1903) | (0.0879) |
| $D_2 y_{t-1}$ | 0.6062*** | 0.8427*** | 0.5633*** |
| | (0.0827) | (0.2259) | (0.0994) |
| $D_3 y_{t-1}$ | 0.0486 | 0.0806 | 0.0372 |
| | (0.0477) | (0.1113) | (0.0545) |
| $D_L\|\pi_{t-1}\|$ | 0.1184*** | 0.1119 | 0.1090*** |
| | (0.0247) | (0.1842) | (0.0332) |
| $D_H\|\pi_{t-1}\|$ | 0.2209*** | 0.4061** | 0.1858*** |
| | (0.0411) | (0.1779) | (0.0441) |
| $i_{t-1}$ | -0.0361* | -0.0935 | -0.0266 |
| | (0.0210) | (0.1279) | (0.0196) |
| $sd_{t-1}^j$ | -0.5481*** | -0.5110*** | -0.5363*** |
| | (0.0749) | (0.1286) | (0.0976) |
| $\alpha$ | -0.2598*** | -0.2505 | -0.2301*** |
| | (0.0587) | (0.3685) | (0.0631) |
| $\ln(\alpha^*)$ | -3.1464 | -2.8773 | -3.4026 |
| | (0.2286) | (0.2960) | (0.3133) |
| $\alpha^*$ | 0.0430 | 0.0563 | 0.0333 |
| | (0.0098) | (0.0167) | (0.0104) |
| $N$ | 14904 | 4968 | 9936 |
| Wald $\chi^2_{(8)}$ | 201.5 | 67.4 | 107.7 |

Table B3: Forecasting accuracy and confidence intervals. Note: Table is based on equation (5). Coefficients are based on random effects Poisson estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| $sip_{t+1\|t}^k$ : | *all* | *treat.A* | $treat.B-L$ | $treat.B-U$ |
|---|---|---|---|---|
| $sip_{t\|t-1}^k$ | 0.4472*** | 0.5530*** | 0.4468*** | 0.0991 |
| | (0.1058) | (0.0817) | (0.0418) | (0.1038) |
| $sdv_{t-1}^j$ | 0.1119** | 0.0993*** | 0.1472*** | 0.2929*** |
| | (0.0448) | (0.0373) | (0.0322) | (0.0788) |
| $\alpha$ | 0.2596*** | 0.2356*** | 0.1665*** | 0.2902*** |
| | (0.0366) | (0.0360) | (0.0277) | (0.0300) |
| $N$ | 14904 | 9936 | 4968 | 4968 |
| Wald $\chi^2_{(2)}$ | 58.9 | 129.8 | 114.6 | 65.5 |

Table B4: Confidence intervals and standard deviation of point forecasts. Note: Table is based on equation: $sip_{t+1|t}^k = \alpha + \beta sip_{t|t-1}^k + \gamma sdv_{t-1}^j + u_t^{em}$. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| $sip^k_{t+1\|t}$ : | *all* |
|---|---|
| $sip^k_{t\|t-1}$ | 0.4153*** |
| | (0.0998) |
| $sd^j_{t-1}$ | 0.1034** |
| | (0.0510) |
| $T2$ | 0.9459* |
| | (0.5606) |
| $T3$ | -0.6684 |
| | (0.6068) |
| $T4$ | -0.6889 |
| | (0.5834) |
| $\alpha$ | 0.3402* |
| | (0.2976) |
| $N$ | 14904 |
| Wald $\chi^2_{(6)}$ | 107.5 |

Table B5: Confidence intervals and standard deviation of inflation Note: Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| $sip^k_{t+1\|t}$ : | *all* | $treat.A$ | $treat.B-L$ | $treat.B-U$ |
|---|---|---|---|---|
| $sip^k_{t\|t-1}$ | 0.4636*** | 0.5719*** | 0.4780*** | 0.1090 |
| | (0.1028) | (0.0726) | (0.0532) | (0.1072) |
| $D^k_4$ | 0.0364* | 0.0228 | 0.0054 | 0.0788** |
| | (0.0215) | (0.0263) | (0.0121) | (0.0320) |
| $D^k_5$ | 0.0669*** | 0.0656** | 0.0735*** | 0.0261 |
| | (0.0233) | (0.0295) | (0.0216) | (0.0264) |
| $\alpha$ | 0.2718*** | 0.2489*** | 0.1802*** | 0.3480*** |
| | (0.0376) | (0.0364) | (0.0264) | (0.0400) |
| $N$ | 14688 | 9792 | 4896 | 4896 |
| Wald $\chi^2_{(3)}$ | 59.0 | 127.2 | 138.5 | 14.5 |

Table B6: Confidence intervals and the effect of forecast errors. Note: Table is based on equation: $sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma D^k_4 + \delta D^k_5 + u^{em}_t$. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| $r_{t+1}^k$ : | *all* | *treat.A* | *treat.B − L* | *treat.B − U* |
|---|---|---|---|---|
| $r_t^k$ | 0.6970*** | 0.6757*** | 0.8521*** | 0.8524*** |
| | (0.1376) | (0.1596) | (0.0250) | (0.0254) |
| $sip_{t+1\|t}^k$ | 0.0559 | 0.0812 | 0.6387*** | -0.1139* |
| | (0.0928) | (0.1102) | (0.1980) | (0.0619) |
| $\alpha$ | -0.0211 | -0.0401 | -0.2016*** | -0.0076 |
| | (0.0513) | (0.0651) | (0.0468) | (0.0299) |
| $N$ | 14688 | 9792 | 4896 | 4896 |
| Wald $\chi^2_{(3)}$ | 26.7 | 21.4 | 6625.1 | 4809.6 |

Table B7: Forecast errors and confidence intervals. Note: Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| **logit** | $D_7$ | $D_8$ | $D_9$ | $D_9$, fe |
|---|---|---|---|---|
| $sip^k_{t\|t-1}$ | 0.2178 | -1.0233* | -0.0194 | 0.0581 |
| | (0.1957) | (0.6213) | (0.2968) | (0.2239) |
| $D_1y_{t-1}$ | 0.2836 | 0.4825** | -0.0122 | -0.0028 |
| | (0.4162) | (0.2245) | (0.2913) | (0.2846) |
| $D_2y_{t-1}$ | -0.3912* | 0.1116 | -0.0490 | -0.0605 |
| | (0.2058) | (0.2339) | (0.1847) | (0.1859) |
| $D_3y_{t-1}$ | -0.3436 | 0.1057** | -0.0277 | -0.0288 |
| | (0.2645) | (0.0441) | (0.1667) | (0.1508) |
| $D_L\|\pi_{t-1}\|$ | 0.2375* | 0.2203*** | 0.1635 | 0.1629 |
| | (0.1354) | (0.0720) | (0.1163) | (0.1143) |
| $D_H\|\pi_{t-1}\|$ | 0.1510 | 0.1858 | 0.1494 | 0.1588 |
| | (0.2850) | (0.1214) | (0.1929) | (0.1842) |
| $i_{t-1}$ | -0.4047 | -0.2827** | -0.2041 | -0.1969 |
| | (0.2570) | (0.1204) | (0.1660) | (0.1637) |
| $sd^k_{t-1}$ | -0.1318 | -0.4759* | -0.2817** | -0.3295* |
| | (0.1381) | (0.2477) | (0.1330) | (0.1905) |
| $\alpha$ | -2.8695*** | -0.1098 | -1.5233*** | - |
| | (0.4649) | (0.3313) | (0.4104) | |
| $\ln(\sigma^2_u)$ | -0.4665 | -1.1088 | -0.7481 | - |
| | (0.2516) | (0.2653) | (0.2443) | |
| $\sigma_u$ | 0.7920 | 0.5744 | 0.6879 | - |
| | (0.0996) | (0.0762) | (0.0840) | |
| $\rho^*$ | 0.1601 | 0.0911 | 0.1258 | - |
| | (0.0338) | (0.0220) | (0.0269) | |
| $N$ | 4968 | 4968 | 4968 | 4968 |
| Wald $\chi^2_{(8)}$ | 48.3 | 72.3 | 29.2 | 34.0 |

Table B8: Determinants of symmetric intervals. Note: Coefficients are based on random effects logit estimations, except for "$D_9$, fe" which is based on fixed effects logit estimation. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

| **Poisson** | $D_7$ | $D_8$ | $D_9$ | $D_9$, fe |
|---|---|---|---|---|
| $sip^k_{t\|t-1}$ | 0.1717 | -0.7314** | -0.0307 | 0.0443 |
| | (0.1412) | (0.3667) | (0.2476) | (0.1641) |
| $D_1 y_{t-1}$ | 0.2215 | 0.2141* | -0.0092 | 0.0032 |
| | (0.3388) | (0.1165) | (0.2046) | (0.1989) |
| $D_2 y_{t-1}$ | -0.2974* | 0.0625 | -0.0298 | -0.0491 |
| | (0.1598) | (0.1137) | (0.1276) | (0.1301) |
| $D_3 y_{t-1}$ | -0.2727 | 0.0704* | -0.0106 | -0.0147 |
| | (0.2216) | (0.0392) | (0.1234) | (0.1050) |
| $D_L\|\pi_{t-1}\|$ | 0.1922* | 0.1043** | 0.1114 | 0.1106 |
| | (0.1092) | (0.0408) | (0.0820) | (0.0801) |
| $D_H\|\pi_{t-1}\|$ | 0.1141 | 0.0917 | 0.1004 | 0.1146 |
| | (0.2330) | (0.0759) | (0.1411) | (0.1353) |
| $i_{t-1}$ | -0.3303 | -0.1336** | -0.1413 | -0.1298 |
| | (0.2126) | (0.0619) | (0.1146) | (0.1115) |
| $sd^j_{t-1}$ | -0.0802 | -0.2374*** | -0.1933** | -0.2476** |
| | (0.1356) | (0.0906) | (0.0841) | (0.1203) |
| $\alpha$ | -2.6585*** | -0.6834*** | -1.6007*** | - |
| | (0.3787) | (0.1597) | (0.2862) | |
| $\ln(\alpha^*)$ | -0.8804*** | -2.8617*** | -1.5552*** | - |
| | (0.2106) | (0.9207) | (0.2275) | |
| $\alpha^*$ | 0.4146 | 0.0572 | 0.2111 | - |
| | (0.0873) | (0.0526) | (0.0480) | |
| $N$ | 4968 | 4968 | 4968 | 4968 |
| Wald $\chi^2_{(8)}$ | 40.3 | 71.3 | 21.9 | 27.4 |

Table B9: Determinants of symmetric intervals. Note: Coefficients are based on random effects poisson estimations, except for "$D_9$, fe" which is based on fixed effects logit estimation. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

## C. Instructions for the Experiment

Thank you for participating in this experiment, a project of economic investigation. Your earnings depend on your decisions and the decisions of the other participants. There is a show up fee of 5 Euros assured. From now on until the end of the experiment you are not allowed to communicate with each other. If you have any questions raise your hand and one of the instructors will answer the question in private. Please do not ask aloud.

**The experiment.** All participants receive exactly the same instructions. You and 8 other subjects all participate as agents in the *same* fictitious economy. You will have to predict future values of given economic variables. The experiment consists of 70 periods. The rules are the same in all the periods. You will interact with the same 8 subjects during the whole experiment.

Imagine that you work in a firm where you have to predict inflation for the next period. Your earnings depend on the accuracy of your inflation expectation.

**Information in each period.** The economy will be described with 3 variables in this experiment: the *inflation rate*, the *output gap*, and the *interest rate*.

- **Inflation** measures a general rise in prices in the economy. In each period it depends on the inflation expectations of the agents in economy (you and the other 8 participants in this experiment), output gap and random shocks which have equal probability to have positive or negative effect on inflation and are normally distributed.
- The **output gap** measures for how much (in percents) the actual Gross Domestic Product differs from the potential one. If the output gap is greater than 0, it means that the economy is producing more than the potential level, if negative, less than the potential level. It depends in each period on the inflation expectations of the agents in the economy, past output gap, interest rate and random shocks which have equal probability to have a positive or negative effect on inflation and are normally distributed.
- The **interest rate** is (in this experiment) the price of borrowing the money (in percents) for one period. The interest rate is set by the monetary authority. Their decision mostly depends on inflation expectations of the agents in the economy.

All given variables might be relevant for inflation forecast, but it is up to you to work out their relation and possible benefit of knowing them. The evolution of variables will partly depend on the inputs of you and other subjects and also different exogenous shocks influencing the economy.

- You enter the economy in period 1. In this period you will be given computer generated past values of inflation, output gap and interest rate for 10 periods back (Called: -9, -8, ... -1, 0)
- In period 2 you will be given all past values as seen in period 1 plus the value from period 1 (Periods: -9, -8, ... 0, 1).

- In period 3 you will see all past values as in period 2 (Periods: -9, -8, ... 1, 2) plus YOUR prediction about inflation in period 2 that you made in period 1.
- In period $t$ you will see all past values of actual inflation up to period $t-1$ (Periods: -9, -8, ... $t-2, t-1$) and your predictions up to period $t-1$ (Periods: 2, 3, ... $t-2, t-1$).

**What do you have to decide?.** Your task is to predict the state of the economy as accurately as possible. Your payoff will depend on the accuracy of your prediction of the inflation in the future period. In each period your prediction will consist of two parts:

a) *Expected inflation*, (in percents) that you expect to be in the NEXT period ($Exp.Inf.$)

b) *Lower bound* (in percents) of your prediction. You must be almost sure that the actual inflation will be higher than your lower bound.

c) *Upper bound* (in percents) of your prediction. You must be almost sure that the actual inflation will be smaller than your upper bound.

Based on b) and c) we determine the confidence interval, $Conf.Int.$ which is equal to

$$Conf.Int. = Upper\ bound\ -\ Lower\ bound$$

**Example 1.** *Let's say you think that inflation in the next period will be* 3.7%. *And you also think there is most likely (*95% *probability) that the actual inflation will not be lower than* 3.2% *and not higher than* 4.0%. *Your inputs in the experiment will be* 3.7 *under a),* 3.2 *under b), and* 4.0 *under c).*

Your goal is to maximize your payoff, given with the equation:

$$W = \max\left\{\frac{100}{1+|Inflation - Exp.Inf.|} - 20, 0\right\} + \max\left\{\frac{100x}{1+\frac{1}{2}Conf.Int.} - 20, 0\right\}$$

where $Exp.Inf.$ is your expectation about the inflation in the NEXT period, $Conf.Int.$ is the confidence interval, Inflation is the actual inflation in the next period and $x$ is variable with value 1 if

$$Lower\ bound \leq Inflation \leq Upper\ bound$$

and 0 otherwise.

The *first part* of the payoff function states that you will receive some payoff if the actual value in the next period will differ from your prediction in this period by less than 4 percentage points. The smaller this difference is, the higher the payoff you receive. With a zero forecast error ($|Inflation - Exp.Inf.| = 0$), you would receive 80 units ($100/1 - 20$). However, if your forecast is 1 percentage point higher or lower than the actual inflation rate, you will get only 30 units ($100/2 - 20$). If your forecast error is 4 percentage points or more, you will receive 0 units ($100/5 - 20$).

The *second part* of the payoff function simply states that you will get some extra payoff if the actual inflation is within your expected interval and if that interval is not larger than 8 percentage points. The more certain of the actual value you are, the smaller interval you give (*Lower bound* and *Upper bound* closer to $Exp.Inf.$), and the higher will be your payoff if the actual inflation indeed is in the given interval, but there will also be higher chances that actual value will fall outside your interval. In our example this interval was 0.8 percentage points. If the actual inflation falls in this interval you would receive 51.4 units $(100/(1+\frac{1}{2}0.8)-20)$ in addition to the payoff from the first part of the payoff function. If the actual values is outside your interval, your receive 0.

In the attached sheet you will find the table showing various combinations of *forecast error* and *confidence interval* needed to earn a given number of points.

**Information after each period.** Your payoff depends on your predictions for the next periods and actual realization in the next period. Because the actual inflation will be known only in the next period, you will also be informed about you current period ($t$) prediction and earnings after the end of the NEXT period ($t+1$). Therefore:

- After Period 1 you will not receive any earnings, since you did not make any prediction for the period 1.
- In any other period, you will receive the information about the actual inflation rate in this period and your *inflation* and *confidence interval* prediction from the previous period. You will also be informed if the actual inflation value was in your expected interval and what your earnings are for this period.

The units in the experiment are fictitious. Your actual payoff (in euros) will be the sum of earnings from all periods divided by 500.

If you have any questions please ask them now!

**Questionnaire**[39]

1. If you believe that inflation in the next period will be _ _4.2%_ _, and you are quite sure that it will not go down for more than _ _ 0.4_ _ nor up for more than _ _ 0.7_ _, you will type:
   Under (1) _ _ _ _ _ _ _ _ _ _ _ for inflation,
   Under (2) _ _ _ _ _ _ _ _ _ _ _ for the lower bound, and
   Under (3) _ _ _ _ _ _ _ _ _ _ _ for the upper bound.
2. If you are now in period _ _ _ _15_ _ _ _, you have information about past inflation, output gap and interest rate up to period _ _ _ _ _ _ _ _ _ _ and you have to predict the inflation for period _ _ _ _ _ _ _ _ _ _ _.

[39] Options (1), (2) and (3) are pointing to the different fields on the screenshot of the experimental interface.