Experimental investigations on market behavior

Blaž Žakelj

TESI DOCTORAL UPF / ANY 2011

DIRECTOR DE LA TESI

Prof. Rosemarie Nagel (Departament d'Economía i Empresa, Universitat Pompeu Fabra)

UNIVERSITAT POMPEU FABRA

To the memory of my grandmother, Ivanka

Acknowledgment

This thesis was written during my graduate studies at Universitat Pompeu Fabra in Barcelona. Department of Economics and Business at UPF provided me with an excellent research environment and education through courses, seminars and workshops.

First of all I would like to thank my supervisor Rosemarie Nagel. She introduced me to the exciting field of experimental economics and supported me throughout my graduate studies. Without her patience and help this work wouldn't have been possible.

I am grateful to Damjan Pfajfar, my coauthor and my friend. Through our numerous discussions and days preparing and analyzing our experiments we learned a lot about the concepts discussed in this thesis.

I have received valuable comments from John Duffy, Frank Heinemann, Ramon Marimon, Antoni Bosch-Domènech, Nagore Iriberri, and Jordi Brandts.

I am indebted to Aniol Llorente-Saguer and Juan Manuel Puerta for their advice and generous help with the execution of the laboratory experiments. Başak Güneş and Zeynep Gürgüç provided valuable help during design stages of my experiments.

I should not forget Marta Araque, Marta Aragay and Gemma Burballa for all their organizational support.

I would also like to thank my colleagues, with which I shared my life at UPF. My office mates, Juan Manuel, Rhiannon, Miloš, Mitja, Margherita and Geronimo, and my friends Raša, Petya, Virginie, Marta, Sandro and Gueorgui all helped broadening my knowledge with serious and less serious conversations and made my time in Barcelona joyful.

Finally, I would like to thank to my parents, who offered unconditional support and encouragement through all the years of my studies.

Abstract

This thesis is a collection of three essays on inflation expectations, forecasting uncertainty, and the role of uncertainty in sequential auctions, all using experimental approach. Chapter 1 studies how individuals forecast inflation in fictitious macroeconomic setup and analyzes the effect of monetary policy rules on their decisions. Results display heterogeneity in inflation forecasting rules and demonstrate the importance of adaptive learning forecasting if model switching is assumed. Chapter 2 extends the analysis from Chapter 1 by analyzing individual inflation forecasting uncertainty. Results show that confidence intervals depend on inflation variance and business cycle phase, have a strong inertia, and are often asymmetric. Finally, Chapter 3 analyzes the role of uncertainty about the number of bidders for the behavior of subjects in a sequential auction experiment. Uncertainty does not aggravate price decline, but it changes individual bidding strategies and auction efficiency.

Resumen

Esta tesis consta de tres ensayos sobre las expectativas de inflación, la incertidumbre de la predicción, y la importancia de la incertidumbre en subastas secuenciales. Todos ellos utilizan un método experimental. El capítulo 1 estudia cómo los individuos predicen la inflación en la economía ficticia y analiza el efecto de las reglas de política monetaria en sus decisiones. Los resultados revelan la heterogeneidad en las reglas de predicción de la inflación y demuestran la importancia del mecanismo de aprendizaje adaptivo si el cambio entre los modelos se supone. Capítulo 2 continúa el análisis del capítulo 1, analiza la incertidumbre individual de las expectativas de inflación. Los resultados muestran que los intervalos de confianza dependen de varianza de la inflación y la fase del ciclo económico, tienen una fuerte inercia, y son frecuentemente asimétricos. Por último, el capítulo 3 analiza la influencia de la incertidumbre sobre el número de oferentes en el comportamiento de los individuos en un experimento de la subasta secuencial. La incertidumbre no agrava la caída de los precios, pero cambia las estrategias de los oferentes y la eficiencia de la subasta.

Foreword

The aim of this thesis is to analyze the economic decisions of people in environments with imperfect information. We design laboratory experiments to demonstrate how participants differently assess a simulated economy environment and how their assessments influence the evolution of the economy. In a different set of experiments, we engage people in a simple auction trading to establish how market uncertainty affects individual participation and resulting prices.

The thesis consists of three essays that address the uncertainty in economic behavior from different perspectives. The first two have been written in collaboration with Damjan Pfajfar. In Chapter 1, “Inflation Expectations and Monetary Policy Design”, we study how subjects form inflation expectations in fictitious macroeconomic setup, and analyze how monetary policy rules affect individual behavior. In Chapter 2, “Uncertainty and Disagreement in Forecasting Inflation”, we extend the experimental setup of the first essay, identify the determinants of individual uncertainty and demonstrate how it influences the variability of inflation. Chapter 3, “Uncertainty about the Number of Bidders in Sequential Auctions with Unit Demand” analyzes the role of uncertain competition size on the bidding behavior and the consequences for the resulting prices.

There are several reasons why expectations are important in economic policymaking. Private agents make current decisions on the basis of beliefs about the future state of the economy, while policy makers have to incorporate the expected behavior of private agents into their policies. Private agents in turn, will also adjust their behavior on the basis of implemented policy changes. The macroeconomic models should therefore be designed with parameters that are independent of the policy change. This was one of the main points of the 1976 Lucas’ critique which importantly contributed to the rise of the micro-founded macroeconomic models. These models are constructed by aggregating models of individual agent’s behavior, where expectations play an important role. Expectations are most frequently modeled by using rational expectations paradigm. Rational expectations assume agents’ knowledge about the underlying model, knowledge that is shared among all participating agents. This enables agents to derive an optimal forecast rule which remains constants under given policy regime. Rather strong assumptions of the rational expectations hypothesis have been subject of criticism that they are unrealistic (see Sargent, 1993). Several behavioral models of expectations have since gained ground, adaptive learning (see Evans and Honkapohja, 2001) being one of the most important ones. These models assume that agents are boundedly rational when forecasting the state of the economy. Allowing for heterogeneity and using expectations models other than rational expectations may be inconvenient for analysis since properties like model consistency, unbiased forecasts, and the fact that prices fully reflect the information available to the market. Moreover, they may not converge to optimal (rational expectations) equilibrium or they may converge to some other equilibria. An empirical investigation about the ways how expectations are formed is an important factor in confirming the validity of alternative behavioral models.

Survey data based analysis is the most common approach of empirically testing the properties of expectations distribution and its evolution. Examples of such works are Carroll (003a) and Branch (2004, 2007). Experimental work about expectation formation has gained increasing intention especially since Marimon and Sunder (1993) demonstrated how adaptive learning can explain experimental subjects converging to low inflation equilibria, thus solving the equilibrium indeterminacy problem. Analysis based on laboratory experiments, has some clear advantages over survey based analysis. Experimentalist has control over the number of forecasters, number of forecasting periods and most importantly, over the information set available to subjects. Notable forecasting experiments include Hommes et al. (2005b), Adam (2007) and Heemeijer et al. (2009).

Our work in Chapter 1 distinguishes from the existing literature by focusing the analysis on the individual behavior and comparing how different monetary policy rules affect the choice of forecasting rules by subjects. We use a simplified version of New Keynesian framework as the underlying model of an experiment where subjects have to forecast inflation on the basis of past values of inflation, output gap, and interest rate. Different treatments are defined by alternative specifications of Taylor rule. These specifications determine how strong the stabilizing effect of the monetary policy on the economy is. We test the behavior of each individual against several simple models of forecasting, including adaptive expectations, rational expectations, autoregressive model, trend extrapolation, and adaptive learning. We show that subjects are indeed heterogeneous in ways how they forecast inflation and that large proportion of subjects use backward looking models for forecasting. The proportion of backward looking subjects increases when the variance of inflation is higher. We also show that there is a significant proportion of subjects that use some form of adaptive learning to make their forecasts.

In Chapter 1 we also demonstrate that models that allow the use different simple models interchangeably in the forecasting process perform better in describing subjects' behavior. Results represent an empirical support for the models that assume endogenous switching between forecasting models. We also analyze the effect of different instrumental rules across treatments. Rules that use actual rather than forecasted inflation produce lower inflation variability and alleviate expectational cycles.

In Chapter 2 the same macroeconomic framework is used to analyze inflation expectations. However, in this chapter the focus of analysis is on the underlying uncertainty, rather than on the point forecasts. Subjects are asked to provide their 95% confidence interval along their point forecast of inflation. The concept of inflation uncertainty is especially important for central banks. It has been shown (for example Levi and Makin, 1980) that higher inflation uncertainty is accompanied with lower output. Inflation targeting central banks can see inflation uncertainty as a measure of effectiveness of their communication strategies. However, analyzing the aggregate distribution of expectations alone does not represent the best way of assessing uncertainty since expectation variance can be caused by both, individual uncertainty and disagreement between individual forecasts. Only comparison of the individual expectations can help to distinguish between the two effects.

Empirical research on the inflation expectation is mostly based on the survey data. Zarnowitz and Lambros (1987) for example show that distinguishing individual uncertainty and interpersonal disagreement is important since the two can vary quite differently. Later research has offered mixed evidence regarding the role of disagreement in forecasting uncertainty. Experimental research in forecasting uncertainty has been done mainly in psychology literature where studies focus mainly on independent event forecasts while only a few studies ask subjects for confidence intervals in experimental economic research.

We add to existing literature by providing a detailed time-series analysis on the determinants that affect the individual uncertainty and those that affect disagreement. We show that there are different factors affecting the two measures which helps to explain different effect they have on explaining inflation variation. These effects are analyzed by using average confidence interval, average forecast error, standard deviation of point forecasts and interquartile range of the aggregate distribution for comparison. We identify the overconfidence effect in the analysis of forecasting accuracy. Individual expectation distribution skewness is also analyzed. We design treatments where subjects have to provide symmetric confidence intervals and those, where subjects can provide lower and upper bound along with their point forecast.

Chapter 3 diverges from the first two chapters, by analyzing the behavior of subjects in a sequential auction experiment. In particular we concentrate on the role of uncertainty about the size of competition on the market and its influence on the auction price. It has been long known that in an auction where identical goods are sold in a sequence, prices for the later goods are usually lower than the prices for earlier goods. Weber (1983) provides the conditions under which sequential sale of identical goods should result in the same price on average for all the goods. Empirical research trying to explain the price decline phenomenon has mostly concentrated on investigating conditions proposed by Weber. The role of uncertainty about the number of competitors for the price formation in sequential unit auctions has received less attention.

We provide a theoretical example that under given assumptions prices of two units of the same good sold in a sequence remain constant even if the number of bidders is unknown, as long as there is common knowledge of the probability distribution function over the number of bidders. We add to the literature by designing an auction experiment which investigates the effect of competition uncertainty on average prices for two units of the same good sold in a sequence. Our experiment consists of three treatments: one with 3 bidders; second with 6 bidders and third with either 3 or 6 bidders that are uncertain about the number of their competitors. We identify price decline phenomenon in all three treatments and analyze mean auction prices, action efficiencies and seller's revenues. We also provide some analysis of individual bidding strategies and show that uncertainty influences subjects' strategies even when optimal strategy does not depend on the number of competitors.

Contents

Abstract **vii**

Foreword **ix**

1 Inflation Expectations and Monetary Policy Design **1**

1.1 Introduction 1

1.2 Model 5

1.2.1 Monetary policy reaction functions 7

1.2.2 Calibration 8

1.3 Experiment 8

1.3.1 Design 8

1.3.2 Treatments 10

1.3.3 Summary of results 12

1.4 Analysis of individual inflation forecasts 15

1.4.1 Tests of rational expectations 16

1.4.2 Sticky information type regression 17

1.4.3 Trend extrapolation rule 17

1.4.4 Estimating simple learning rules 18

1.4.5 "General" models of expectation formation 20

1.4.6 "Classical econometrician" and rational expectations 21

1.4.7 Discussion 24

1.5 Switching between different models 26

1.5.1 Unrestricted switching 27

1.6 Monetary policy in the presence of heterogeneous expectations 31

1.7 Conclusion 37

2 Uncertainty and Disagreement in Forecasting Inflation **39**

2.1 Introduction 39
2.2 Experimental design 43
2.2.1 Experimental procedures 45
2.3 Individual uncertainty 46
2.3.1 Forecasting accuracy 49
2.3.2 Determinants of individual uncertainty 52
2.3.3 When are confidence intervals (a)symmetric? 57
2.4 Disagreement and aggregate expectation distribution 58
2.4.1 Dispersion of aggregate distribution 60
2.5 Discussion 62
2.6 Conclusion 66

3 Uncertainty about the Number of Bidders in Sequential Auctions with Unit Demand 69
3.1 Introduction 69
3.2 Related literature 70
3.3 Theoretical considerations 74
3.3.1 Assumptions 74
3.3.2 Auction with given number of bidders 75
3.3.3 Uncertainty about the number of bidders 78
3.4 Experimental design 81
3.5 Results 83
3.5.1 First- versus second-unit bids 83
3.5.2 First- versus second-unit prices 84
3.5.3 Individual bids 86
3.6 Conclusion 89

References 91

A Additional information for Chapter 1 103

B Additional information for Chapter 2 115

C Additional information for Chapter 3 131

Chapter 1

Inflation Expectations and Monetary Policy Design

1.1 Introduction

This chapter discusses an experimental study on the expectations formation process within a macroeconomic framework. Recently, with the development of explicit microfounded models expectations have become pivotal in the modern macroeconomic theory. Central banks increasingly attribute more importance to the developments of households' inflation expectations as they signal future inflationary risks. In line with this development, several theoretical models concerning expectations formation process have been proposed. They postulate informational frictions and heterogeneity of expectations as the main features of the expectation formation process. However, so far these models and their main assumptions have not been subject to rigorous empirical tests. A thorough assessment must rely on micro-level data and the associated distribution, while empirical contributions so far mostly employ aggregate data.[1] Moreover, to evaluate some new theories of expectation formation, e.g. adaptive learning,[2] we need to assure that agents' current information sets encompass all the information from the previous periods. Controlled laboratory environment avoids these methodological issues that are present in the survey data. In Chapter 1 we analyze individual data on inflation expectations gathered from an experimental economy and test them for different theoretical models. Insights into households' expectation formation provide useful guidance to central banks how to anchor inflation expectations.

[1] Recently, there have been some studies based on micro survey data, e.g. Branch (2004, 2007) and Pfajfar and Santoro (2010). These studies have confirmed that agents only infrequently update their information sets and that they use different theoretical models to forecast inflation.

[2] Adaptive learning assumes that subject are acting as econometricians when forecasting, i.e. reestimating their model each time new data becomes available. See Evans and Honkapohja (2001).

After establishing some stylized facts about that we focus on the relationship between policy actions and the formation of inflation expectations. Better understanding of this relationship has been stressed by the Chairman of the Federal Reserve Bernanke (2007) as crucial for the conduct of monetary policy. Advantage of our experiment lies in the usage of the New Keynesian framework and in possibility to compare the aggregate dynamics of inflation and output gap and the effectiveness of monetary policy with the results from the theoretical analysis. We study this question by employing several simple monetary policy rules in different treatments and examine potential implications of the design of monetary policy for forecasting inflation.

We provide substantial evidence in support of heterogeneity in the forecasting process. When analyzing individual responses from students of the Universitat Pompeu Fabra and Tilburg University, we find that agents form expectations in accordance with different theoretical models. In our sample approximately $30-45\%$ of agents are rational and around $30-35\%$ of agents predominately extrapolate trend. In addition, $15-25\%$ of agents mostly behave in line with adaptive learning and sticky information type models and about $5-10\%$ of agents forecast in an adaptive manner, updating their forecast with respect to the last observed error. Therefore, contrary to the findings of previous experimental studies, we observe a significant proportion of rational subjects. However, as it is not straightforward to define a rational subject, we explore different definitions in order to establish some robustness of our conclusions. Adaptive learning results are also novel as our work represents one of the first estimations of the gain parameter. The average gain of agents that employ adaptive learning models is around 0.045. Furthermore, when we allow agents to switch between different models, we find that adaptive learning models are the most popular models for forecasting inflation: 36.7% of all forecasts in our experiment are made with this class of models.

Rather than sticking to one model, switching between alternative models seems to describe subjects' behavior better. We observe that on average subjects switch every 4 periods. Therefore, Chapter 1 provides an empirical support for models that postulate endogenous switching, and assume that it is not always optimal to form beliefs in a rational way (e.g. Brock and Hommes, 1997). It could be optimal for some agents in at least some periods to commit to systematic errors as this might be less costly than using a rational rule. Furthermore, we also show that agents use different models as on average in each period 4.5 different models are used in groups of 9 subjects. This suggests that observed heterogeneity is pervasive.

Only a few experimental studies investigate the expectation formation process. The first experiments were performed in a no-feedback environment (e.g. Schmalensee, 1976) and lately some studies have also incorporated a feedback effect in their framework. However, these tend to analyze the expectation processes in an asset pricing setup. Some tests of the rational expectation hypothesis have been conducted within a simple macroeconomic setup (e.g. Williams, 1987; Marimon et al., 1993; Evans et al., 2001; Adam, 2007).[3] These studies mainly focus on aggregate expectations formation and tend to reject the rational expectations assumption in favor of adaptive way of forming beliefs. On the contrary, we focus on the analysis on the behavior of individuals. Our framework allows us to ask the same agents to provide their forecasts over the whole time span. Some analysis of the micro expectations data is conducted by Marimon and Sunder (1995) and Bernasconi and Kirchkamp (2000) in an overlapping generations framework. These authors estimate several different regressions in order to study inflation expectation formation and find that most subjects behave adaptively, although Bernasconi and Kirchkamp (2000) provide evidence that adaptive expectations are not of first order degree as argued in Marimon and Sunder (1995). Arifovic and Sargent (2003) also address the issue of inflation expectations formation and study the adaptive hypothesis on individual responses. They focus on the time inconsistency problem, asserting that in many cases policy makers achieve time-inconsistent optimal inflation rate, although in some treatments the economy moves towards sub-optimal (Nash) time consistent outcomes. They find support of adaptiveness and some evidence of heterogeneity of forecasts. Similarly, Fehr and Tyran (2008) suggest that expectations of individuals are heterogeneous. They study the adjustments of nominal prices after the anticipated monetary shock. "Learning to forecast" experiments are also conducted within the asset pricing framework characterized, as in our case, by positive feedback (see e.g. Hommes et al., 2005b and Haruvy et al., 2007).[4] These studies conclude that most subjects (90%) use simple rules to forecast prices looking at one, two or, at most, three lags of prices.

The baseline experiment described below is repeated under different monetary policy regimes to assess how alternative conducts of monetary policy influence the expectation formation process and the degree of heterogeneity. Monetary policy is modeled using different Taylor-type rules that are commonly used in the literature. Their effectiveness is then compared in terms of variability of inflation and inflation forecasts. We explore how different monetary policy settings anchor inflation expectations. We find that the variability of inflation is significantly affected by the degree of aggressiveness of monetary policy. Our results also suggest that instrumental rules responding to contemporaneous inflation perform better than rules responding to inflation expectations. Furthermore, the

[3]See Duffy (2008) for a survey on experimental macroeconomics. Most studies have been conducted in OLG economies with seignorage. Thus our framework is most closely related to the framework of Adam (2007) who studies the expectation formation process in a monetary sticky price environment.

[4]See also Hommes (2011) for a short survey.

design of monetary policy significantly affects the degree of heterogeneity – especially the proportion of trend extrapolation rules – and thus the stability of the main macroeconomic variables. The proportion of trend extrapolation rules increases in an environment characterized by excessive inflation variability and expectational cycles and then further amplifies the cycles. Thus, it is imperative to design a monetary policy that is robust to different expectation formation mechanisms.

Marimon and Sunder (1995) compare different monetary rules in the overlapping generations (OLG) framework to see their influence on the stability of the inflation expectations. In particular, they focus on the comparison between Friedman's k-percent money rule and the deficit rule where the government is fixing the real deficit and finance it through the seigniorage. They provide some evidence in support of Friedman's rule which could help to coordinate agents beliefs and help to stabilize the economy. Similar analysis is also performed in Bernasconi and Kirchkamp (2000). The latter argues that Friedman's money growth rule produces less inflation volatility, but higher average inflation compared to constant real deficit rule.[5]

Closer to our framework is the experiment by Adam (2007). He conducts experiments in a sticky price environment where inflation and output depend on expected inflation and analyzes the resulting cyclical patterns of inflation around its steady state. These cycles exhibit significant persistence and he argues that they closely resemble an restricted perception equilibrium[6] where subjects make forecasts with simple underparameterized rules. In our experiment we also detect cyclical behavior of inflation and output gap in some treatments, however we show that these phenomena are not only associated with underparameterization but also with heterogeneity of expectations, the design of monetary policy and (its influence on) the degree of backward-looking behavior.

This chapter is organized as follows: Section 1.2 describes the model for experimental analysis. Section 1.3 outlines the experimental design. In Section 1.4 we focus on the analysis of individual responses while in Section 1.5 we analyze switching dynamics between different models. Section 1.6 links the results to the monetary policy design and Section 1.7 concludes.

[5]The effects of monetary policy design on expectations were also examined in Hazelett and Kernen (2002) were they search for hyperinflationary paths in the laboratory.

[6]Restricted perception equilibrium is generally more volatile than rational expectation equilibrium (for more details see Evans and Honkapohja, 2001).

1.2 Model

In our experiment we use a forward-looking sticky price New Keynesian (NK) monetary model with different monetary policy reaction functions. The advantage of the NK model is that it is widely used in policy analysis and allows us to compare our experimental results with those obtained theoretically. However, there are two implicit complications for the participants. First, it requires them to forecast two periods ahead. It would definitely be easier for participants to produce a one period ahead forecast (sometimes called "nowcasting") as they would observe the realizations immediately after their forecasts are made. This would also enable us to simplify the analysis of individual responses, especially in the case of adaptive learning. The second complication is that forward-looking NK models assume that agents have to forecast both inflation and the output gap. We were afraid that this would represent an overly difficult task for the subjects. This was a considerably more difficult decision to make as asking the participants to only forecast inflation meant departing from the standard macro model. Nevertheless, we decided not to compromise the experiment by complicating the task for the subjects. We leave the fully forward-looking NK model for future work.

The baseline framework in the NK approach is a dynamic stochastic general equilibrium model with money, nominal price rigidities, and rational expectations (RE). Lately, some authors have augmented this model for adaptive learning and also for heterogeneous expectations (e.g. Branch and McGough, 2009). The model consists of a forward-looking Phillips curve (PC), an IS curve, and a monetary policy reaction function.[7]

In Chapters 1 and 2 we decided to focus on the reduced form of the NK model, where we can clearly elicit 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 the microfoundations. However forecasts are difficult to elicit in this environment, where subjects act as producers and consumers and interact on both the labor and final product markets and do not explicitly provide their forecasts (for this approach, see Noussair et al., 2011). Therefore, an appropriate framework for the question that we address in Chapter 1, is the "learning to forecast" design where incentives are set in order to induce forecasts as accurate as possible. In this framework, we do not assign subjects a particular role in the economy; they rather act as "professional" forecasters. One way to think about the relationship between "professional forecasters" and consumers/firms is that these economic subjects employ professional forecasters to pro-

[7] Detailed derivations can be found in, e.g., Woodford (1996), or textbooks such as Walsh (2003) or Woodford (2003).

vide them with forecasts of inflation.

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 $E_t\pi_{t+1}$, although strictly speaking it should be denoted as $E_t\left(\pi_{t+1}|\mathscr{I}_{t-1}\right)$. In fact, E_t (a forecast made in period t with the information set $t-1$) might not be restricted to just rational expectations.

The IS curve is specified as follows:

$$y_t = -\varphi\left(i_t - E_t\pi_{t+1}\right) + y_{t-1} + g_t, \tag{1.1}$$

where the interest rate is i_t, π_t denotes inflation, y_t is output gap, and g_t is an exogenous shock. The parameter φ is the intertemporal elasticity of substitution in demand. It will be noted that we do not include expectations of the output gap in the specification. Instead, we have a lagged output gap. In principle, one could argue that this specification of the IS equation corresponds to the case when agents have naive expectations about the output gap. The main reason for including a lagged output gap in our specification is that we want another endogenous variable to influence the law of motion for inflation. Even in the case when agents have rational expectations they have to use the observed information on the output gap to forecast inflation as this enters into the perceived law of motion of the rational expectations form. Compared to purely forward-looking specifications, our model might display more persistence in the output gap. This is our most significant departure from an otherwise standard macroeconomic model.

Aggregating across the price setting decisions of individual firms yields the linear relationship in the equation (1.2). Thus, the supply side of the economy is summarized in the following PC:

$$\pi_t = \beta E_t\pi_{t+1} + \lambda y_t + u_t. \tag{1.2}$$

On average, the longer prices are fixed, i.e. the smaller λ is, the less sensitive inflation is to the current output gap. The parameter β is the subjective discount rate. The shocks g_t and u_t are unobservable to subjects and follow the following process:

$$\begin{aligned}\begin{bmatrix} g_t \\ u_t \end{bmatrix} &= \Omega \begin{bmatrix} g_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} \widetilde{g}_t \\ \widetilde{u}_t \end{bmatrix}; \\ \Omega &= \begin{bmatrix} \kappa & 0 \\ 0 & \nu \end{bmatrix},\end{aligned}$$

where $0 < |\kappa| < 1$ and $0 < |\nu| < 1$. $\widetilde{g}_t$ and $\widetilde{u}_t$ are independent white noises, $\widetilde{g}_t \backsim N\left(0, \sigma_g^2\right)$

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 seen as a government spending shock or a taste shock and the standard interpretation of u_t is a technology shock. All these shocks are found to be quite persistent in the empirical literature (see e.g. Cooley and Prescott, 1995 or Ireland, 2004). In the experimental context it is important to have some exogenous unobservable component in the law of motion for endogenous variables, so we avoid the extreme case where all agents coordinate on the forecasts identical to the inflation target. Without AR(1) shock this would represent the dominant strategy. This is especially relevant concern as we initialize the model in a rational expectations equilibrium (REE).

1.2.1 Monetary policy reaction functions

To close the model, we have to specify the interest rate rule.[8] We use two alternative Taylor-type rules in different treatments. Most of our attention is devoted to forward-looking reaction functions: inflation forecast targeting where the interest rate is set in response to inflation expectations. We study three parametrizations of this rule and investigate how different degrees of central bank's aggressiveness in stabilizing inflation influence inflation expectations. Successively, we ask whether it is better for the central bank to respond to current or expected inflation, and therefore we also analyze the pure inflation targeting.

We start with the following interest rate rule (*Inflation Forecast Targeting*):

$$i_t = \gamma(E_t\pi_{t+1} - \overline{\pi}) + \overline{\pi}. \tag{1.3}$$

In this version, the central bank responds to deviations in inflation from the target, $\overline{\pi}$. It is implicitly assumed that the central bank observes the average prediction of subjects. We vary γ in different treatments and study the stability of the system under alternative reaction coefficients attached to inflation.

The second alternative specification is *Inflation Targeting*, where the monetary authority is assumed to respond to deviations in contemporaneous inflation from the inflation target:

$$i_t = \gamma(\pi_t - \overline{\pi}) + \overline{\pi}. \tag{1.4}$$

[8] Engle-Warnick and Turdaliev (2010) conduct experiments on monetary policy rules. Their subjects are only told to act as policymakers and to stabilize inflation. Most of the subjects control inflation relatively well and the authors argue that Taylor rules provide a good description of the subjects' policy decisions.

1.2.2 Calibration

We use the McCallum and Nelson (2004) calibration. This is a standard calibration for NK models. In order to have positive inflation for most of the periods we set the inflation target to $\overline{\pi} = 3$. A summary of the calibration is reported in Table 1.1.

$\beta = 0.99$	$\overline{\pi} = 3$	$\nu = 0.6$
$\varphi = 0.164$	$\lambda = 0.3$	$\kappa = 0.6$

Table 1.1: McCallum-Nelson Calibration

The different treatments are fully comparable as they all have exactly the same shocks. In particular, κ and ν are calibrated to 0.6, while their standard deviations are 0.08.

1.3 Experiment

1.3.1 Design

The experimental subjects participated in a simulated economy with 9 agents.[9] Each session of a treatment has 2 independent groups ("economies"), and therefore 18 subjects participate in each session. All the participants were recruited through recruitment programs for undergraduate students at the Universitat Pompeu Fabra and the University of Tilburg. Invitations to apply were sent to all of the approximately 1300 students in a database at Pompeu Fabra and to about 1200 students at Tilburg, except to those that had already participated in one of our sessions before. There are 70 periods in each treatment. We scaled the length of each decision sequence and the number of repetitions in such a way that each session lasted approximately 90 to 100 minutes, including the time for reading the instructions and 5 trial periods at the beginning. On average, the participants earned around €15, depending on the treatment and individual performance. The program was written in Z-Tree experimental software (Fischbacher, 2007).

The subjects are presented with a simple fictitious economy setup. As shown above, the economy is described with three macroeconomic variables: inflation, the output gap and

[9] Most of the learning to forecast experiments are conducted with 5-6 subjects, e.g. Hommes et al. (2005b), Adam (2007), Fehr and Tyran (2008).

the interest rate. The participants observe time series of these variables in a table up to period $t-1$. 10 initial values (periods $-9,\ldots,0$) are generated by the computer under the assumption of rational expectations. The subjects' task is to provide inflation forecasts for the period $t+1$. The underlying model of the economy is qualitatively described to them. We explain the meaning of and the relationship between the main macroeconomic variables and inform them that their decisions have an impact on the realized output, inflation and interest rate at time t. Omitting details of the underlying model is a common strategy in the learning to forecast experiments (see Duffy, 2008, and Hommes, 2011). 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.[10]

In every period t, subjects have to decide about: *i)* a prediction of the $t+1$ period inflation; and *ii)* the 95% confidence interval of their inflation prediction. The way how they report the confidence interval depends on the treatment and is described in the section below.

After each period, the subjects receive information about the realized inflation in that period, their prediction of it, and the payoff they have gained. The subjects' payoffs depend on the accuracy of their predictions. The accuracy benchmark is the actual inflation rate computed from the underlying model on the basis of the predictions made by all the agents in the economy. We replace $E_t\pi_{t+1}$ in equations (1.1), (1.2), and (1.3) by $\frac{1}{K}\sum^k \pi^k_{t+1|t}$, where $\pi^k_{t+1|t}$ is subject k's point forecast of inflation (K is the total number of subjects in an economy). Mean, rather than median forecast of subjects is used as an input to the model since it better complies with the assumptions of the underlying New Keynesian model. In the subsequent rounds the subjects are also informed about their past forecasts. They do not observe the forecasts of other individuals or their performance.

[10] Kelley and Friedman (2008) provide a survey of experiments that support the theoretical result above. Examples of learning to forecast experiments are e.g., Adam (2007) and Hommes et al. (2005b).

The payoff function, W, is the sum of two convex components as described below:

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

The first, W_1, depends on the subjects' forecast errors and is designed to encourage them to give accurate predictions. It gives subjects a payoff if their forecast errors, f, are smaller than 4. The second, W_2, depends on the width of their confidence interval and is intended to motivate subjects to think about the variance of actual inflation since it is more rewarding when it is narrower. CI is either equal to their point estimate of confidence interval or half of the difference between the upper and the lower bound. The subjects receive a reward if their confidence intervals, CI, are no greater than ± 4 percentage points, conditional on the fact that actual inflation falls within the given interval: $CI \geq \left|\pi_t - \pi^k_{t+1|t}\right|$. With this setup we restrict payoffs to positive values. Compared to more standard quadratic payoff functions, ours gives a greater reward for more accurate predictions and provides an incentive to think also about small variations in inflation, which may be important. As this experiment can potentially produce quite different variations in inflation between different sessions it is important to keep the incentive scheme quite steep. The payoff function is non-linear. Therefore, we accompanied it with a generous explanation and a payoff matrix on a separate sheet of paper to make sure all the participants understood the incentives. A similar approach is used in Adam (2007).

The participants received detailed instructions before the experiment started. They can be found in appendices A and B. To ensure understanding of the task, we read instructions out loud and presented the task descriptively along with examples. The subjects also filled in a short questionnaire after they had read the instructions and answered the questions about the procedures to make sure that they understood them.

1.3.2 Treatments

There are eight treatments in our experiment. Each treatment can be denoted as Xp, where p represents one of the four monetary policy regimes and X represents one of the two types of the confidence interval input. Subjects thus report their forecasting confidence in two ways: in the first, call it A, subjects report the interval as the number of percentage points within which their prediction is accurate. In the second, call it B, subjects are simply asked for the lower and the upper bound of their inflation prediction interval. For each p, there are 4 independent groups for A, and two independent groups for B.

Since only the point forecast of each subject is used for calculating $E_t\pi_{t+1}$ in our model, confidence intervals are irrelevant for the outcome of the model. Our analysis in Chapter 1 will therefore treat the treatments with the same policy specification as equal. For convenience reasons we will refer in Chapter 1 to treatments Ap and Bp as "treatment p". The analysis of confidence intervals and comparison of treatments Ap to Bp is performed in Chapter 2 where we concentrate specifically on forecaster's uncertainty.

Each of the four monetary policy regimes is defined by a different monetary policy reaction function. There are 6 independent groups for each policy.

Treatment	Groups	Taylor rule (equation)	Parameters
Inflation forecast targeting (1)	1-6	Forward looking (1.3)	$\gamma = 1.5$
Inflation forecast targeting (2)	7-12	Forward looking (1.3)	$\gamma = 1.35$
Inflation forecast targeting (3)	13-18	Forward looking (1.3)	$\gamma = 4$
Inflation targeting (4)	19-24	Contemporaneous (1.4)	$\gamma = 1.5$

Table 1.2: Treatments

The first three treatments, which are shown in Table 1.2, deal with the parametrization of the inflation forecast targeting given in equation (1.3). In this setup, the coefficient γ determines central bank's aggressiveness in response to deviations of inflation from its target. It is also believed that the higher the γ is, the stronger the stabilizing effect of the monetary policy rule is. It is of key interest to see how subjects react to more and less aggressive interest rate policies. Moreover, in a controlled environment we test whether different slope coefficients indeed have the expected stabilization effect.

The majority of empirical findings agree that the magnitude of the slope coefficient is around 1.5. Generally, when $\gamma > 1$ the interest rate rule is E-stable and produces a determinate outcome[11] (Taylor principle), while when $\gamma \leq 1$ it is E-unstable and indeterminate. When the Taylor principle holds all our treatments yield determinate and E-stable REE.

[11]E-stability is asymptotic stability of an REE under least sqares learning. By determinacy we mean the existence of a unique dynamically-stable REE. For more detailed definition see Evans and Honkapohja (2001).

Initially, we planned to perform a treatment with $\gamma < 1$ to check whether this leads to instability, but findings from the pilot treatments convinced us this was not a suitable choice as subjects quickly reached extremely high levels of inflation, which never returned to the target level. In this case the effect of output gap on inflation never outweighs the expected inflation effect. Such explosive behavior of the system suggest that the Taylor principle holds.

For our first and benchmark treatment we decided to follow Taylor and chose $\gamma = 1.5$. The average behavior of groups in this treatment showed no convergence to target inflation, so we choose $\gamma = 1.35$ as a sufficiently different case for a comparison. Alternatively, we chose $\gamma = 4$ as a parametrization with a high stabilizing effect where convergence to the target inflation should be faster.

In treatment 4 we focused on what measure of inflation central banks should target: that expected by the subjects or actual inflation. We performed a treatment using an inflation targeting rule where the central bank reacts to current inflation, with $\gamma = 1.5$, as in our benchmark case.

1.3.3 Summary of results

Below we present a brief summary of our experimental results. In 4 treatments of our experiment and 24 independent groups we gathered $15,120$ point forecasts of inflation from 216 subjects. The mean inflation forecast for all treatments is around 3.06% and the mean inflation is 3.02% when the inflation target is set at 3%. The standard deviations of inflation and inflation expectations vary substantially across groups. For inflation expectations the largest is 6.31 and the lowest 0.23 while for inflation the largest is 5.83 and the smallest is 0.24. Standard deviations of inflation forecasts are usually higher than standard deviations of inflation for groups with higher volatility, while for groups with lower volatility this might not necessarily be the case. Table 1.3 displays summary statistics for each group, while Figure A.1 in Appendix A displays the distribution of inflation forecasts in each treatment.

Inflation paths of all independent groups in each treatment are shown in Figure 1.1. One can observe that there are substantial differences in variability between different groups in each treatment and also that treatments 3 and 4 resulted in less volatile inflation on average. These differences may be due to treatment effects, different learning models used by subjects, group effects based on the learning models used, or some arbitrary subject effects. In group 2 of treatment 1, for example, one subject mistakenly inserted value 52 instead of 5.2, which resulted in sudden increase in inflation. In later sessions

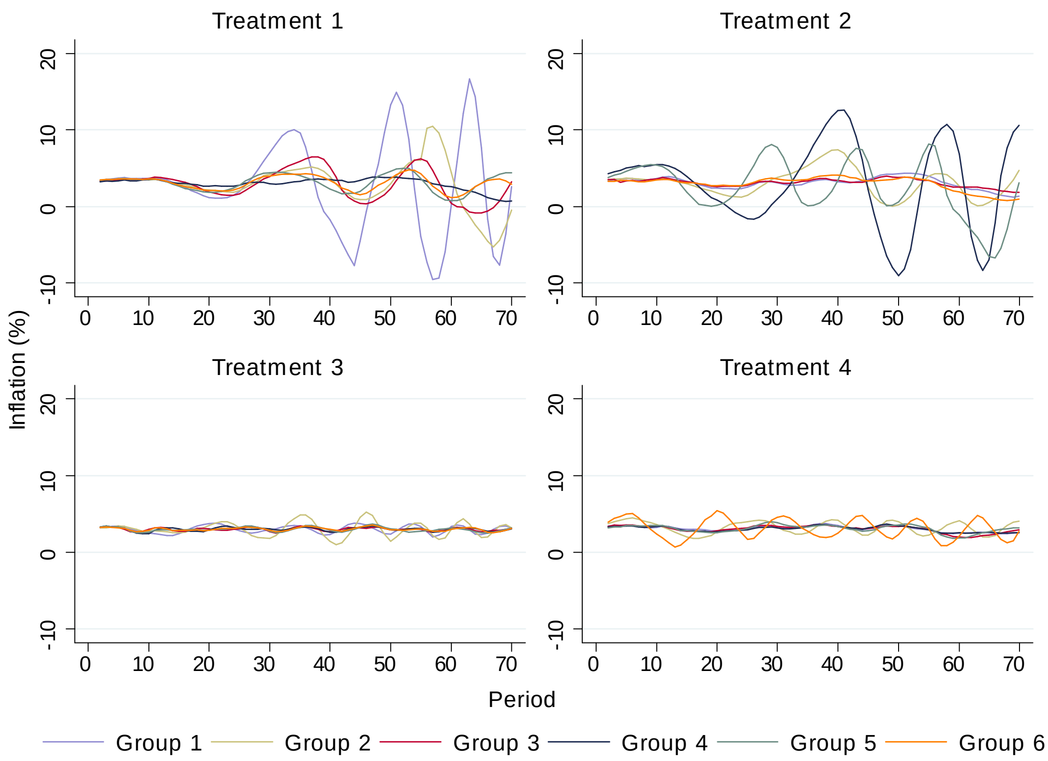

Figure 1.1: Group comparison of realized inflation by treatment. Treatment 1 has inflation forecast targeting (IFT) with $\gamma = 1.5$. Treatment 2 has IFT with $\gamma = 1.35$. Treatment 3 has IFT with $\gamma = 4$. Treatment 4 has inflation targeting with $\gamma = 1.5$.

we prevented that kind of mistakes by implementing a warning window when a value sufficiently differed from the previous input. Individual learning effects are presented in Section 1.4 while treatment effects are discussed in Section 1.6.

In Figure 1.2 it is possible to distinguish signs of a rounding effect (or digit preference).[12] This is especially evident for the responses below 0 and above 6, where we can observe a clear pattern that resembles rounding: the frequency of responses are significantly higher for round numbers than for the neighboring decimal numbers. A closer inspection reveals that rounding is also present for the responses between 0 and 6, only that rounding here does not take place only for responses such as 2, 3 and 4, but also for 2.5 and 3.5. This is due to the fact that in treatments where variability is lower subjects round on the basis of a smaller grid. Overall, we notice that 72% of all responses are reported to within one decimal point accuracy, while 13% of them are to an accuracy of 2 decimal points. The remaining 15% of forecasts are rounded as integers. The overall share of the latter is significantly higher for the groups with higher volatility compared to the groups displaying lower volatility.

[12]The full range of the responses reported is between -13.9 and 24. However we restrict this histogram to responses between -3 and 10.

Statistics by groups	Treatment 1						Treatment 2						Treatment 3						Treatment 4						
	Inflation forecast targeting, $\gamma = 1.5$						Inflation forecast targeting, $\gamma = 1.35$						Inflation forecast targeting, $\gamma = 4.0$						Inflation targeting, $\gamma = 1.5$						
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	Avg
Inflation Expectations																									
Mean	2.94	3.00	3.04	3.01	3.12	3.14	3.11	3.09	3.12	3.18	2.72	3.04	3.02	3.03	3.01	3.00	3.00	3.00	3.12	3.29	3.07	3.05	3.10	3.15	3.06
StdDev	6.31	3.48	2.02	0.73	1.12	0.93	0.75	1.87	0.49	5.77	3.76	0.86	0.57	1.07	0.26	0.30	0.30	0.23	0.38	0.89	0.48	0.36	0.54	1.42	2.20
Min	-13.9	-6.10	-2.50	0.40	0.30	0.50	1.00	-0.70	0.20	-12.0	-8.80	0.52	1.70	0.00	2.00	1.20	2.10	2.40	2.30	1.00	1.60	2.30	0.50	0.00	-13.9
Max	24.0	52.0	7.50	3.98	5.40	5.20	4.50	9.50	4.20	16.1	10.5	4.50	4.76	6.90	3.80	4.50	4.00	3.70	4.20	5.20	4.00	3.90	4.40	7.00	52.0
Inflation																									
Mean	2.85	2.88	2.92	3.00	3.13	3.12	3.12	3.09	3.13	3.02	2.52	3.03	3.01	3.02	2.99	3.00	2.99	3.01	3.09	3.23	3.05	3.05	3.09	3.11	3.02
StdDev	5.83	2.89	1.95	0.75	1.09	0.90	0.76	1.81	0.51	5.50	3.56	0.88	0.51	0.94	0.24	0.26	0.31	0.24	0.39	0.81	0.48	0.38	0.51	1.28	1.37
Min	-9.53	-5.27	-0.84	0.67	0.81	1.16	1.26	0.06	1.84	-9.04	-6.74	0.80	2.00	0.97	2.49	2.41	2.48	2.51	2.40	1.77	1.88	2.46	1.77	0.68	-9.53
Max	16.7	10.5	6.51	3.89	5.01	4.78	4.38	7.42	3.98	12.6	8.17	4.13	3.84	5.28	3.44	3.46	3.74	3.56	3.78	4.49	3.62	3.70	4.00	5.46	16.7

Table 1.3: Preliminary statistics

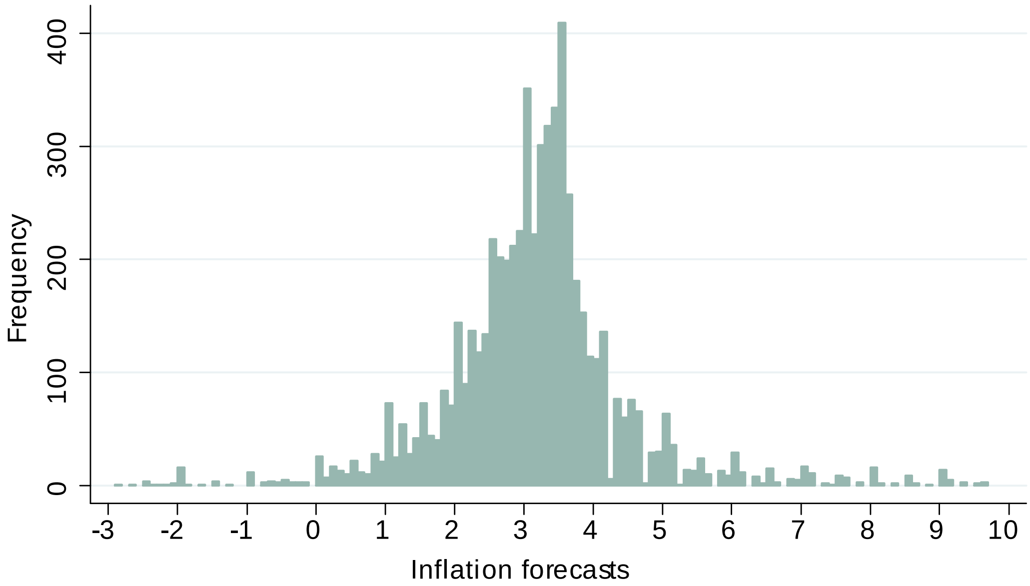

Figure 1.2: Histogram of inflation forecasts for all treatments.

However, it should be noted that survey data usually display more rounding, particularly the Michigan survey (see Curtin, 2005, Bryan and Palmqvist, 2005). Subjects in experiments are paid according to their performance and thus the accuracy of their forecasts always matters. On the contrary, in survey data we can observe the effect of inattentiveness[13] when inflation is low and stable. In this environment it can be said that the forecast accuracy is relatively less important than in the periods when inflation is more volatile and higher. The mean of forecast errors in our experiment is 0.04 and the standard deviation is 1.23. Thus, there is only a slight positive bias of errors. Furthermore, our subjects overpredict in 51.2% of cases and underpredict in 48.8%. A detailed analysis of the confidence intervals is presented in Chapter 2.

1.4 Analysis of individual inflation forecasts

The analysis of individual responses focuses in the first part on learning dynamics. Several learning models are simulated in order to find the best fit of each individual series on expectations. We also estimate other standard models of expectation formation including common rationality tests. All these models are estimated for each individual using OLS. Reported results are with robust standard errors that, where appropriate, take into account the presence of clusters in groups (or treatments). Below we present each of these tests and briefly comment estimation for all subjects while in the discussion we determine the

[13] Inattentiveness was first discussed by Mankiw and Reis (2002).

best-performing model for each subject. In the next section we dig deeper and investigate potential switching of subjects between different models.

1.4.1 Tests of rational expectations

Several econometric tests are designed to check the rationality of forecasts. In this subsection we apply some standard tests commonly employed in the survey data literature.[14] We assess different degrees of forecast efficiency and check whether forecasts yield predictable errors. The simplest test of efficiency is a test of bias:

$$\pi_{t+1} - \pi^k_{t+1|t} = \alpha, \tag{1.5}$$

where π_{t+1} is inflation at time $t+1$ and $\pi^k_{t+1|t}$ is k^{th} subject's inflation expectations for time $t+1$ made at time t (with information set $t-1$). By regressing expectational errors on a constant we check whether inflation expectations are centred around the right value. Majority of agents produce unbiased estimates of inflation. Overall, only 7.9% of them produce biased estimates at a 5% significance level and only 4.6% at a 1% threshold. Most of them are from treatments 2 and 4.

The next regression represents a further test for rationality:

$$\pi_{t+1} = a + b\pi^k_{t+1|t}. \tag{1.6}$$

As in Mankiw et al. (2004) the last expression can be simply augmented to test whether information in forecasts are fully exploited:

$$\pi_{t+1} - \pi^k_{t+1|t} = a + (b-1)\,\pi^k_{t+1|t}, \tag{1.7}$$

where rationality implies jointly that $a = 0$ and $b = 1$. As in the test for bias, under the null of rationality these regressions are meant to have no predictive power. The latter model is a more strict test of rationality and is seldom fulfilled in the survey data literature. On the contrary, our results suggest that 28.7% of agents exploit all the available information at a 5% significance level and 42.1% of them when we decrease the threshold to 1%. Treatment 2 is associated with the highest proportion of rational agents (48% and 57%, accordingly). Compared to other experimental studies, these tests suggest that a significant proportion of subjects behave rationally, although in asset pricing experiments Heemeijer et al. (2009) find a significant proportion of fundamental traders. These can be

[14] See Pesaran (1987), Mankiw et al. (2004) and Bakhshi and Yates (1998) for a review of these methods.

associated with rational expectations. Also Roos and Luhan (2008) show that about 23% of subjects do not have biased price expectations.[15]

1.4.2 Sticky information type regression

In this section we estimate a simple weighted average regression similar in formulation to sticky information model by Carroll (003a) with adaptive expectations. In our framework we have forecasts derived under the assumption of rational expectations while Carroll (003a) implements professional forecaster predictions. We estimate the following equation:

$$\pi^k_{t+1|t} = \lambda_1 \pi^{RE}_{t+1|t} + (1-\lambda_1)\,\pi^k_{t|t-1}; \tag{1.8}$$

$$\pi^k_{t+1|t} = \lambda_1 \eta_0 + \lambda_1 \eta_1 y_{t-1} + (1-\lambda_1)\,\pi^k_{t|t-1}, \tag{1.9}$$

where $\pi^{RE}_{t+1|t}$ is a rational expectations prediction of inflation for period $t+1$ at period t. This type of model is important for forecasting, especially in our framework, where some agents are backward-looking and rational agents have to incorporate also the backward-looking behavior into their forecasts. Thus we estimate a model (1.9) that is stated in terms of observable variables with the restrictions on all coefficients, where η_0 and η_1 are REE coefficients. Our formulation is inherently different than the one by Carroll (003a, 003b) as epidemiological framework that he proposes is no longer valid in our setup where subjects in principle observe all relevant information.[16] About 97% of agents display a significantly positive λ_1, while the average λ_1 is 0.20. Groups in treatment 3 had the highest average λ_1 (0.37), while subjects in treatment 2 had the lowest (0.11). It is not straightforward to define rationality in our framework and thus the results can be challenged on these grounds. The definition used in this subsection corresponds to REE if all agents in the group form expectations rationally. Similar weighted average regressions are estimated in Heemeijer et al. (2009), where they replace RE prediction with the equilibrium price.

1.4.3 Trend extrapolation rule

We also evaluate simple trend extrapolation rules. These are pointed out as particularly important rules for expectation formation process in Hommes et al. (2005b). We specify

[15] In field experiments by Berlemann and Nelson (2005) similar rationality tests were conducted suggesting that most participants exploit all available information.

[16] He argues that news about inflation spreads slowly across agents and reaches only a fraction λ_1 of population in each period.

the following process:

$$\pi^k_{t+1|t} - \pi_{t-1} = \tau_0 + \tau_1 (\pi_{t-1} - \pi_{t-2}), \tag{1.10}$$

where we estimate τ_0 and τ_1. We find that constant is significant at 5% level in 28.7% of cases while the τ_1 is significant in 78.2% of cases at the same level. Most of the times τ_1 is between 0 and 1, but there are a few cases when τ_1 is significantly lower than 0 (6.9%) and for 15.3% of subjects it is significantly higher than 1. We refer to the latter rules as strong trend extrapolation. Hommes et al. (2005b) find that about 50% of subjects in their experiment behave consistently with the trend extrapolation rule.

1.4.4 Estimating simple learning rules

In order to test for adaptive behavior, we apply different learning rules to experimental data. For a discussion on learning rules and convergence to rational expectations see Evans and Honkapohja (2001). We first test learning on a model with constant gain updating (CGL), where subjects learn from their past observed errors. The model below is equivalent to the adaptive expectations formula:

$$\pi^k_{t+1|t} = \pi^k_{t-1|t-2} + \vartheta \left(\pi_{t-1} - \pi^k_{t-1|t-2} \right), \tag{1.11}$$

where ϑ is the constant gain parameter. Under this learning rule agents revise their expectations according to the last observed error. In the experiment subjects are asked to forecast inflation in the next period (hence they make their forecast for period $t+1$ at time t), therefore the revision regards their previous period's forecast ($t-1$), which is made at time $t-2$. Note that this rule corresponds to the second order adaptive scheme in Marimon et al. (1993). All participants have ϑ positive and significant at a 5 percent level. 13.4% of participants have a constant gain parameter significantly lower than 1, while 53.7% of them update their forecasts with an error correction term significantly greater than 1. This means that the latter agents possibly overreact to their past errors. Their prevalence might imply problems with dynamic stability in certain treatments.

Below we present a learning mechanism with decreasing gain parameter (DGL):

$$\pi^k_{t+1|t} = \pi^k_{t-1|t-2} + \frac{\iota}{t} \left(\pi_{t-1} - \pi^k_{t-1|t-2} \right). \tag{1.12}$$

If the estimated parameter (ι in this version) is significantly different from 0, we conclude that agents actually learn from their past mistakes with a decreasing gain over time. Our tests do not support the hypotheses that the coefficient decreases over time as the R^2 is always greater (for all subjects) for a constant gain model.

Several versions of these models are estimated in Arifovic and Sargent (2003), Hommes et al. (2005b), Marimon and Sunder (1995) and Bernasconi and Kirchkamp (2000). Hommes et al. (2005b) argue that some subjects (about 5%) behave consistently with this rule, while Marimon and Sunder (1995) and Bernasconi and Kirchkamp (2000) put forward that most subjects in their OLG experiments use either first or second order adaptive expectations.

Recursive representation of simple learning rules

The above specification mainly aims at testing whether data support the existence of adaptive behavior. In this subsection, as in the adaptive learning literature, we assume that subjects behave like econometricians, using all the available information at the time of the forecast. In the following specifications, we test whether agents update their coefficients with respect to the last observed error. We assume four different perceived laws of motion (PLM):

$$\begin{aligned} \pi^k_{t+1|t} &= \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-1} + \varepsilon_t. && (1.13)\\ \pi^k_{t+1|t} &= \phi_{0,t-1} + \phi_{1,t-1}y_{t-1} + \varepsilon_t. && (1.14)\\ \pi^k_{t+1|t} &= \phi_{0,t-1} + \phi_{1,t-1}\pi^k_{t|t-1} + \varepsilon_t. && (1.15)\\ \pi^k_{t+1|t} - \pi_{t-1} &= \phi_{0,t-1} + \phi_{1,t-1}\left(\pi_{t-1} - \pi_{t-2}\right) + \varepsilon_t. && (1.16) \end{aligned}$$

Note that equation (1.14) represents a PLM of the REE form and equation (1.16) a version of the trend extrapolation rule. When agents estimate their PLMs they exploit all the available information up to period $t-1$. As new data become available they update their estimates according to a stochastic gradient learning (see Evans et al., 2010) with a constant gain. Let X_t and $\widehat{\phi}_t$ be the following vectors: $X_t = \begin{pmatrix} 1 & \pi_t \end{pmatrix}$ and $\widehat{\phi}_t = \begin{pmatrix} \phi_{0,t} & \phi_{1,t} \end{pmatrix}'$. In this version of constant gain learning (CGL) agents update the coefficients according to the following rule:

$$\widehat{\phi}_t = \widehat{\phi}_{t-2} + \vartheta X'_{t-2}\left(\pi_t - X_{t-2}\widehat{\phi}_{t-2}\right). \quad (1.17)$$

The empirical approach consists of searching for the parameter ϑ that minimizes the sum of squared errors (SSE), i.e. $\left(\pi^s_{t+1|t} - \pi^k_{t+1|t}\right)^2$ (see Pfajfar and Santoro, 2010 for details). The implicit problem in this approach is that we have to assume the initial values for $\widehat{\phi}_t$ for 2 periods. Setting up the initial values is one of the main problems when we recursively estimate learning. This issue is extensively discussed in Carceles-Poveda and Giannitsarou (2007). Strictly speaking, this problem should not occur in our case since we simply try to replicate our time-series data as closely as possible. Thus, in the following recursive learning estimations, we design an exercise in order to search for the best combinations of the gain parameter and initial values to match each subjects' expectations as closely as possible. This strategy can also be considered as a testing procedure for the detection of the learning dynamics for each individual. If the gain is positive under this method of initialization, then the series should exhibit learning for all other initialization methods with higher (or equal) gain.

We find that 56.5% of participants learn according to the first setup with lagged inflation as in model (1.13). The gain parameter ϑ is in the range between 0.0001 and 0.1000, with a mean value of 0.02900 and the median is 0.01125. We also estimate adaptive learning with the PLMs consistent of the REE form and AR(1) form, however these models rarely outperform other models studied here. In the learning version of the trend extrapolation model (1.16) 31.5% of subjects have positive gains. The optimal gains are on average slightly higher than before as they range between 0.0003 and 0.7900 with a mean value of 0.0654 (the median is 0.0310).

This version of the PLM (1.16) often performs better than previous versions of learning in terms of SSE. Below we compare different models and find that this version of constant gain learning indeed best represents the behavior of a significant proportion of our subjects. For a comparison with other studies, we exclude from our sample all subjects for which learning does not represent the best model.[17] In this case, we find that the average gain of these subjects is 0.0447 with a standard deviation of 0.0537 (median is 0.0260). The standard deviation is quite high as there are a few very high values, but most of the gains fall in the range between 0.01 and 0.07.

There are only a few estimates of the gain coefficient in the literature. Orphanides and Williams (005a) suggest a gain between 0.01 – 0.04 and Milani (2007) estimates it at 0.0183, while Pfajfar and Santoro (2010) find smaller gains (around 0.00021 for a similar version of learning). Our results suggest slightly higher gains than most of the above papers, but our data might be more volatile than the actual US inflation.

1.4.5 "General" models of expectation formation

Simple learning rules do not capture all macroeconomic factors that can affect inflation forecasts. In this subsection we estimate some general models of expectation formation.

[17] We will consider Comparison 1 in the **??** and exclude model (1.14) as it is generally associated with extremely high values of gain parameter.

We specify the following regression:[18]

$$\pi^k_{t+1|t} = \alpha + \gamma\pi_{t-1} + \beta y_{t-1} + \mu i_{t-1} + \zeta\pi^k_{t|t-1} + \varepsilon_t. \tag{1.18}$$

We find that 81.9% of agents take into account inflation when making their predictions. About 56.0% of the subjects take interest rate into account, while 66.7% also regard their own forecast from the previous period. Under some restrictions this equation could represent the form of the RE solution of the model ($\zeta = 0$).

For a comparison we also estimate a simple AR(1) model:

$$\pi^k_{t+1|t} = \phi_0 + \phi_1\pi^k_{t|t-1} + \varepsilon_t. \tag{1.19}$$

Similar model was already estimated with recursive learning. Model with constant coefficients, in general, is not often used by subjects for forecasting inflation. The behavior of forecast error is investigated in depth in the next chapter, in Subsection 2.3.1.

1.4.6 "Classical econometrician" and rational expectations

Before we discuss the best performing model for each individual we ask ourselves how would a "classical" econometrician forecast inflation in this environment. We estimate a regression for each period in time using only the available information that is on the subjects' screens. Of course, a more "sophisticated econometrician" could do a better job. For example, exogenous shocks are not observable in our framework, but a better econometrician could design an unobserved components model to extract information about the autoregressive shocks and then use them in these regressions. In the RE paradigm shocks play a significant part in the formation of expectations. In some treatments it is possible to observe that at least some agents extract information about the shock in the PC and at least partly use this information when forecasting. This is especially evident in treatment 4.

Therefore, we estimate rolling regressions and make one-step-ahead forecasts. A similar approach is used by Branch (2004) in the survey data literature for proxying rational expectations. Branch uses a trivariate VAR model and estimates it recursively. In our case, because of degrees of freedom problems, we have to resort to a univariate model (1.18). For a comparison, we also recursively estimate the adaptive expectations model (1.11) and a version of the trend extrapolation rule without the restrictions on coefficients.

[18]The models in groups 19-24 do not have the interest rate as dependent variable as this would imply multicolinearity due to the design of monetary policy in our framework.

Sum of sqare errors	Treatment 1						Treatment 2						Treatment 3						Treatment 4						
	Inflation forecast targeting, $\gamma = 1.5$						Inflation forecast targeting, $\gamma = 1.35$						Inflation forecast targeting, $\gamma = 4.0$						Inflation targeting, $\gamma = 1.5$						
Group	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	Avg
Subjects min	524	112	29.6	3.92	9.87	8.83	4.89	22.8	3.54	133	155	4.26	7.47	40.9	2.48	3.56	3.77	2.87	2.72	10.9	2.21	2.12	3.55	25.9	-
Subjects max	2354	1812	83.0	14.2	27.2	32.6	37.5	76.4	14.4	908	474	12.7	30.8	80.6	10.3	15.1	7.12	4.97	4.84	59.5	8.70	5.70	16.3	120	-
Subjects mean	1050	352	59.8	6.07	18.4	16.7	10.0	40.8	6.32	522	219	6.02	15.4	61.0	5.38	5.85	5.73	4.28	3.46	24.2	3.78	3.42	6.67	55.4	-
Sticky info. (1.8)	2110	1317	270	33.8	61.2	40.4	38.1	268.1	14.4	1720	1017	40.4	11.5	32.3	3.12	4.04	5.39	3.25	8.92	77.7	14.8	9.88	16.4	222	305
Gen. mod. (1.18), ζ=0	881	355	67.8	5.84	15.7	13.8	6.75	59.3	5.83	451	315	6.70	8.40	16.5	2.37	3.21	3.63	2.23	3.77	34.0	3.30	3.59	5.87	60.2	97
Trend ext. (1.10)	558	184	27.0	5.73	9.59	9.31	7.77	23.9	5.83	260	158	7.00	7.81	18.6	2.08	2.46	2.55	1.96	3.55	12.2	3.69	3.29	4.77	27.3	56
General model (1.18)	755	310	54.2	6.15	15.2	13.2	6.89	49.1	4.82	445	246	5.67	6.67	13.6	2.52	2.44	3.07	1.99	2.59	22.4	2.85	3.76	6.48	$5.3e^8$	88
Adaptive exp. (1.11)	973	210	67.8	6.15	21.3	15.2	8.63	65.2	5.68	805	313	8.04	12.8	53.6	4.36	5.78	5.77	3.41	3.67	21.7	3.71	3.30	5.80	61.7	112

Table 1.4: Comparison between subjects and classical econometrician

In practice, this rule is equivalent to the AR(2) model. We evaluate the general model with and without the restriction: $\zeta = 0$. Then we compare these forecasts with the actual realizations and compute the SSE, which are presented in Table 1.4 for five competing models. Before starting the analysis, it is worth pointing out that in treatments where the variation in inflation is greater the mean SSE is also higher (the correlation coefficient is 0.91). In two thirds of our groups the trend extrapolation rule performs best. However, in more stable treatments the general model can outperform the trend extrapolation rule.

Table 1.4 gives us a benchmark for evaluating the subjects' prediction accuracies. It is noteworthy that the best performing subjects often outperform our classical econometrician (best performing model). This occurs in practically all the groups except those comprising treatment 3, where a high frequency of cycles is observed (see Figure 1.1). There are two possible explanations for this: first, some subjects might be at least weakly rational; and second, subjects might be switching between different expectation formation mechanisms. We start by investigating the first possibility and then in Section 1.5 we dig deeper regarding the second possible explanation.

There are two definitions of rationality: the statistical and the "economic" definition. The former is defined and discussed in Subsection 1.4.1, while the latter interpretation argues that expectations should be consistent with the underlying economic model. Strictly speaking, where all agents know the macroeconomic model and behave accordingly, we know exactly the form of RE and the actual coefficient values.[19] However, in our experiment subjects are not familiar with the underlying macroeconomic model, and they might reasonably believe that other subjects potentially do not use RE. They have to take this into account when producing inflation forecasts. Even more, if rational agents understand the informational content of the interest rate, especially in treatments 1-3, they could implement this information into their decisions. Thus, in the environment of heterogeneous forecasts the REE PLM may be of a different form than the REE PLM in the case of homogeneous forecasts. This issue is further discussed in Nunes (2009) and Molnár (2007) where it is conjectured that some proportion of agents use adaptive learning to forecast, while the remaining agents are rational. Nunes (2009) studies this problem in the context of a forward-looking NK model and shows how to solve the model under the assumption of heterogeneous expectations. Our case is slightly different as the information sets of individuals do not include other subjects' forecasts. These could only be observed indirectly through interest rate in treatments 1-3, however subjects do not know that the interest rate setting depends on their forecasts. Nevertheless, if some agents use the PLM with last observed inflation, and the rational agents are aware of that, then they have to include the last observed inflation in their PLMs as well.

[19] As can be seen below, this REE PLM model (1.14) never outperforms the other models.

As there is not possible to calculate RE as a benchmark in our heterogeneous environment we have two different possibilities: (i) to use the statistical definition of rational expectations mentioned above, or (ii) to estimate the ALM (actual law of motion) for inflation in each group and check whether the estimated coefficients of the corresponding PLM entail statistically different coefficients to the ones of ALM. The problem here is that it is not straightforward how to define the form of the ALM as discussed above. We assume that the ALM is of the following form:

$$\pi_{t+1} = \gamma_0 + \gamma_1 \pi_{t-1} + \gamma_2 \pi_{t-2} + \gamma_3 y_{t-1} + \gamma_4 i_{t-1} + \varepsilon_t, \tag{1.20}$$

and the corresponding correctly parameterized PLM is:

$$\pi^k_{t+1|t} = \beta_0 + \beta_1 \pi_{t-1} + \beta_2 \pi_{t-2} + \beta_3 y_{t-1} + \beta_4 i_{t-1} + \varepsilon_t. \tag{1.21}$$

In order to be able to claim that a subject has model consistent (or rational) expectations, the estimated coefficients in both regressions should not be statistically different. To test for this we estimate the following equation:

$$\pi_{t+1} - \pi^k_{t+1|t} = \mu_0 + \mu_1 \pi_{t-1} + \mu_2 \pi_{t-2} + \mu_3 y_{t-1} + \mu_4 i_{t-1} + \varepsilon_t, \tag{1.22}$$

where $\mu_i = \gamma_i - \beta_i$. For subject to forecast rationally none of the estimated coefficients in equation (1.22) should be statistically significant. In the discussion below, we compare these definitions of RE. Rationality is in this case "superimposed" as we classify all agents that satisfy the requirements as rational, irrespective of their expectation formation mechanism.

1.4.7 Discussion

In this section we determine which theoretical model best describes the behavior of each individual on average. We compare the SSE[20] of each individual for the 10 models of expectation formation described above. A subject is regarded as using the model which produces the lowest SSE between the model predictions and their actual predictions.

We compare 9 models of inflation expectation formation that best describe the behavior of at least 1 participant. Model (1.12) is not used as it is always outperformed by other models.

[20] Our results and conclusions are the same irrespectively of whether we use RMSE (root mean square error), R^2 or SSE as they are all monotonic transformations of each other.

model (eq.)	Comparison 1	2	3	4	5	6
Rational expectations: Stat (1.7)	28.7	42.1	-	-	-	-
Rational expectations: Theory (1.22)	-	-	40.7	44.9	-	-
AR(1) process (1.19)	0.5	0.5	0.5	0.5	0.5	0.5
Sticky information type (1.8)	6.5	5.6	4.2	3.2	10.2	6.5
Adaptive expectations (1.11)	7.4	5.1	4.2	4.2	11.6	9.3
Trend extrapolation (1.10)	30.1	25.5	28.2	26.9	36.6	26.9
Recursive - lagged inflation (1.13)	11.6	7.9	8.8	8.3	21.8	9.3
Recursive - REE (1.14)	2.8	2.3	2.8	1.9	4.2	1.4
Recursive - AR(1) process (1.15)	0.5	0.5	0.5	0.5	0.5	
Recursive - trend extrapolation (1.16)	12.0	10.6	10.2	9.7	14.8	12.0
General model (1.18), $\zeta = 0$	-	-	-	-	-	34.3

Table 1.5: Inflation expectation formation (percent of subjects)

In Table 1.5 we present 6 different comparisons using different definitions of the RE. In comparisons 1 and 2 we define the RE based on statistical properties while in comparisons 3 and 4 based on theory as outlined above in Subsection 1.4.6 in equation (1.22): in comparisons 1 and 3 at 5% significance level, while in comparisons 2 and 4 at 1% significance level. In comparison 5 we exclude the general model from the set of alternative models, while in comparison 6 we compare all the empirical models.

We can observe that the results are indeed quite similar across alternative definitions of RE, although the theoretical definition (comparisons 3 and 4) suggests a slightly higher proportion of rational subjects. One possible reason is that we estimate the model (1.22) under the assumption of common AR(1) errors, as the experiment design embeds unobserved AR(1) shocks. Without this assumption, comparisons 3 and 4 would imply 27.3% and 31.0% of the subjects are rational. Generally, there is evidence that in all the treatments about $30-45\%$ of the subjects are rational and about $25-35\%$ simply extrapolate trend. Around $5-10\%$ of the subjects employ adaptive expectations while the remaining $15-25\%$ mostly behave in accordance with new theories of expectation formation, adaptive learning and sticky information type models. As mentioned before, most of other papers in the experimental literature stress the importance of adaptive expectations.

We report results from the literature on heterogeneous expectations for illustration, although these results might not be fully comparable to ours. Expectation formation of prices is, for example, studied in the US beef market. Chavas (2000) estimates that 81.7% of agents are boundedly rational using simple univariate models to forecasts prices and

18.3% of agents are rational. Contrary to that Baak (1999) finds that the proportion of rational agents is higher, i.e. about two thirds of agents are rational, while others are boundedly rational. Branch (2004) presents the results for 3 competing models of expectation formation (VAR, adaptive, and naive) estimated based on Michigan survey data. He finds that about 48% of agents use a VAR predictor and 44% of agents behave adaptively, while 7% are naive.

The availability of information is probably the main reason why our results suggest a higher degree of rationality than some previous studies on the inflation expectation. We must bear in mind that the subjects in our experiment always have available historical series of all the relevant macroeconomic variables and their past predictions. In the real world not all the variables might be readily observable or the information cost of collecting them might play an important role. The other reason for the high degree of rationality is that we initialize the model under RE. All these increase the possibility of not rejecting the assumption of rationality.

We further study the degree of heterogeneity by analyzing each treatment separately. We present comparison 1 across all the treatments in Table A.1 in Appendix A, where we can observe that there is quite a lot of heterogeneity across treatments. We further discuss this in the next section, where we analyze switching between different rules.

1.5 Switching between different models

The aim of this section is to further investigate how subjects form expectations. Do they consistently use one model or do they switch between different models? We mentioned before that switching might be one of the explanations for better performance of some individuals compared to the "classical econometrician." There are some attempts in the literature to link the performance of forecasting rules to the share of agents using that rule. Models that explore this issue are generally labelled as rationally heterogeneous expectations models. Some examples of these models are Brock and Hommes (1997), Branch and McGough (2008) and Pfajfar (2008). Their main argument is that it is not always optimal from an utility maximization point of view to forecast rationally as this might entail some costs.

In this section we tackle the problem from a slightly different perspective, as we only have 9 subjects in each group. Their information sets are different as the subjects do not directly observe the past forecasts of other subjects. Thus, it is not possible to compare these different models of dynamic predictor selection in our setup. We focus instead on establishing some stylized facts about "unrestricted" switching on an individual basis. An alternative approach, where all agents have the same information set is investigated in Anufriev and Hommes (2008). They provide support for switching based on a version of the predictor dynamics analyzed in Hommes et al. (2005a) and show that in an asset pricing environment the model with switching between simple heuristic rules can replicate the main results of the Hommes et al. (2005b) experiment in terms of individual behavior and aggregate dynamics. This approach is also followed in Assenza et al. (2011), where the environment is more similar to ours. However, we proceed in this analysis somewhat differently as our results above postulate that several of the rules employed are based on personal information; i.e. subjects include their own past forecast (which is unobservable to others) in their forecasting rule. In essence, we look at the roots of the switching behavior, and we do not impose a particular switching mechanism.

1.5.1 Unrestricted switching

We start this analysis by determining the optimal model for each individual in each period with a recursive estimation of the models specified above. Our approach consists of recursively computing the SSE up to a period t and then comparing it with period t for each individual. This comparison is performed for all periods except for the first 4 periods. Therefore, we can determine which model best fits the actual forecasting series in each point in time and whether there is any switching observed among these models. As many models' predictions are very similar at least in some episodes, we assume that there is no switching if the model that performs best in the previous period is not outperformed in the current period by 0.1 percentage points in terms of forecasts accuracy or 0.01 in terms of SSE. The rationale behind this choice is that the majority of forecasts are reported to one decimal point accuracy and the subjects are not able to differentiate between these competing models. In Table 1.6 we report the relative shares of the usage of each forecasting model considered, along with descriptive statistics for inflation.

We can observe that higher proportion of all forecasts are made using one of the stochastic gradient learning algorithms. Depending on the treatment, in 23 to 45% of all cases agents use these algorithms to forecast. If we average this across groups, 36.7% of the forecast decisions are best explained with adaptive learning. This means that, on average, adaptive learning is the most popular way of forming beliefs.

In around 17% of cases subjects use the general model, and in about 12% of all forecast decisions they behave in accordance with the sticky information type model. The remaining third of all forecasts are best explained with some sort of backward-looking model.

by group	Treatment 1						Treatment 2						Treatment 3						Treatment 4						
	Inflation forecast targeting, $\gamma = 1.5$						Inflation forecast targeting, $\gamma = 1.35$						Inflation forecast targeting, $\gamma = 4.0$						Inflation targeting, $\gamma = 1.5$						
model (eq.)	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	Avg
Gen. mod. (1.18), ζ=0	17.7	18.5	24.6	16.2	14.8	11.4	13.3	14.5	14.3	18.5	19.5	16.8	12.1	16.3	27.9	18.4	10.3	26.4	11.4	15.7	13.6	22.2	12.8	19	16.9
Sticky info. type (1.8)	12.3	6.4	8.2	5.9	10.1	14.3	11.6	10.8	13.6	9.6	9.9	8.4	14.8	11.1	19.2	23.9	15.8	18.2	9.4	8.9	12.6	8.6	15.8	11.1	12.1
Ad. exp. CGL (1.11)	16	12.6	13	15.8	15	11.4	16.7	13.8	10.8	17.8	14.3	18.2	15.2	14.3	12.5	11.1	12.6	5.1	11.1	15.8	17.5	11.4	11.6	20.4	13.9
Ad. exp. DGL (1.12)	7.2	5.4	4.7	8.4	2.7	5.9	8.2	6.2	9.6	5.6	5.4	6.6	6.7	4.9	5.6	8.8	8.8	7.6	7.7	5.2	5.6	6.4	5.1	4	6.3
Trend extr. (1.10)	23.6	20	14.1	15	16.2	11.8	15.5	13.8	13.8	16.5	20.9	13.6	13.3	19.2	6.1	10.4	12.6	6.4	7.6	13.1	10.1	12.5	11.8	18.4	14
Rec. AR(1) (1.15)	7.9	12.1	11.1	19	19.2	12.8	19.2	9.4	20	8.4	8.8	18.9	11.1	10.8	12.5	13.5	22.2	18.7	26.1	16	29.1	19.7	19.7	7.1	15.6
Rec. trend extr. (1.16)	15.3	24.9	24.2	19.7	22.1	32.3	15.5	31.5	17.8	23.6	21.2	17.5	26.8	23.4	16.3	14	17.7	17.7	26.6	25.3	11.4	19.2	23.2	20	21.1
inflation min	-9.53	-5.27	-0.84	0.67	0.81	1.16	1.26	0.06	1.84	-9.04	-6.74	0.8	2	0.97	2.49	2.41	2.48	2.51	2.4	1.77	1.88	2.46	1.77	0.68	-9.53
inflation max	16.68	10.46	6.51	3.89	5.01	4.78	4.38	7.42	3.98	12.56	8.17	4.13	3.84	5.28	3.44	3.46	3.74	3.56	3.78	4.49	3.62	3.7	4	5.46	16.68
inflation s.d.	5.83	2.89	1.95	0.75	1.09	0.9	0.76	1.81	0.51	5.5	3.56	0.88	0.51	0.94	0.24	0.26	0.31	0.24	0.39	0.81	0.48	0.38	0.51	1.28	2.05

Table 1.6: Inflation expectation formation (percent of all cases)

	Gen. mod. (1.18), ζ=0		Sticky info. (1.8)		ADE CGL (1.11)		ADE DGL (1.12)		Trend ext. (1.10)		Rec. AR(1) (1.15)		Rec. trend. extr. (1.16)	
s.e.: method	robust	cluster	robust	cluster	robust	cluster	robust	cluster	robust	cluster	robust	cluster	robust	cluster
p_{js}	46.285	46.285	-12.2631***	-12.2631*	16.6381*	16.6381*	-18.7886	-187.886	25.2361***	25.2361**	-15.6069***	-15.6069	25.229	25.229
	(-5.675)	(-7.324)	(-4.354)	(-4.569)	(-8.920)	(-6.612)	(-12.449)	(-17.302)	(-6.227)	(-7.832)	(-5.392)	(-7.622)	(-5.545)	(-6.258)
cons	0.5816	0.5816	2.8509***	2.8509**	-0.9502	-0.9502	2.5567***	25.567	-2.1698***	-2.1698	3.7924***	3.7924*	0.8321	0.8321
	(-0.8462)	(-0.8648)	(-0.7462)	(-0.8232)	(-1.041)	(-0.6043)	(-0.8706)	(-1.268)	(-0.7605)	(-0.9535)	(-1.043)	(-1.504)	(-1.311)	(-1.586)
Observations	24		24		24		24		24		24		24	
R^2	0.0192		0.1115		0.1173		0.0407		0.4989		0.3525		0.0072	

Table 1.9: Relation of standard deviation to certain behavioral types as de ned in Table 1.6. Notes: Standard errors in parentheses. */**/*** denotes signicance at 10/5/1 percent. Standard errors take into account potential presence of clusters in treatments.

Specifically, in around 14% of cases subjects use simple trend extrapolation rules while in the remaining 20% of cases they behave in an adaptive manner. Compared to the results outlined above for "average" best model, we can immediately observe that there is approximately the same proportion of backward-looking cases as there are subjects that use backward-looking rules. However, when allowing for switching there are more forecast decisions made in an adaptive way. Also model (1.15) is only a predominant model for one subject, but when we allow for switching it is used on average in 15.6% of all forecasts.

Generally, we can observe that when we allow subjects to switch between different models, they are in fact using alternative models to forecast. Under this assumption, agents use between 1 and 7 different models (average number of models used for forecasting is 6.5) and they on average switch every 4 periods. However, switching occurs less frequently in treatments 3 and 4 compared to treatments 1 and 2 (significant at 5% level with different tests of equality of medians).[21] Only one subject did not switch between models. Overall, these results support the idea of intrinsic heterogeneity that is theoretically modelled in Branch and Evans (2006) and Pfajfar (2008).

To further analyze the degree of heterogeneity in the data, we compute the average number of models used in each period. We find that on average 4.5 different models (between 2 and 7) are used within a group in each period. This additionally supports the above conjecture that heterogeneity is pervasive as there are not significant differences across treatments. The average number of models employed for forecasting within a group only varies (in each period) between 4.2 and 5.3. Furthermore, there is no "smoothing" employed across different subjects in the same group. We have only employed some "smoothing" within a subject as some models perform quite similarly and cannot be differentiated at one decimal point accuracy.

We also investigate the pattern (timing) of switching with panel probit and logit models (with random, and fixed effects, and population averages), where dependant variable, z_t^k, is 1 when switching occurs and 0 otherwise. We estimate the following regression:

$$z_t^k = \alpha_1^k + \alpha_2 \pi_{t-1} + \alpha_3 y_{t-1} + \alpha_4 i_{t-1} + \alpha_5 \left(\pi_{t-1} - \pi_{t-1|t-2}^k \right)^2 + \varepsilon_t^k. \tag{1.23}$$

We find that subjects decide to switch according to developments of inflation, the output gap, and the interest rate. The different models exhibit similar effects of the explanatory variables. The most pronounced effect expectably comes from the output gap, which has a strong negative impact on the probability of switching. A positive change

[21] Using the Kruskal-Wallis test, we show that switching occurs on average every 6.1 periods in treatment 4, 3.7 period in treatment 3, 2.6 period in treatment 2, and 2.9 periods in treatment 1.

	Probit RE	Probit PA	Logit RE	Logit PA	Logit FE
Cons.	-0.2502***	-0.2210***	-0.4139***	-0.3552***	
	(0.0836)	(0.0749)	(0.1449)	(0.1188)	
$\Delta\pi_{t-1}$	0.0422	0.0402	0.0661	0.0639*	0.0545
	(0.0293)	(0.0247)	(0.0482)	(0.0388)	(0.0354)
π_{t-1}	-0.0568***	-0.0533***	-0.0919***	-0.0857***	-0.076**
	(0.0219)	(0.0190)	(0.0345)	(0.0302)	(0.0383)
y_{t-1}	-0.1702***	-0.1596***	-0.2747***	-0.2577***	-0.2540***
	(0.0391)	(0.0381)	(0.0674)	(0.0623)	(0.0591)
i_{t-1}	0.0440**	0.0415**	0.0715**	0.0670***	0.0575**
	(0.0181)	(0.0161)	(0.0286)	(0.0254)	(0.0275)
er^2	0.0061	0.006	0.011	0.0099	0.0089
	(0.0171)	(0.0143)	(0.0248)	(0.0260)	(0.0359)
$\ln(\sigma_u^2)$	-1.5874***		-0.5814***		
	(0.1996)		(0.2064)		
σ_u	0.4522***		0.7478***		
	(0.0441)		(0.0783)		
ρ	0.1670***		0.1453***		
	(0.0270)		(0.0256)		
N	14040	14040	14040	14040	13975
Groups	216	216	216	216	215
Obs. per Gr.	65	65	65	65	65
Wald $\chi^2(9)$	34.0	31.8	31.2	32.6	36.2

Table 1.7: Determinants of swithing behavior. Notes: RE stands for random effects, PA population averages, while FE is for fixed effects model. Standard errors in parentheses. */**/*** denotes significance at 10/5/1 percent level. Standard errors are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups.

in the inflation trend increases the probability of switching, although, higher inflation decreases it. This demonstrates that there exists a certain pattern in the structure of individual switching. There are also some differences across treatments. In particular in treatment 4 the pattern of switching is different. However, treatment dummies are insignificant if we insert them in the above regression. The results are reported in Table 1.7, where $\Delta\pi_{t-1} = |\pi_{t-1} - \pi_{t-2}|$, and $er^2 = \left(\pi_{t-1} - \pi^k_{t-1|t-2}\right)^2$.

1.6 Monetary policy in the presence of heterogeneous expectations

Woodford (2003) showed that within a standard NK model monetary policy should minimize the variance of inflation and the output gap as this corresponds to maximizing the utility of consumers. Therefore we start this section with an analysis of variance of inflation since a monetary authority only cares about inflation in instrument rules under scrutiny. Testing for differences in the medians across treatments, where the null hypothesis that the medians are the same in all treatments is rejected at 1% significance with Kruskal-Wallis and van der Waerden tests (see Conover, 1999). Therefore, we can argue that the design of monetary policy matters in our framework. The following table shows the comparison of median standard deviations of inflation in treatments 2, 3, 4 with treatment 1. We report p-values from the Kruskal-Wallis test in Table 1.8.[22]

Treatment	Groups	Comparison with Treatment 1 (p-value)
Inflation forc. targ. $\gamma = 1.5$	$1-6$	–
Inflation forc. targ. $\gamma = 1.35$	$7-12$	0.6310
Inflation forc. targ. $\gamma = 4$	$13-18$	0.0104
Inflation targeting $\gamma = 1.5$	$19-24$	0.0250

Table 1.8: Comparison of standard deviations using Kruskal-Wallis test

We also find that there is a significant difference between treatments 2 and 3 (p-value is 0.0250). Thus, we can argue that treatments 3 and 4 produce significantly lower inflation variability than treatments 1 and 2. Now that we establish that there is a difference in variance of inflation between treatments we further analyze the roots of these differences between and within treatments.

For an illustration how important expectations are for the stability of the system we simulate our treatments with different forecasting rules under the assumption of homogeneous expectations (see Figures A.4 and A.5 in Appendix A). We can immediately observe that adaptive expectations (with a gain coefficient higher than 1) and trend extrapolation rules can lead to pronounced cyclical variability in inflation. It is also possible to observe that treatments 2 and 4 perform better than 1 and 3 in stabilizing those expectation formation mechanisms. However, the evidence might be reversed with respect to "stable" expectation formation mechanisms.

[22]Other nonparametric tests perform very similarly.

The proportion of backward-looking (especially trend extrapolation) agents plays a particularly important role for the stability of the system. We can observe that there is a considerable degree of heterogeneity across treatments. Even more, differences in the degree of backward-looking subjects can explain the differences in variability between groups in the same treatment. The results are intuitive as we find that there is a strong correlation between the stability of the system and the degree of trend extrapolation behavior. We further test these conjectures regarding the relationship between the variability and proportion of different groups of subjects with cross-sectional and panel data regressions. With former we find that especially increasing proportion of trend extrapolation behavior is increasing the variance. Also increasing proportion of CGL adaptive expectations rules is increasing the variance as most of the estimated coefficients ϑ in equation (1.11) are higher than 1 while the proportion of recursive learning (1.15) and also sticky information rules (1.8) is reducing it. We estimate the following regressions:

$$sd_s = \eta_0 + \eta_1 p_{js} + \varepsilon_s,$$

where sd_s is standard deviation of group s, and p_{js} is proportion of agents using j-th model for forecasting in group s. The set of alternative models is the same as in Table 1.6 above. Regression results are reported in Table 1.9, both with robust and clustered standard errors. Initially, we added treatment dummies to the above regression, however they were insignificant in almost all cases. We have to point out that all estimated coefficients (that are significantly different than 0) have the expected signs.

These results are confirmed also with the system GMM estimator of Blundell and Bond (1998) for dynamic panels. To construct the panel we compute the $sd_{s,t}$, standard deviation from the first period up to period t. Using the switching analysis we similarly compute $p_{js,t}$, the share of model j in group s up to the period t. We estimate the following model:

$$sd_{s,t} = \eta_0 + \eta_L sd_{s,t-1} + \sum_j \eta_j p_{js,t} + \varepsilon_{st}.$$

Results are reported in Table Table 1.10. Different variants are tested, depending on the inclusion of models for recursive learning, adaptive expectations and general model. The only intriguing result is about the coefficient on the proportion of the general model (1.18) which is insignificant in the cross sectional regression and significantly positive or insignificant in dynamic panel data models. Therefore it is difficult to say from this analysis what is the effect of the proportion of usage of general model (1.18) to the stability of inflation. Although these agents use all relevant information to forecast inflation, simulation exercise shows that at low values of γ this forecasting model (if used exclusively) will result in high variability of inflation (see Figure A.6 in Appendix A). Furthermore, theoretical analysis shows that as soon as one uses past inflation to forecast the model exhibits indeterminacy, i.e. there might be a multiple equilibria problem.

$sd_{s,t}$:	(a)	(b)	(c)	(d)
$sd_{s,t-1}$	1.0147***	1.0121***	1.0121***	1.0099***
	(0.0085)	(0.0073)	(0.0069)	(0.0066)
Gen. mod. (1.18), $\zeta = 0$	0.0018***	0.001	0.0031*	
	(0.0007)	(0.0013)	(0.0017)	
Sticky info. (1.8)	-0.0029*	-0.0039	-0.0018	-0.0043**
	(0.0016)	(0.0025)	(0.0019)	(0.0020)
ADE DGL (1.12)	-0.0023**	-0.0030**	-0.0008	-0.0027**
	(0.0009)	(0.0013)	(0.0015)	(0.0014)
Trend Ext. (1.10)	0.0067***	0.0055***	0.0077***	0.0055***
	(0.0015)	(0.0018)	(0.0023)	(0.0014)
ADE CGL (1.11)		-0.0011	0.001	
		(0.0018)	(0.0015)	
Recursive V1 (1.13)		-0.0021		-0.0025
		(0.0025)		(0.0018)
Recursive V4 (1.16)			0.0021	
			(0.0025)	
cons	-0.0759*	0.0219	-0.1895	0.0373
	(0.0417)	(0.1378)	(0.1449)	(0.0556)
N	1560	1560	1560	1560
χ^2	67328.4	54449.2	65883.1	79094.9

Table 1.10: Decision model's influence on standard deviation of inflation. Notes: Estimations are conducted using system GMM estimator of Blundell and Bond (1998) for dynamic panels. Standard errors in parentheses. */**/*** denotes significance at 10/5/1 percent level. Standard errors are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in treatments.

The result regarding the influence of the proportion of trend extrapolation rules to the standard deviation of inflation is very robust across these different techniques as the coefficients are always very significant and positive. The proportion of these agents probably plays the most important role for the stability of inflation. It also helps us to explain the differences among groups within the same treatment. Generally, we note that the group with lower proportion of trend extrapolation rules is more stable compared to other groups in the same treatment.

Heemeijer et al. (2009) compare experimental results in positive and negative expectations feedback models.[23] In a positive expectations system, e.g. asset pricing model, they

[23] Also Fehr and Tyran (2008) compare the two environments, although in a different context.

observe similar aggregate behavior to ours and note that when there is a stronger positive feedback more agents resort to backward-looking, especially trend following rules. In our case, by changing the monetary policy, we augment the degree of positive feedback from inflation expectations to current inflation. Therefore, the design of monetary policy is important for the prevailing expectation formation mechanism and vice versa, as can be seen if we compare results within the same treatment. The graphical analysis of the evolution of inflation across treatments is reported in Figure 1.1.[24]

However, this is only a part of the story in our experiment. We expected that the treatment 2 where monetary authority does not react too strongly to inflation expectations ($\gamma = 1.35$) performs better regarding the stability of inflation than the benchmark treatment, although the theory under rational expectations suggests that higher γ leads to lower variability of inflation. This is not confirmed in our analysis above as the median standard deviation is not statistically different than in treatment 1. This might be due to expectations of cycles by some individuals in groups 4 and 5 of this treatment and extensive use of strong trend extrapolation rules at the beginning of the experiment. In order to study the relationship between γ and the variance of inflation under different expectation formation mechanisms we design simulation exercises that exactly replicate the design, parametrization and shocks employed in the experiment. When all subjects have rational expectations we confirm the theory that higher γ leads to lower variability of inflation while many other expectation formation mechanisms produce non-monotonic, often U-shaped behavior of the inflation variance. On one hand, rules that we labelled as stable in regressions above produce decreasing variability of inflation when increasing γ, although sometimes non-monotonic. On the other hand, especially trend extrapolation rules will lead to U-shaped behavior and eventually higher variability when increasing γ (see Figure A.6)**.** The minimum variability of inflation with sticky information and trend extrapolation rule is achieved at $\gamma = 1.1$. For naive expectations the minimum is around $\gamma = 3$ (non-monotonic U-shaped). This can be also observed from Figures A.4 and A.5 in Appendix A.

Therefore, the relationship between the variability of inflation and different rules is non-trivial and the question whether treatment 2 should produce lower variability compared to treatment 1 depends particularly on the proportions of alternative rules used. Based on simulation results and observed behavior of individuals we can argue that in the presence of heterogeneous expectations instrumental rules that are less aggressive have the potential to produce lower variability of inflation, however there is the risk that, e.g. after a shock, the amplitude of inflation increases significantly as monetary policy is not aggres-

[24]Detailed figures with the evolution of inflation and inflation forecasts in each treatment are reported in Appendix A (Figures A.2 and A.3).

sive enough. Thus, one could argue that non-linear Taylor-type rules would perform best in this environment, although the literature in monetary economics has not attached much attention to this type of instrumental rules.

As we have seen so far, the expectational feedback is not the only source of instability in our multivariate system where we also have lagged endogenous variables.[25] Treatment 3 produces lower variability of inflation compared to treatment 2, but in the former case the frequency of cycles is significantly higher as the monetary authority is (too) strongly responding to deviations from inflation target. After some threshold of response to inflation forecast (depends on the proportion of agents using each rule) the resulting amplitude of the inflation variability decreases, while the frequency of cycles increases. The latter makes it more difficult to forecast and more participants resort to simpler rules. Using simulations explained above we can identify two effects of increasing γ on the variability of inflation: (i) this always increases the frequency of cycles irrespective of the expectation formation mechanism and (ii) it is increases or decreases the amplitude of the cycle. The latter result depends on the expectation formation mechanism and can produce non-monotonic or even U-shaped responses of variability, except for rational expectations where it decreases monotonically (see Figure A.6).

Also treatment 4 performs better than the benchmark treatment. Responding to contemporaneous inflation (as in treatment 4) turns out to be a better practise for central banks compared to responding to inflation expectations.[26] Moreover, this treatment resembles quite closely the behavior of survey forecasts, as there are periods when subjects systematically overpredict inflation (low and stable inflation) and underpredict inflation (when inflation is high). This is evident in Figure A.3 in Appendix A. In this treatment there is the highest proportion of biased agents and also results from the general model suggest similar behavior of these agents to the results obtained in the survey data literature. Moreover, if we compare the means of inflation forecasts in treatments 1 and 4 we find that the mean of inflation forecasts of groups in treatment 4 is significantly higher than the mean of inflation forecasts of groups in treatment 1 (at 10% significance with Kruskal-Wallis test). Also average inflation in treatment 4 is higher (3.10 in treatment 4 compared to 3.00 in treatment 1), however the difference is statistically insignificant with nonparametric tests. Comparison between treatments 1 and 4 implies that significantly lower standard deviation of inflation (and inflation forecasts) for treatment 4 (see Table 1.8) comes at a "cost" of higher inflation expectations (and possibly inflation). This result is similar to Bernasconi and Kirchkamp (2000) as they suggest Friedman's money growth rule pro-

[25] Generally, as γ is increasing the positive feedback is decreasing.

[26] Pfajfar and Santoro (2008) and Muto (2011) reach similar conclusion in different versions of the NK model: Muto (2011) in case when agents learn from central banks' forecasts, while Pfajfar and Santoro (2008) when they introduce the cost channel and capital market imperfections.

duces less inflation volatility, but higher average inflation compared to constant real deficit rule.

We can also observe that generally the variability of inflation is lower than the variability of inflation expectations. This provides an explanation to the fact that responding to current inflation stabilizes the system in a more efficient way compared to reacting to expected inflation. Moreover, by reacting to current inflation we decrease the expectational feedback compared to responding to the expected inflation. As a result, in treatment 4 we reduce the size of the expectational cycles as in booms monetary policy overreacts less than in the case when interest rate is set to respond to expected inflation (in presence of backward-looking agents). At the root of this pattern is that backward-looking subjects do not observe the informational content of output gap and do not predict the change in the growth rate of inflation. They still expect that inflation will accelerate as in the last few periods. Then, if the monetary authority is reacting with respect to the expected inflation, they do not change the stance of monetary policy in time. The economy is pushed in the recession where the backward-looking agents underpredict inflation and the recession is more severe than if all agents were rational. The whole process repeats in the next cycle. We have to point out that the causality goes in both directions as the proportion of backward-looking agents (especially strong trend extrapolation agents) depends on the design of monetary policy (degree of aggressiveness) and also the stability of the economy is influenced by the degree of backward-looking agents.

Adam (2007) obtains similar dynamic pattern of inflation and inflation expectations, especially to our treatment 3. He argues that the cause for observed behavior is the subjects' reliance to simpler underparameterized rules for forecasting inflation. Thus, he characterizes the dynamics of inflation as a restricted perception equilibrium, as inflation exhibits excessive volatility around its REE. Our results support his findings as some agents do not take into account output gap when forecasting. However, we also show that the volatility of inflation depends on the way monetary policy is designed and conducted. We argue that the proportion of backward-looking subjects plays an important role, especially those that use strong trend extrapolation rule.[27]

[27] Also several asset pricing experiments have observed the dynamics of aggregate price exhibiting bubbles (see eg. Smith et al., 1988 and Hommes et al., 2005b). Even more, Lei et al. (2001) show that this can occur also in an environment where speculation is not possible. They conclude that this occurs due to systematic errors in decisions.

1.7 Conclusion

In Chapter 1 we preset a design of macroeconomic experiment where subjects are asked to forecast inflation. The underlying model of the economy is a simple NK model which is commonly used for the analysis of monetary policy. The focus of present chapter is on the formation of inflation expectations and monetary policy design. In different treatments we employ various modifications of the original Taylor rule and study the influence of alternative monetary policy designs to inflation formation process and also vice versa. Therefore, we also try to determine the design of monetary policy which would effectively stabilize and anchor the process of inflation expectations. It is clear that monetary policy influences the expectation formation process. We find that the variability of inflation is significantly lower in treatments 3 and 4 compared to treatments 1 and 2. The cyclical behavior of inflation is also studied in the experimental study by Adam (2007). When we set interest rate with respect to current inflation, we observe the dynamics of inflation expectations that most closely resembles the behavior of survey data. Generally, this setup performs better in terms of inflation variability than responding to the expected inflation as the variability of inflation is lower than the variability of inflation forecasts. Thus, we reduce the amplitude of expectational cycles.

However, we can point out that the underlying process of inflation expectation formation depends also on the way monetary policy is conducted. The proportion of backward-looking agents, especially trend extrapolating subjects, plays an important role, as in some environments it is more difficult to forecast inflation rationally. In these cases more subjects resort to simpler backward-looking rules. We find that roughly $30-35\%$ of subjects predominately use trend extrapolation rules and additionally $5-10\%$ of subjects use adaptive expectations. Contrary to previous studies, our results suggest that there is a significant and relatively large share of agents that predominately use rational expectations. The share of these agents is about $35-45\%$. The remaining agents use some version of adaptive learning or sticky information type models. Furthermore, we also find that most agents tend to switch between different rules. When we take into account this possibility, we get slightly different results. Most notably, adaptive learning models become more important as this mechanism for forecasting is used in 36.7% of all forecasting decisions. Chapter 1 is one of the first empirical contributions to postulate that these models represent one of the most popular ways of forecasting inflation. The average proportion of trend extrapolative decisions is smaller when we allow for switching (14%), but in accordance to our conjecture above it varies significantly across treatments (between 6.1 and 23.6%). In 16.9% of cases agents use the general model, 20.2% adaptive expectations, and the remaining 12.1% of cases agents use sticky information type model.

Chapter 2

Uncertainty and Disagreement in Forecasting Inflation

2.1 Introduction

This chapter focuses on an individual forecasting uncertainty and complements the analysis in Chapter 1. The importance of inflation uncertainty has been recognized at least since Friedman's Nobel Lecture (Friedman, 1977). Friedman argued 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 to a reduction in output due to institutional rigidities. Indeed, Levi and Makin (1980) and Mullineaux (1980) found empirical support for the hypothesis that higher inflation uncertainty is associated with lower output. This represents a clear rationale for central banks to care about inflation uncertainty. Moreover, inflation-targeting central banks, in particular, trust that inflationary expectations can be importantly shaped by their communication strategies. They have recently increased their interest in the distribution of inflation expectations, in part because both individual uncertainty and disagreement (interpersonal uncertainty) can be viewed as measures of the effectiveness of their communication strategies. For some central banks these strategies also include publishing their probabilistic forecasts of inflation in the form of fan charts. More generally, the credibility of inflation targets can be assessed using both point forecasts 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) stressed that the cost of inflation should be viewed both in terms of its level and of its uncertainty. This claim is in fact consis-

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

tent with the standard New Keynesian dynamic stochastic general equilibrium (DSGE) model, which shows that in order to maximize consumer welfare the central bank should minimize the variation of inflation (see e.g. Woodford, 2003).[2]

Subjects in the experiment are presented with fictitious economy described by series of inflation, interest rates and the output gap. They are asked to forecast inflation and to provide 95**%** confidence intervals around their point forecasts. Compared to the experiment above, we introduce here a new treatment in which subjects are allowed to report different uncertainty below and above their point forecast. The background New Keynesian model that produces realizations for inflation, the output gap and interest rates is the same as in Chapter 1. Realized values are displayed to the subject and the process is iterated. This allows us to study both individual uncertainty about forecasts and interpersonal variation - disagreement on the point forecasts. Our analysis is based on the same data set as the analysis presented in Chapter 1. Here we focus on the measures of uncertainty and disagreement, we compare them with survey data results, and evaluate 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 than others at stabilizing both inflation and its uncertainty. 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 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 a lower variability in actual inflation. However, there are some treatment effects that go beyond the interest rate channel. Similar evidence is also observed for a treatment where the central bank reacts more strongly to the deviations of inflation expectations from the inflation target.

The results of an analysis of the behavior of individual confidence intervals suggest that the width of the confidence interval is highly inertial and it increases when inflation is below the target level. This contrasts with the results of the survey data literature, where

[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 tradeoff between aggregate and cross-sectional efficiency.

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 the monetary authority? The forecasting ability of different measures has mostly been examined using the survey data of professional forecasters.[3] Three measures have been predominantly used in the survey data literature: the standard deviation of point forecasts, the average individual forecast error variance, and the variance of the aggregate distribution. These measures are complementary in terms of informative value. The first describes disagreement but says little about uncertainty, and the second captures uncertainty but disregards disagreement. Zarnowitz and Lambros (1987) show that there can be substantial differences between the variation in disagreement and the variation in 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 their 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 the measure that has the highest correlations with the variability of inflation.

Several dynamic panel data regressions have been designed to identify the determinants of the three measures discussed above. Disagreement among subjects measured with the standard deviation of point forecasts increases when the average group forecast error increases and when inflation is below the target level. Similar explanatory variables also affect individual uncertainty although disagreement is arguably less inertial. All the factors that significantly affect the specification of uncertainty and disagreement are by definition also important for the interquartile range. Indeed, inflation, the mean forecast error and the 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 the results fall within the specified 95% interval. Giordani and Söderlind (2003) reach similar conclusions when analyzing the survey data of professional forecasters. That subjects tend to report narrower confidence intervals than that asked for is a well-known fact, labelled as the "overconfidence effect." This issue has been extensively debated in the experimental

[3]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 forecasts. Furthermore, the panel pool changes continuously. See Engelberg et al. (2009) for some other methodological issues involved. A laboratory environment presents a potential solution to this problem.

psychology literature. A common approach in this literature is to frame the experiment in the context of stock market forecasting exercises.[4]

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. Studies in the psychology literature also usually assume symmetric confidence intervals (see O'Connor et al., 2001 for a discussion). Symmetric intervals are easier to handle in empirical analysis when the aim is to construct the aggregate distribution of expectations, because it can simply be assumed that an 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, in particular asymmetric monetary policy effects over the business cycle. We have decided to perform treatments Ap with a restriction to symmetric confidence intervals while in treatments Bp we test this assumption and allow subjects to have potentially asymmetric intervals.[5] Only 12.5% of confidence intervals are symmetric when we allow subjects to report asymmetric confidence intervals. Du and Budescu (2007) and O'Connor et al. (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 the 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 subjects to provide descriptive measures of their confidence level (but do not perform any analysis of them), 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 the average forecast but not of studying the behavior of uncertainty or disagreement. They argue that asking for the confidence intervals instead of point predictions in asset pricing framework has the effect of reducing price fluctua-

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

[5]Engelberg et al. (2009) document 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.

tions and increasing subjects' coordination on a common prediction strategy. Our focus is also quite different to that of psychology experiments. The latter usually limit their attention to independent event forecasts, while the present study concentrates on a series of (dependent) forecasts. This allows us to perform a 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 chapter is organized as follows: Section 2.2 describes the model and the experimental design; in Section 2.3 we focus on the analysis of the individual responses while in Section 2.4 we analyze disagreement and the properties of aggregate distribution; Section 2.5 discusses and assesses the forecasting ability of different measures, while Section 2.6 concludes.

2.2 Experimental design

We first provide the summary of the model described in Section 1.2 and describe the design of the treatments depending on the way how confidence intervals are reported. Subjects participate in a fictitious economy and are asked to provide inflation forecasts and a measure of uncertainty about their forecasts. The mean of the point forecasts is then used by the data generating process to calculate inflation, the interest rate, and the output gap. These variables are available to subjects before the next period forecast. We use the reduced form of the forward-looking sticky price New Keynesian (NK) model with different monetary policy reaction functions as an underlying model. The model is discussed more in detail in Chapter 1.

As above in equations (1.1) and (1.2), the IS curve (output gap) and Phillips curve (inflation) are specified as:

$$y_t = -\varphi\left(i_t - E_t\pi_{t+1}\right) + y_{t-1} + g_t, \tag{2.1}$$

$$\pi_t = \lambda y_t + \beta E_t\pi_{t+1} + u_t. \tag{2.2}$$

where i_t is the interest rate, π_t denotes inflation, y_t is the output gap, and g_t is an exogenous shock. McCallum-Nelson calibration that is used, is described in Table 1.1. while shocks g_t and u_t are uncorrelated and unobservable to subjects and follow $g_t = \kappa g_{t-1} + \widetilde{g}_t$ and $u_t = \nu u_{t-1} + \widetilde{u}_t$. $0 < |\kappa| < 1$ 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 $\widetilde{u}_t \backsim N\left(0, \sigma_u^2\right)$. $E_t\pi_{+1} = \frac{1}{K}\sum^k \pi_{t+1|t}^k$, is the mean point forecast of inflation in an economy with K forecasters. Monetary policy reaction functions define four different types of treatments: three with (i) inflation forecast targeting and different levels of γ, and one with (ii) inflation targeting:

$$i_t = \gamma(E_t\pi_{t+1} - \overline{\pi}) + \overline{\pi}. \tag{2.3}$$

$$i_t = \gamma(\pi_t - \overline{\pi}) + \overline{\pi}. \tag{2.4}$$

The slope coefficient γ determines the central bank's aggressiveness in response to deviations in inflation (or inflation expectations) from its target. A higher γ implies a stronger stabilizing effect of the Taylor-type monetary policy rule. As discussed in Chapter 1, there are eight treatments in our experiment, depending on a monetary policy regime and a type of the confidence interval input. In treatment Xp, p represents one of the four monetary policy regimes and X represents one of the two types (A or B) of the confidence interval input.

policy regime p		treatments Ap symmetric confidence interval	treatments Bp asymmetric confidence bounds
Taylor rule (equation)	Parameters	Groups	Groups
1 – Forward looking (2.3)	$\gamma = 1.5$	1-4	5-6
2 – Forward looking (2.3)	$\gamma = 1.35$	7-10	11-12
3 – Forward looking (2.3)	$\gamma = 4$	13-16	17-18
4 – Contemporaneous (2.4)	$\gamma = 1.5$	19-22	23-24

Table 2.1: Treatments

In treatments Ap we restrict inputs to symmetric confidence intervals. Subjects report the difference from their point forecast, which can be interpreted as 1.96 standard errors of their expectation, assuming it is represented by a normal distribution. This condition is relaxed in treatments Bp, 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). Table 2.1 provides a summary.

In the analysis of this chapter our focus is the individual forecasting confidence and the differences between treatments Ap and Bp. Unless otherwise noted, we disregard the potential influence of the monetary policy regimes $p = 1...4$. We will refer to treatments A1...A4 as Ap, and to treatments B1...B4 as Bp.

2.2.1 Experimental procedures

The experimental subjects participate in a simulated economy of 9 agents. The experiment consists of 12 sessions each containing 2 independent groups, thus making 24 groups in total. The participants were enlisted through a recruitment program for undergraduate students at the Universitat Pompeu Fabra and the University of Tilburg. The participants remain in the same group throughout the experiment. They earn on average around €15, depending on the treatment and individual performance. The participants receive detailed instructions and a quiz questionnaire, and play 5 practice rounds before the start of the experiment to make sure they fully understand their task. Instructions to subjects can be found in Appendix A for treatments Ap and in Appendix B for treatments Bp. The program is written in Z-Tree experimental software (Fischbacher, 2007).

The participants observe time series of inflation, the output gap and the interest rate and their past forecasts, up to the period $t-1$. They do not observe the forecasts of other individuals or their performance. 10 initial values are generated by the computer under the assumption of rational expectations. The underlying model of the economy is qualitatively described to them. The subjects' task is to provide inflation forecasts for the period $t+1$ with a 95% confidence interval. After each period subjects receive information about the inflation in that period, their prediction of it, and the payoff they have gained. The payoff function is in essence the same for treatments Ap and for treatments Bp. The only difference between the two types is in the way how a confidence interval, CI, is calculated. For treatments Ap, CI is subject's direct input, it measures the distance of point forecast from the upper and the lower confidence bound of 95% confidence interval. In treatments Bp, subjects have to provide the lower (CB_L) and the upper (CB_U) confidence bound of their 95% interval. The CI is than calculated as $CI = \frac{1}{2}(CB_U - CB_L)$.

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

The first component, W_1 is designed to encourage subjects to give accurate point forecasts, while the second component, W_2, is intended to motivate subjects to think about the variation in actual inflation since it is more rewarding when subject's interval is narrower. There is thus a trade-off between the width of this interval and its accuracy. The W is defined on the interval $[0,4]$.

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 an 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%.[6] 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 decided to formulate the question in terms of 95% confidence intervals.

2.3 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 in disagreement and the variation in uncertainty. Therefore, both might not be appropriate measures for forecasting the 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 the aggregate distribution of forecasts and disagreement.

Figure 2.1 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.[7] The average symmetrical confidence

[6] This is also supported by previous evidence in the literature. We show in Chapter 1 that in our experiment a 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).

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

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 treatments Ap (treatments Bp) 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 et al. (2009).

Average confidence interval	All	Treat. Ap (symmetric)	Treat. Bp (asymmetric)
1 – Forward looking (2.3), $\gamma = 1.5$	0.564	0.669	0.352
2 – Forward looking (2.3), $\gamma = 1.35$	0.776	0.914	0.500
3 – Forward looking (2.3), $\gamma = 4$	0.395	0.466	0.254
4 – Contemporaneous (2.4), $\gamma = 1.5$	0.430	0.410	0.471

Table 2.2: Width of confidence intervals across treatments. Note: The width of asymmetric confidence intervals is calculated as (Upper b. - Lower b.)/2.

The average confidence intervals in each treatment are listed in Table 2.2, while a per-group summary is presented in Table 2.3. In general, confidence intervals are narrower in treatments Bp than in treatments Ap at 1% significance using nonparametric tests (Wilcoxon/ Mann-Whitney). In Section 3.1 we show that treatments Ap and Bp also

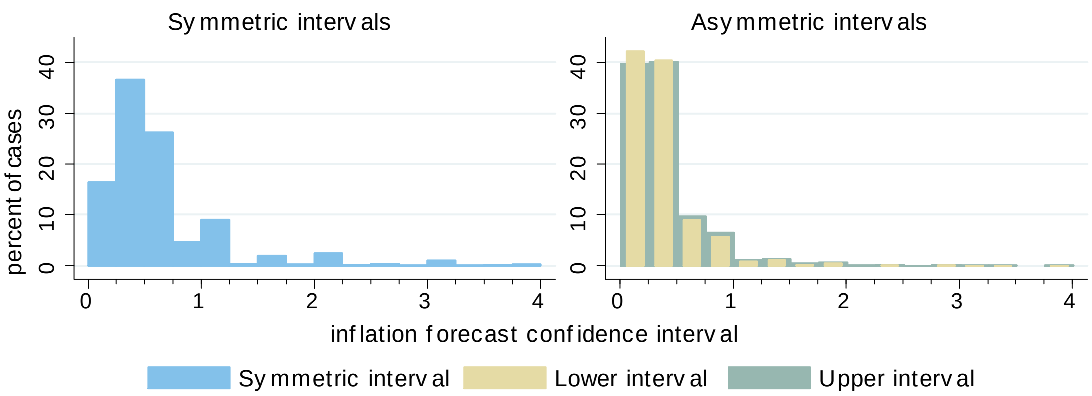

Figure 2.1: Histogram of confidence intervals for treatments Ap (left) and treatments Bp (right), across all subjects and periods.

differ in the 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 to 0.2046 for treatments 1-3 and 0.2081 for treatment 4.[8] 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 γ, while in treatment 4 it depends on γ: higher γ 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. On the contrary, we find that the average confidence interval is narrower in treatment 4 compared to the 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 the other treatments, we observe that while it is significantly narrower than treatments 1 and 2 it is wider than treatment 3. Theory also suggests that in treatments 1-3 all confidence intervals should have the same width; however this is strongly rejected by our experimental data. We can conclude that 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 target.[9]

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 attitude to risk. With professional forecasters it can be claimed that their probability and point forecasts are correlated because they interact and influence each other.[10] 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

[8]The unconditional variances of the residuals following the AR(1) process are $vr_g = \sigma_g^2/(1-\kappa^2)$ and $vr_u = \sigma_u^2/(1-\nu^2)$. The 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 $1.96 \cdot \sqrt{((\lambda\gamma\varphi+1)^{-2} vr_g + (\lambda\gamma\varphi+1)^{-2} vr_u)} = 0.2081$.

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

[10]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 of this.

		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 2.3: Confidence bounds, summary statistics.

information about each other's expectations, nor is the average aggregate prediction directly observable.

2.3.1 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 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) obtain a very similar result (72%).[11] Our results confirm the overconfidence effect in an even stronger manner than survey data results. Only 60.5% of the times do subjects manage to set confidence bounds that include the actual inflation in the next period.[12] This proportion is higher in treatments Ap, where 64.3% correctly specify confidence intervals, while in treatments Bp 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 treatments Ap (Bp) 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 mentioned in the introduction, this overconfidence effect has attracted a lot of attention in the 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 et al., 1982).[13] The most striking example of this bias has been recently documented by Ben-David et al. (2010) who assembled a panel of forecasts by top financial executives. They show that the market returns realized 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 et al. (2005) argue that traders who underestimate risk are prone to the winner's curse.[14]

The accuracy of confidence intervals also differs across different monetary policies. We find that in treatments 3 and 4, subjects are more accurate (62.9% and 69.4% accuracy respectively) than in the benchmark treatment 1 (51.7% accuracy). The differences are significant at a 10% level with the 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 treatments Ap and 2.8% (1.4%) of the subjects in treatments Bp for the lower (upper) bound. Closer inspection allows us to conclude that on average only about 9.0% of the subjects in treatments Ap and 1.4% (8.4%) of the subjects in treatments Bp for the lower (upper) bound on average report the confidence bounds that are not significantly different from 95% confidence intervals based on actual forecast errors. The rest of the subjects on average forecast confidence bounds that are significantly lower than the actual forecast errors. Per-group statistics are reported in

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

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

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

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

Table B.1 in Appendix B.

We check how the volatility of inflation, the width of confidence bounds, and macroeconomic variables affect the likelihood of inflation falling within the specified confidence bound.[15] We estimate the following regression:

$$\begin{aligned} x_t^k &= \alpha^k + \beta sip_{t|t-1}^k + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \varepsilon D_3 y_{t-1} + \\ &\quad + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \zeta i_{t-1} + \delta sd_{t-1}^j + u_t^{em}, \end{aligned} \tag{2.5}$$

where x_t^k takes 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 treatments Bp it is $(CB_U - CB_L)/2$), y_t is the output gap, π_t is actual inflation and i_t is the interest rate for group j. $D_1, \ldots, D_3$ 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; D_3 equals 1 when $D_1 = 0$ and $D_2 = 0$ jointly and is 0 otherwise. sd_{t-1}^j is the 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.

The results for fixed effects logit estimation are reported in Table 2.4 while those for Poisson fixed effects and random effects are reported in Tables B.3-B.5 in Appendix B. As one would expect, when there is a higher volatility of inflation there are more results outside the interval, especially in treatments Bp. This is well documented in the psychology literature as greater volatility leads to overconfidence (e.g. Lawrence and Makridakis, 1989, Lawrence and O'Connor, 1992).[16] However, some studies also find 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 a 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.[17]

[15] Frequencies of forecast errors depending on the inflation cycle can be found in the Table B.2.

[16] 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).

[17] In Table B.9 in Appendix B we also report the results of the relationship between individual k's forecast error $r_{t+1}^k = \pi_{t+1|t}^k - \pi_{t+1}$, and the confidence interval as a measure of uncertainty.

x_t^k :	*all*	*treat.Ap*	*treat.Bp*
$sip_{t\|t-1}^k$	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 2.4: Forecasting 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 the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

2.3.2 Determinants of individual uncertainty

Below we analyze the determinants of confidence bounds using panel data. All the regressions below are estimated using the system GMM estimator of Blundell and Bond (1998) for dynamic panel data. They are replicated for the whole sample (*all*), treatments A*p* (*treat.Ap*), and separately for the part of the interval below the point forecast ($treat.Bp - L$) and above the point forecast ($treat.Bp - U$) in treatments B*p*. 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 the confidence interval and the standard deviation of inflation. We estimate the following regression:

$$sip_{t+1|t}^k = \alpha + \beta sip_{t|t-1}^k + \gamma sd_{t-1}^j + u_t^{em}, \tag{2.6}$$

where individual k's current perceived uncertainty in period t is measured by her confidence interval, $sip_{t+1|t}^k$. The results are reported in Table 2.5.

$sip^k_{t+1|t}$:	all	$treat.Ap$	$treat.Bp-L$	$treat.Bp-U$
$sip^k_{t|t-1}$	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)
α	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 2.5: Confidence intervals and standard deviation of inflation. Note: The 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 the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

We find that confidence intervals are highly inertial. This has previously been documented in Bruine de Bruin et al. (2011), and Giordani and Söderlind (2003). A higher standard deviation of inflation leads to wider confidence intervals, although with a smaller effect in treatments Ap. Du and Budescu (2007) find no relationship between these variables. A positive correlation between the self-reported range of responses and the underlying uncertainty is also found for survey data in Bruine de Bruin et al. (2011).

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 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 underestimates inflation; while $D^k_5 = 1$ when subject k overestimates inflation. We run the following regression:

$$sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma D^k_4 r^k_{t-1} + \delta D^k_5 r^k_{t-1} + \varepsilon D^k_6 r^k_{t-1} + u^{em}_t. \quad (2.7)$$

The results shown in Table 2.6 suggest that subjects increase their confidence intervals after the last observed inflation is outside the interval.[18] This holds for both "undershooting" and "overshooting." In the latter case r^k_{t-1} is negative, so a negative coefficient δ implies that confidence intervals are widened after $\left|r^k_{t-1}\right| > sip^k_{t-1}$. Positive or negative

[18]Table B.8 in Appendix B reports regression with dummies without interaction with actual forecast errors.

$sip^k_{t+1|t}$:	*all*	*treat.Ap*	*treat.Bp − L*	*treat.Bp − 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)
α	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 2.6: Confidence intervals and phases of the economic cycle. Note: $treat.Bp-L$ ($treat.Bp-U$) 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 the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

errors do not result in any significant change in confidence intervals in the next period when inflation falls within the interval. It is also interesting to note that the confidence intervals in treatments Bp exhibit less inertia, especially at the upper bound, compared to treatments Ap. Moreover, the interval above the 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 et al. (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 ones are not.

Several studies have established that there are significant variations in uncertainty over the business cycle; in particular, uncertainty is found to be countercyclical. Bloom (2009) and Bloom et al. (2010) build theoretical models where uncertainty shocks play a key role in business cycle fluctuations. We estimate equation (2.8), where we control for the path of the output gap. In addition, specification (2.8) also allows for the possibility that subjects change their interval forecasts on the basis of their last point forecast errors:

$$\begin{aligned} sip^k_{t+1|t} &= \alpha + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \varepsilon D_3 y_{t-1} \\ &\quad + \zeta i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi \left|r^k_{t-1}\right| + \vartheta T2 + \iota T3 + \kappa T4 + u^{em}_t, \end{aligned} \quad (2.8)$$

where y_t is the output gap, i_t is the interest rate, and $D_1,\ldots,D_3$ are dummy variables as identified with equation (2.5). The estimation results are given in Table 2.7. $T2$, $T3$ and $T4$ are treatment dummies.[19]

$sip^k_{t+1\|t}$:	*all*	*treat.Ap*	*treat.Bp − L*	*treat.Bp − 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)			
α	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 2.7: Confidence intervals and macroeconomic variables. Note: *treat.Bp − L* (*treat.Bp − U*) only includes the 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 the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

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

Friedman (1968) points out 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 (2.8) with inflation (π_{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 the 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 decreases in inflation below the inflation target also increase uncertainty. With specification (2.8) we concentrate on the absolute deviations of inflation from the 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, when inflation is above the target inflation only the upper part of the confidence interval will be widened, whereas when it is 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, although 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, although 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 2.7 we also demonstrate that confidence intervals depend on the last observed absolute forecast error.

The above regressions confirm the asymmetries between the upper and lower confidence bound demonstrated in Table 2.2. 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 γ and more generally on the monetary policy.[20] However, there also exist other treatment effects as can be observed when controlling for a standard deviation of inflation. Table B.7 in Appendix B demonstrates that if we include treatment dummies in regression (2.6) we find that the dummy variable for treatment 2 is significant.

[20] Due to the presence of heterogeneous expectations this relationship is not monotonic. It is found that the relationship between γ and the variability of inflation is U-shaped (see Chapter 1 for further details).

2.3.3 When are confidence intervals (a)symmetric?

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 treatments Bp. Let us first analyze the proportion of subjects that systematically choose either a wider interval above the point forecast as compared to the one below point forecast or vice-versa. It is clear from Table 2.8 that when subjects are given the option to choose an asymmetric confidence interval they often do so, especially in treatments B1 and B2. Moreover, among more than 40% of the subjects who systematically choose asymmetric intervals, fewer than 6% perceive higher uncertainty on the left-hand side of their point forecast. We can also observe that the proportion of subjects choosing symmetric intervals is the highest in treatment 4. Table 2.8 shows that the behavior of subjects in the inflation targeting treatment is more in line with theory than in the 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 (2.3), $\gamma = 1.5$	44.4	50.0	5.6
2 – Forward looking (2.3), $\gamma = 1.35$	50.0	44.4	5.6
3 – Forward looking (2.3), $\gamma = 4$	33.3	66.7	0.0
4 – Contemporaneous (2.4), $\gamma = 1.5$	16.7	72.2	11.1
All	36.1	58.3	5.6

Table 2.8: Proportions of subjects from treatments Bp, depending on the difference between their upper (C_U) and lower (C_L) confidence intervals. When $C_L < C_U$, the subject choose 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 an 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 gave their responses with one or two decimal points accuracy. It is therefore reasonable to define symmetry as $|C_L - C_U| \leq 0.1$; in this case we set $D_8 = 1$.[21] According to this definition 47.2% of our responses in treatments Bp are approximately symmetric. We estimate the following regressions:

$$\begin{aligned} D_z &= \alpha + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \varepsilon D_3 y_{t-1} + \zeta i_{t-1} \\ &\quad + \eta D_L |\pi_{t-1}| + \theta D_H |\pi_{t-1}| + \phi sd^j_{t-1} + u^{em}_t; \;\; z \in \{7, 8\}. \end{aligned} \tag{2.9}$$

[21] Alternatively, we also tried $D_9 = 1$ if $0.9 \leq \left|\frac{ConfIntH_{n-1}}{ConfIntL_{n-1}}\right| \leq 1.1$. The results can be found in Tables B.10 and B.11 in Appendix B.

The results for the logit fixed effects estimator are reported in the first two columns of Table 2.9, while logit and Poisson random effects estimations can be found in Tables B.10 and B.11 in Appendix B. While the above regressions inform us about the likelihood that subjects choose symmetric intervals, they are not suitable for measuring the magnitude of the asymmetry of the individual forecast distributions or their 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_t^k by subtracting the point forecast from the midpoint of the confidence interval. If skw_t^k is smaller (greater) than 0, then the interval is left (right) skewed, and the confidence interval below the point forecast is wider (narrower) than the one above. If $skw_t^k = 0$ then the interval is symmetric. The factors affecting skewness are analyzed on the right-hand side of Table 2.9 using the Blundell-Bond system GMM estimator.

$$\begin{aligned} skw_t^k &= \alpha + \eta skw_{t-1}^k + \beta sip_{t|t-1}^k + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \varepsilon D_3 y_{t-1} \\ &\quad + \zeta i_{t-1} + \eta D_L |\pi_{t-1}| + \theta D_H |\pi_{t-1}| + \phi sd_{t-1}^j + u_t^{em}. \end{aligned} \tag{2.10}$$

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

The skewness measure, on the other hand, also gives us an indication of the 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 the standard deviation of inflation and the skewness of confidence distribution, while we find this relationship only for the case of D_7.

2.4 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 the probability density functions of individual distributions, add them up and analyze the features of aggregate distribution.

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

	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_1y_{t-1}	0.3345	0.4867**	-0.0496
	(0.4229)	(0.2048)	(0.0415)
D_2y_{t-1}	-0.4418**	0.1259	-0.0447
	(0.2093)	(0.2308)	(0.0337)
D_3y_{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)
α	-	-	0.1037**
			(0.0519)
N	4968	4968	4968
Wald $\chi^2_{(8,9)}$	79.3	58.3	156.3

Table 2.9: 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 the 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 the standard deviation of point forecasts to the phases of the economic cycle, interest rate, inflation and the mean forecast error:

$$\begin{aligned} sdv_{t+1|t}^j \quad = \quad & \alpha + \beta sdv_{t|t-1}^j + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \varepsilon D_3 y_{t-1} \\ & + \zeta i_{t-1} + \eta D_L |\pi_{t-1}| + \theta D_H |\pi_{t-1}| + \phi mr_{t-1}^j + u_t^{em}, \end{aligned} \quad (2.11)$$

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

The regressions based on (2.11) are displayed on the left-hand side of Table 2.10. The 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 only inertia and sensitivity to the business cycle play an important role. Disagreement increases when the 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 low inflation can also generate higher uncertainty.

There are some treatment differences regarding the determination of the standard deviation of point forecasts (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. (2.11) with treatment dummies using robust standard errors.

2.4.1 Dispersion of aggregate distribution

Several central banks have started to put the data on the 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 of the adoption of new communication strategies by many central banks. Thus, it is often desirable to aggregate individual distributions and analyze them, rather than calculate 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 by using a triangles approach similar to Engelberg et al. (2009). The mode is set to be equal to the point forecast, while 95% of the derived triangular distribution is set to be between the lower and the upper confidence bound. In this way we generate probability density functions for each forecast by an individual. The distributions are then aggregated cross-sectionally, across the individuals in a group.

We choose the interquartile range (IQR)[22] as an apropriate measure as it is less sensitive

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

	$sdv^j_{t+1\|t}$			$IQR^j_{t+1\|t}$		
	all	*treat.Ap*	*treat.Bp*	*all*	*treat.Ap*	*treat.Bp*
$sdv^j_{t\|t-1}$	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)
α	-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 2.10: 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 the potential presence of clusters in treatments. */**/*** denotes significance at 10/5/1 percent level.

to small variations in the tails of the estimated density compared to the cross-sectional standard deviation of the aggregate distribution.[23] Nevertheless, it is useful to show that the variance of aggregate distribution is related to the two measures that we study above. Boero et al. (2008) show explicitly that the variance of the aggregate distribution can be decomposed into the average individual uncertainty and disagreement of point forecasts.

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

To discover the properties of the aggregate distribution, we run the following regression:

$$\begin{aligned} IQR_t^j &= \alpha + \zeta IQR_{t-1}^j + \beta D_1 y_{t-1} + \gamma D_2 y_{t-1} + \delta D_3 y_{t-1} \\ &\quad + \varepsilon i_{t-1} + \eta D_L |\pi_{t-1}| + \theta D_H |\pi_{t-1}| + \eta m r_{t-1}^j + u_t^{em}, \end{aligned} \tag{2.12}$$

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

Equation (2.12) considers the sources of divergences in expectations, such as the output gap, the interest rate and the 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 the disagreement of expectations (see Giordani and Söderlind, 2003). We therefore also include the previous period interquartile range among the independent variables and find them highly significant. The results on the right-hand side of Table 2.10 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 the variance of inflation expectations.[24] We observe that the interquartile range is positively correlated with the absolute level of inflation when inflation is below the target level. In treatments Ap, the 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.

2.5 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. Various studies argue that disagreement measured as the standard deviation of point forecasts lacks a theoretical basis and is therefore 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 et al. (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

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

survey data articles point out that the advantage of the measure of disagreement among forecasters (*sdv*) is that it is available in any survey, whereas only a limited number of surveys ask for measures of individual uncertainty. Our design thus allows us to also use the average confidence interval (*asip*) for comparison. A proxy for the uncertainty may be the average absolute forecast error across individuals (*mr*). A measure of the variation in the aggregate distribution of forecasts gives information about both uncertainty and disagreement. The interquartile range (*IQR*) is a proxy for that. Figures B.1 and B.2 in Appendix B display a timewise comparison between the average confidence interval, the standard deviation of point forecasts and the interquartile range for each group.

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

	mr_t^j	$asip_t^j$	sdv_t^j	IQR_t^j	π_{t+1}	i_{t+1}	y_{t+1}
mr_t^j	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			
π_{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 2.11: Pairwise correlation coefficients. 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 2.11, there is a significant correlation coefficient (about 0.5) between the average width of the confidence interval and the standard deviation of point forecasts.[25] Rich and Tracy (2010) and Boero et al. (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. The present analysis suggests that uncertainty and disagreement are modestly correlated.

A positive correlation between the interquartile range and individual uncertainty can be observed. The correlation coefficient (around 0.7) is higher than that reported in the

[25]Table B.6 in Appendix B depicts the relationship between confidence bounds and the dispersion of point forecasts in more detail. We find no evidence of this relationship for symmetric intervals, while for asymmetric there is a positive relationship.

previous paragraph. As shown in the statistical analysis by Boero et al. (2008), there exists a "structural" relationship between these two variables so a positive relationship is expected. For similar reasons there is also a correlation between the disagreement and the interquartile range. The latter correlation is of similar magnitude to the former. Therefore, one could argue that the 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 the standard deviation of point forecasts is a useful proxy for uncertainty and that disagreement tracks uncertainty better than the GARCH model; however, this view is questioned by Rich and Butler (1998).[26]

Policymakers are interested in inflation uncertainty and in obtaining proxies for it. 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 2.11, 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 the average perceived uncertainty of subjects. In order to further assess the forecasting performance of these measures we estimate the following regression:

$$\begin{aligned} sd_t^j &= \alpha + \beta sd_{t-1}^j + \gamma asip_{t-1}^j + \varepsilon sdv_{t-1}^j + \delta IQR_{t-1}^j \\ &\quad + \zeta i_{t-1} + \eta \pi_{t-1} + \phi y_{t-1} + u_t^{em}, \end{aligned} \tag{2.13}$$

where $asip_{t-1}^j$ is the average confidence interval in period $t-1$ for group j. Table 2.12 reports the results. We estimate three different specifications, which are a subset of the above equation. In variant (a) we include all three measures, while in variant (b) we include only measures of individual uncertainty and disagreement. Variant (c) embeds only the 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 the standard deviation of inflation.

The 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 the interquartile range has a marginally significant negative effect. This may be due to a degree of multicollinearity between the IQR and the standard deviation of point forecasts and/or mean confidence intervals. In specification (c) the effect of the IQR is insignificant, while in specifications (a) and (b) we observe that only the

[26] Lahiri and Sheng (2010) point out that disagreement is useful for forecasting in stable periods 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)
π_{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)
α	-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 2.12: 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 the potential presence of clusters in treatments. */**/*** denotes significance at 10/5/1 percent level.

average individual confidence interval has a positive and highly significant effect on *sd*. Therefore, we can conclude that to forecast 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 the survey data literature, as we reach similar conclusions to those of Zarnowitz and Lambros (1987), Boero et al. (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 the 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 have significantly positive effects with negative development having a more profound effect. The output gap exerts a negative effect on *sdv*.

2.6 Conclusion

In Chapter 2 we have presented 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 of the analysis has been on 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 have focused on various modifications of the original Taylor rule and studied the influence of different monetary policy designs on the formation of confidence bounds. We have found that inflation targeting produces lower uncertainty and higher accuracy of intervals than inflation forecast targeting. The treatment that reacts strongly to deviations in inflation expectations from the inflation target also produces similar effects as stated above, compared to treatments that do not react as strongly to deviations in inflation forecasts. This effect not only channels through the variability of inflation, but there is also evidence that there are additional effects, for example the monetary policy rules that were discussed in Chapter 1.

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

We have also analyzed measures of individual uncertainty, disagreement among forecasters and the properties of aggregate distribution. All these measures are related, as argued in Boero et al. (2008), although they have very different features. The interquartile range is a measure of both uncertainty and disagreement. We first analyzed the formation of confidence intervals. We found that confidence intervals are positively related to inflation variability, that they are highly inertial and that they 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 causes wider confidence intervals. Below target inflation also causes the interval to increase and absolute deviations from the inflation target is an appropriate variable to take into account. Furthermore, we have been able to establish some facts about the differences between the formation of lower and upper bounds. In particular, we have found that the upper bound is more sensitive to the stage of the business cycle while the lower bound exhibits significantly more inertia**.**

More generally, we have also studied the determinants of the choice of asymmetric interval. In our treatments B*p*, subjects have the possibility of choosing an asymmetric confidence interval, while in treatments A*p* they are restricted to symmetric intervals. Wehave found that in only about 12.5% of cases subjects choose symmetric intervals when they have the possibility of choosing an asymmetric interval. Moreover, in treatments B*p* more than 35% of subjects report higher upper bounds than the lower ones, while only about 5% of subjects show the opposite pattern. Symmetric intervals are more likely to be observed 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 the standard deviation of point forecasts and the interquartile range of the aggregate distribution? We have documented that *IQR* is more inertial than *sdv*, while they both increase when inflation is below the target level. We have also compared forecasting performance of these measures and observed 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 affects our forecasting specifications signifficantly. Therefore, we confirm the previous results from the survey data literature that more central banks should design their surveys in such a way that each individual provides their whole distribution of forecasts or at least some measure of the uncertainty of their forecasts. In this sense it might be enough if they were asked for their confidence intervals as in our treatments A*p*. Generally, this would greatly enhance the informativeness of these surveys as central banks would also receive a proxy for forecasting inflation uncertainty.

Chapter 3

Uncertainty about the Number of Bidders in Sequential Auctions with Unit Demand

3.1 Introduction

Empirical data becomes of key importance to auction economists if it shows that it persistently contradicts the theoretical models. With the growing availability of trading platforms and increasing use of auction mechanisms as the method of exchange, divergence from theoretical predictions poses a challenge to mechanism designers. Can they incorporate this behavior in their design and increase expected revenues? Price behavior in sequential auctions is one of the cases where this divergence is particularly obvious. It has been proved (Weber, 1983) that prices of identical goods sold sequentially in an auction should, under certain assumptions, remain constant on average for all the goods sold in a sequence. Following that, numerous empirical studies have recorded that price in fact declines for subsequent units sold. This has been particularly true in auctions for the works of art (Beggs and Graddy, 1997), wine auctions (Ashenfelter, 1989 and McAfee and Vincent, 1993) and jewelry (Chanel et al., 1996). Paper of Ashenfelter and Graddy (2003) for example, gives a good review of art auctions and anomalies present there.

Given discrepancy between theory and empirical findings was also my main motivation for research. Works so far have tried to explain price decline with the risk aversion of participants, the fact that many goods have significant common value component, not perfectly symmetric and identical units in a sequence, asymmetric bidders' valuations, presence of absentee bidders or even specific auction rules.

Our analysis will rather concentrate on the role of uncertainty of the number of bidders for the empirical evidence on price decreases. One of the key assumptions of the Weber's theorem is that the number of participants in an auction is exogenously fixed and known to all participants. Releasing this assumption and comparing the outcome under uncertainty with that under certainty should shed some light on reasons for price decreases. We demonstrate that when rational risk-neutral bidders know the ex-ante probabilities of the number of competitors, the equilibrium price should remain constant for all units sold in a sequence.

The second advantage of the present approach is that it uses experiment to obtain the evidence on price paths. Experiment allows us to control for many variables that were otherwise disputed in other attempts to explain declining price trend and foremost it enables us to have true symmetric and independent private values. In particular we will consider sequential second-price independent private value auctions where values are fixed and bidders have demand for only one unit of each good.

The structure of the chapter is as follows. The literature review discusses the background and the related research. In the theoretical part equilibrium bids are calculated and predicted price path for the case where bidders know the size of the competition and for the case where they only know its probability distribution. Assumptions on the experimental outcome are also given. In the second part of the chapter, the laboratory experiment on bidding behavior in two-stage second-price independent private value auctions is presented. Experimental design is followed by the analysis of the results and the conclusion with comments.

3.2 Related literature

Sequential auctions have in last decade received considerable attention from experimental economics, largely due to empirical evidence that frequently shows divergence from equilibrium prices. Price behavior has been studied mostly in the setting of sequential auctions with multiunit demand. In uniform auctions of that kind, there is an inherent incentive that bidders reduce their demand for subsequent units. Kagel and Levin (2001) study sensitivity of bidders in two different auction types and find demand reduction in both of them. List and Lucking-Reiley (2000) make a field experiment and arrive to the similar conclusion. Engelmann and Grimm (2006) compare different auction types in an auction with 2 bidders and 2-unit demand. They test various types for allocative efficiency, demand reduction and overbidding. Their results mostly confirm those of Kagel and Levin, despite they also obtain some contradictions.

Our analysis rather concentrates on the price behavior in the multiunit auctions with single unit demand; in particular, to the origins of the "price decline anomaly". Weber (1983) published an influential paper where he studies sequential auctions with independent private values. He shows that under risk neutrality, known number of bidders, fixed values and unit demand, the price tendency should be a martingale, with other words, should remain constant among auctions. The logic behind is as follows: Bidders in the earlier auctions bid less than their valuations since they have a positive expected profit from participation in the subsequent auctions. Therefore they decrease their bid for exactly that amount. Bidders with high valuations have higher opportunity cost on non-participation in the later rounds so they will discount their bids in the first round for more than bidders with low valuations. These tendencies counterbalance and the prices in the sequence of auctions remains the same in the two auctions.

This, rather strong conclusion has been studied extensively ever since Ashenfelter (1989) reported an evidence from vine auctions, that prices are twice as likely to decrease than increase for identical bottles of vine sold in the same lot sizes. Ashenfelter called it a price decline anomaly and afterwards it has been identified in various auction types, and auctions of different types of goods and most of the work confirms the results by observing price declines.

Several proposals have been made in attempt to explain this price movement, majority by releasing one or more assumptions of the original Weber's model. For example Black and de Meza (1992) claims that price decline only occurs in auctions where the winner of the first auction has an option to buy all objects at the same price. However, later it has been shown (Ashenfelter and Graddy, 2003) that price anomaly exists even in auctions which do not allow for such option. Since Weber's model assumes risk neutrality McAfee and Vincent (1993) tries to explain declining prices by the presence of risk averse bidders, who prefer buying earlier than later. They show that in the case of first and second price auctions in the presence of risk averse bidders, prices can have a declining trend. The logic behind the decline is that bidders are uncertain about the prices in the later auctions and are therefore prepared to pay a premium in the earlier auctions. The drawback of their solution is the assumption of non-decreasing absolute risk aversion required for existence of pure strategy equilibrium bidding function. This is usually not considered to be a very realistic assumption. Authors also present an example with decreasing absolute risk aversion for which mixed strategy equilibria exist, but these are ex post inefficient.

Another potential explanation why prices tend to decline is offered by Ginsburgh (1998). He analyzes several wine auctions at Christie's London and observes significant decline in prices for the identical wine lots sold in a sequence. He also notices that 60% of all the lots were sold to absentee and he attributed a decline in price to them. The absentee bidders use simple strategies that significantly differed from theoretical predictions. The drawback of Ginsburgh explanation is the fact that the strategy that dominated roughly 65% of the auctions was to bid prices for several lots and stop bidding after having won given number of them. However, this is in fact an auction with multiunit demand and in this case it is also theoretically optimal for some bidders to reduce demand for additional items in order to pay less for their winnings.

Several other works try to access problem of declining prices releasing the assumption about independence of bidders' valuations. Engelbrecht-Wiggans (1994) shows that when objects are stochastically equivalent (but not identical), Weber's conjunction no longer holds. The paper demonstrates that price trend depends on the distribution of objects' values, however, when number of repetitions is large enough, prices on average decrease. Bernhardt and Scoones (1994) take similar assumption. In their model, each bidder's valuations are identically distributed across the objects but are not perfectly correlated. They show that even if bidders are risk neutral, mean prices fall. The intuition behind their claim is the following: where in Weber's model bidders with high valuations also discount more than bidder's with low valuations, here everybody discounts the same. In the two stage auction game all bidders expect the same profit from bidding on the second object so they also discount their bids in the first stage for the same amount. They claim this is because the bidders with high valuations determine the price in the first auction, if there are sufficiently many bidders.

Experimental work dealing with price decline has so far confirmed the phenomenon[1]. Burns (1985) made experiments of simulated sequential wool auctions where he had treatments both with students and professional wool traders as subjects. Prices declined in both of the treatments; however it is interesting that they disappeared with repetition in the treatment with students, while they persisted with experienced wool bidders. None of the groups noticed the price decline, when asked in the questionnaire after the experiment. Keser and Olson (1996) make laboratory experiment designed to address specifically the price decline phenomena. They construct simple environment, in line with the Weber's setup. Their conjecture confirms previous findings. They still find declining prices. They also find correlations between the number of price declines and position in sequence in which the units are sold. Their experiment controls for several previously suggested reasons such as uncertainty of supply, option to buy the whole lot, or the presence of buyer's agents on the bidding pool. However, they have not been able to identify the actual source

[1] In certain cases the contrary can be true. Neugebauer and Pezanis-Christou (2007) show that if demand is uncertain, prices for additional units will rise. Milgrom and Weber (2000) demonstrate similar effect for a model with affiliated values.

of price declines. Wang (2006) in his sequential second price auction experiment identifies price declines. He demonstrates that applying cognitive hierarchy model gives much better explanation of the price path.

The role of uncertainty about the number of bidders has received much less attention in the auction theory. McAfee and McMillan (1987) show that in a first-price sealed bid auction with bidders having independent private values and constant or decreasing risk aversion the expected revenue is strictly higher when the bidders do not know how many competitors there are then when they do know it. Bidders will bid more aggressively when number of bidders is high and less aggressively when the number of bidders is low. They demonstrate that releasing information on the degree of competition increases value variance ex ante which implies a price decrease on average. They call this a bid dispersion effect. Even though the effect is only shown in the case of single unit auction, it can be applied to sequential auctions where some information is released during the auction process.

In more recent paper Pekec and Tsetlin (2008) compare the effect of uncertainty of number of bidders in uniform and discriminatory auctions. They show that releasing the assumption about certainty of number of bidders in a general model can result in a situation that symmetric increasing equilibrium might not exist. Furthermore, whereas in the case of no uncertainty uniform auctions yield greater revenues than discriminatory auctions, revenue ranking of uniform and discriminatory auctions can be reversed in the case of uncertainty. It has to be noted that results of Pekec and Tsetlin are based on assumption that a good has private and common value properties. In the case of pure private values, an increasing bidding function is a sufficient condition for an increasing symmetric equilibrium.

Experiments regarding different number of bidders have so far only been done in single unit auctions. For example, Cox et al. (1988) analyzed the effect of changing the number of bidders on the individual's bid functions in the independent private value auctions. Assuming first price sealed bid auction and risk neutral bidders, the risk neutral Nash equilibrium bidding function as developed by Vickrey clearly increases in slope if the number of bidders is increased. Cox at al. showed that the extent of this influence depends on individual's risk aversion.

Dyer et al. (1989) tries to show the effect of number of bidders being concealed on the expected revenue. Their experimental design in independent private value first-price sealed bid auction construct requires subjects to submit bids contingent on given number of bidders in first treatment, whereas in the second treatment the subjects were at the same time also asked to submit unconditional bid, with the same uncertainty of distribution of bidders as before. Results of the experiment confirm their predictions that the non-contingent bid is on average greater than the bid contingent on 3 participants and lower than the bid contingent on 6 participants. It has also confirmed the hypothesis of McAfee and McMillan (1987) that the bidding with uncertainty of number of bidders generates on average higher expected revenues than the bidding contingent on given number of bidders. However, regarding the individual bidding Dyer at al. rejected the Nash equilibrium bidding hypothesis since less than half of all bids satisfy full inequality requirements, which include the condition of non-contingent bid being greater than the weighted average of the two contingent bids. Experimental design of this chapter resembles the one of Dyer at al. applied to sequential auction environment.

3.3 Theoretical considerations

For the experiment, we will consider a sale of two identical objects, which are sold sequentially in a second price sealed bid auctions. Each bidder i has a value x_i, known only to him. He also has wants to obtain only one of the two goods offered.

3.3.1 Assumptions

A1 *Private values* model. Bidders know their own valuation of the object with certainty, but they do not know the values of all the other bidders. However, they know the distribution of other bidders' values and this knowledge is common to all the bidders.

A2 *Independently distributed values* $x_i, \dots, x_n$ of the bidders.

A3 *Symmetry of bidders.* Value of each bidder is drawn from the same distribution and is the only distinction among bidders.

A4 Sequence of two auctions with *fixed values*. Bidders keep the same value, x_i in both auctions.

A5 *Unit demand.* Bidders only have a demand for one unit of goods. Once a bidder has one an item, her utility of the second unit is zero, and will leave the game.

A6 *Risk neutrality* of bidders.

Every bidder's utility depends on both auctions. We assume Von Neumann utility function $u(x_i - q)$, where, for the first auction, bidder receives utility from the difference between her valuation and the paid price conditional that she wins the auction; and 0 otherwise, whereas in the second auction bidder receives the utility from the difference between her valuation and the price paid conditional that she won the second auction and did not won the first one. Otherwise she receives 0.

In our analysis, two distinct games are taken in consideration. The first one assumes that bidders are fully aware of the number of bidders, whereas the second supposes that there exists uncertainty regarding the number of bidders. We assume there exists pure strategy symmetric equilibrium bidding function $b(x_i)$. A bidder i bids $b = b(x_i)$ in the first auction and her value, x_i in the second. Bidders in the second price auctions have a dominant strategy to bid their true valuations if only a single unit is sold. The same is the bidding strategy for the last unit in sequential action with given assumptions. In the two-unit auction bidders have an incentive to decrease their bid in the first auction, in order to keep an option of participating in the second auction.

3.3.2 Auction with given number of bidders

We define X_i as an independent variable, x_i its realization, $P(x)$ its a cumulative distribution function and $p(x)$ its probability distribution function. Payoff function of a bidder who bids r, has a valuation x, and Von Neumann utility function $u(x-y)$ can be written as follows:

$$V(r;x) = \int_0^r u(x - b_1(y)) f_1(y)dy + \int_0^x u(x-y) f_2(y)dy \tag{3.1}$$

The first part of the equation is the bidder's payoff if she wins the first auction, whereas the second part is her payoff from the second auction, conditional on losing the first auction and winning the second. We assume that the bidders' values $X_1, \ldots, X_n$ have the order statistics defined as $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$. The expected revenues from the auction for the first unit depend on the bid of the second highest bidder, assuming that the bidder with value x won the auction.

$$f_1(y) = f\left(y = X_{(n-1)} \middle| X_{(n-1)} < x\right) \cdot f(x = X_{(n)})$$

It can be shown that the conditional distribution of $X_{(j)}$, given that $X_{(i)} = x_i$ for $j < i$, is the same as the distribution of the j-th order statistic in a sample of size $i-1$. Using this property, we can calculate $f_1(y)$ as PDF of the beta distribution with parameters $\alpha = k = n-1$ and $\beta = (n-1)-k+1 = 1$.

$$f_1(y) = (n-1)P(y)^{n-2}p(y) \tag{3.2}$$

By analogy we can calculate $f_2(y)$ with beta distribution parameters $\alpha = n-2$ and $\beta = 2$.

$$f_2(y) = (n^2 - 3n + 2)P(y)^{n-3}(1 - P(y))p(y) \tag{3.3}$$

Therefore the expected utility can be written as:

$$\begin{aligned} V(r;x) &= \int_0^r u(x - b(y))(n-1)P(y)^{n-2}p(y)dy + \\ & [1 - P(r)] \int_0^x u(x-y)(n^2 - 3n + 2)P(y)^{n-3}p(y)dy \end{aligned} \tag{3.4}$$

Differentiating and setting the payoff function $V'(x,x) = 0$, we get a bidding function $b^*(y)$ that satisfies

$$u(x - b^*(x)) = \int_0^x u(x-y)\frac{(n-2)P(y)^{n-3}p(y)}{P(x)^{n-2}}dy \tag{3.5}$$

McAfee and Vincent (1993) showed that symmetric increasing pure strategy equilibrium bidding function b exists for every distribution P if and only if u displays non-decreasing absolute risk aversion. Therefore, in an simplified case, when $u(x-y) = x-y$, we get that optimal bid in first period equals

$$b^*(x) = x - \frac{(n-2)}{P^{n-2}(x)}\int_0^x (x-y)P(y)^{n-3}p(y)dy \tag{3.6}$$

Proposition 1 *Let assumptions A1–A6 hold and let the number of bidders be fixed and known to all. Also let q_n denote price paid for the nth item, $X^{(1)}, \ldots, X^{(n)}$ denote the order statistics in decreasing order and I_n the sequence of the past prices $\{q_1, \ldots, q_n\}$. Then for all $m \leq n \leq k$ the unique symmetric Nash equilibrium for the second price sequential auction satisfies: $E[q_n|I_{m-1}] = E\left[v(t^{(k+1)})|I_{m-1}\right] = q_{m-1}$.*

Example. For a general proof see Weber (1983). Here we demonstrate the result for the case of the two stage Vickrey auction. Let's assume Von Neumann utility function to satisfy $u(x_i - y) = x_i - y$. We also assume values to be uniformly distributed on an interval $[0, x_H]$. The density function is therefore a constant, $p(x) = x_H^{-1}$ and cumulative distribution function is $P(x) = x/x_H$. Inserting this into equation (3.6) gives us the optimal bid function in the first stage

$$b^*(x) = \frac{n-2}{n-1}x \tag{3.7}$$

The proportion for which agents decrease their valuation to form a bid is therefore exactly inverse to the number of their opponents. The expected price paid by the winner of the first stage equals to the second highest bid: $E(q_1) = E(b_{(n-1)})$. Since the bid function (3.7) is monotonic and increasing, the ranking of the bids will be the same as the ranking of the corresponding values, $E\left[b_{(k)}(x)\right] = E\left[b(X_{(k)})\right]$. Moreover, the bid function is linear and $E\left[b_{(k)}(x)\right] = b\left[E(X_{(k)})\right]$ equality also holds. Therefore $E(q_1) = b\left[E(X_{(n-1)})\right]$. Given that the expected value of k-th order statistics equals $E(X_{(k)}) = k/(n+1)x_H$, the expected second highest value in the auction equals

$$E(X_{(n-1)}) = \frac{n-1}{n+1}x_H \tag{3.8}$$

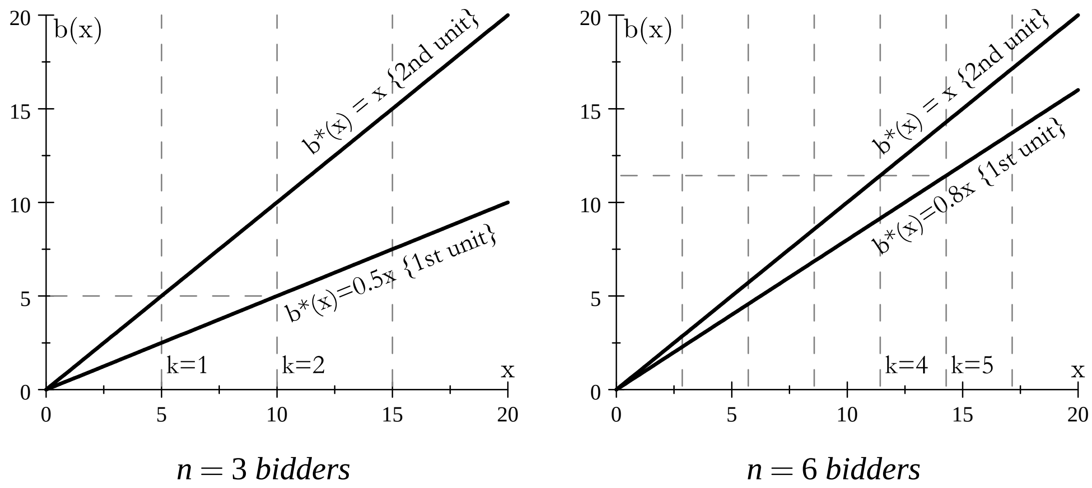

Figure 3.1: RNNE bid functions on value interval [0, 20] for auctions with 3 bidders (1a) and 6 bidders (1b) for the two units. Vertical lines represent the expected value, $E(X_{(k)})$ of the k-th order statistics. Horizontal line represents the expected auction price, the same for both units $E(q_1) = E(q_2)$.

By inserting equation (3.8) into equation (3.7), we get that the expected price for the first unit sold equals to:

$$E(q_1) = \frac{n-2}{n-1}E(X_{(n-1)}) = \frac{n-2}{n+1}x_H \tag{3.9}$$

In the second stage auction the bidders have a dominant strategy to bid their own valuation. A bidder with the highest value, $X_{(n)}$ won the first unit and does not bid for the second. A bidder with second highest value wins the second unit and pays the price equal to the third highest value:

$$E(q_2) = E(X_{(n-2)}) = \frac{n-2}{n+1}x_H \tag{3.10}$$

which clearly confirms the equivalence $E(q_1) = E(q_2)$.

We demonstrate the risk neutral Nash equilibrium (RNNE) bid functions from our example, equation (3.7) for the first unit, and $b^*(x) = x$ for the second unit in the Figures 3.1 (left) and 3.1 (right). In an auction with 6 bidders for the two units, the average valuation of the second highest bidder, $X_{(5)}$ is 14.2857, while her average RNNE bid, $E(b_{(5)})$ equals 11.4286. This is also the average price (equation 3.9) paid for the first unit $E(q_1)$. In bidding for the second unit only 5 bidders remain and each bids her value. Unit is won by the highest bidder $k = 5$, who pays an average price $E(q_2)$ equal to the third highest value $E(X_{(4)}) = 4/7x_H = 11.4286$. Prices for both units are therefore equal $E(q_1) = E(q_2)$. Same logic applies to the auction with 3 bidders for the two units, where resulting prices equal $E(q_1) = E(q_2) = 5$.

■

3.3.3 Uncertainty about the number of bidders

The optimal bidding strategy of a bidder changes when she doesn't know the exact number of bidders, and only knows the distribution function of the number of bidders. Matthews (1987) and Harstad et al. (1990) show that unique symmetric Nash equilibrium payoff function for risk-neutral bidders can also be written as a weighted average of her payoffs conditional on the number of bidders:

$$V(r;x) = \sum_{n=1}^{N} \pi_n V_n(r;x) \tag{3.11}$$

where $N < \infty$ is the maximum number of bidders possible and π_n is the ex-ante probability for having n bidders in the auction which equals

$$\pi_n = \frac{n\beta_n}{\sum_{n=1} i\beta_i} \tag{3.12}$$

and where β_i is an exogenous probability of having i participants in the auction. Thus, the unconditional payoff function becomes:

$$V(r;x;n) = \sum_{n=1}^{N} \pi_n \left(\int_0^r u(x - b_1(y|n)) f_1(y|n)dy + \int_0^x u(x-y) f_2(y|n)dy \right) \tag{3.13}$$

Similar as before we calculate the derivative of expected utility function and set it to 0 to obtain that optimal bid satisfies:

$$u(x-b^*(x)) = \frac{\int_0^x u(x-y)\sum_{n=1}^N \pi_n(n-1)(n-2)P(y)^{n-3}p(y)dy}{\sum_{n=1}^N \pi_n(n-1)P(y)^{n-2}} \tag{3.14}$$

Using our simple utility function example $u(x-y) = x-y$, we obtain:

$$\begin{aligned} b^*(x) &= x - \frac{1}{\sum_{n=1}^N \pi_n(n-1)P^{n-2}(x)} \times \\ & \int_0^x (x-y)\sum_{n=1}^N \pi_n(n-1)(n-2)P(y)^{n-3}p(y)dy \end{aligned} \tag{3.15}$$

Example. Let's use the utility function given with equation (3.15), and assume only two possible states, $n_1 = 3$ and $n_2 = 6$ with according probabilities $\beta_{n_1} = \beta_{n_2} = 0.5$. Values are drawn independently from the uniform distribution with support $[0, x_H]$. Therefore PDF and CDF are $p(x) = x_H^{-1}$ and $P(x) = x/x_H$ respectively and the ex-ante probabilities are $\pi_{n_1} = 1/3$ and $\pi_{n_2} = 2/3$. Plugging these values in the equation (3.15) we obtain that the risk neutral optimal unconditional bidding function equals

$$b^*(x) = \frac{8x^4 + x_H^3 x}{10x^3 + 2x_H^3} \tag{3.16}$$

■

In the Figure 3.2 (right) risk neutral Nash equilibrium bidding functions under all 3 different procedures are displayed and compared to the optimal bidding in the second stage. The calculation of expected price paid by the highest bidder is not straightforward here since the bidding function is not linear. We can generally write it as:

$$E(q_1) = \int_0^{x_H} b^*(x)\sum_{n=1}^N \beta_n f_{X_{(n-1)}}(x|n)dx \tag{3.17}$$

The probability for each possible state, $f_1(x|n)$ follows PDF of the beta distribution with (α, β) parameters equal to $(2,2)$ for $n = 3$ and $(5,2)$ for $n = 6$. A probability weight is therefore:

$$\begin{aligned} \sum_{n=1}^N \beta_n f_{X_{(n-1)}}(x|n) &= \beta_3 \frac{1}{B(2,2)} P(x)(1-P(x))p(x) + \\ & \beta_6 \frac{1}{B(5,2)} P(x)^4(1-P(x))p(x) \end{aligned} \tag{3.18}$$

Applying the results from Example 2, we get:

$$E(q_1) = \int_0^{x_H} \frac{8x^4 + x_H^3 x}{10x^3 + 2x_H^3} 3(1+5(\frac{x}{x_H})^3)(1-\frac{x}{x_H})\frac{x}{x_H}\frac{1}{x_H}dx$$

which integrates to:

$$E(q_1) = \frac{23}{56}x_H \tag{3.19}$$

The solution of the given expression for the $x_H = 20$ is $E(q_1) = 8.2143$. In the second auction bidders bid their valuation, regardless the number of bidders present. Therefore the expected price under uncertainty is the linear combination of the two expectations under certainty:

$$E(q_2) = \sum_{n=1}^{N} \beta_n \frac{n-2}{n+1} x_H \tag{3.20}$$

Applied to Example 2, the equation simplifies to $E(q_2) = \frac{23}{56}x_H$ which yields $E(q_2) = 8.2143$ for $x_H = 20$. Comparing equations (3.19) and (3.20) we clearly see that $E(q_1) = E(q_2)$ which speaks against the assumption that the uncertainty of the number of bidders may be the reason for decreasing prices in the present auction design.

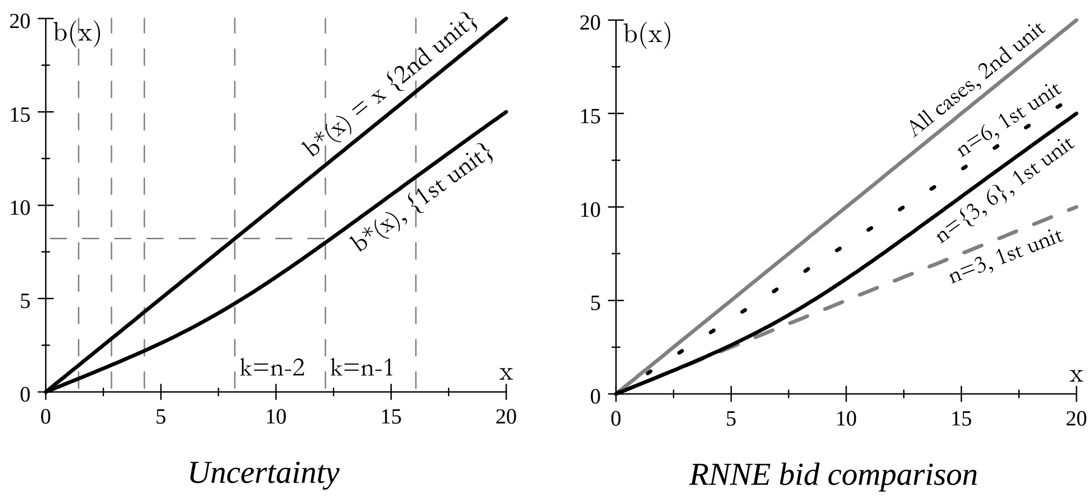

Figure 3.2: Left graph shows RNNE bid functions on value interval [0, 20] for auctions of two units with uncertainty between 3 and 6 bidders with equal probability. Vertical lines represent the expected value, $E(X_{(k)})$ of the k-th order statistics. Horizontal line represents the expected auction price, the same for both units $E(q_1) = E(q_2)$. The graph on the right compares RNNE bidding functions under certainty and uncertainty.

Optimal bidding function for our example is depicted in Figure 3.2 (left). The bidding function $b(x)$ for the first unit corresponds to equation (3.16). Note that due to convexity of the later, $b(E[X_{(n-1)}])$ is slightly smaller than $E[b(X_{(n-1)})] = E(q_2)$. Equilibrium prices are therefore the same for the first and the second unit in all three examples.

3.4 Experimental design

An auction experiment is designed with an attempt to trace the price path of the selling identical good sold in a sequence. In particular, we are interested in the influence of the uncertainty over the number of bidders participating in the auction on the selling price. The auction is sealed bid, second price, independent private value model, with two-unit supply.

Despite there are a few experiments trying to capture price decline anomaly in sequential auctions, see for example Keser and Olson (1996) or Burns (1985), these studies assume the number of bidders as fixed and a prior knowledge of all participants in the auction. There exist also experimental studies regarding the uncertainty of the number of bidders, see for example Dyer et al. (1989), but the latter are made in a single-unit supply environment, and rather focused in the optimal design and revenue maximizing strategies for the seller. As opposed to that, multi-unit auction with demand uncertainty is investigated, with focus on price development in the sequential offers of the same "good".

In each experimental treatment we had either 18 or 12 participants, recruited among undergraduate students of Universitat Pompeu Fabra in Barcelona. The students had mostly (but not exclusively) economics and business background. Subjects were randomly matched after each period into groups of 3 or 6 bidders. Each of these groups constitutes a separate auction for the two identical fictitious goods sold sequentially. The reason for choosing that quantity of participants is simple. 3 is the smallest number of participants to make a two-unit auction nontrivial. 6 subjects on the other side represent sufficient increase from 3 to make significant change in optimal bid/value ratio. Every subject i is randomly assigned a value (reservation price) x_i, drawn from uniform distribution defined on the interval $[0, 20]$. As in standard private value construct, every subject is informed about the distribution of values and upper and lower limit of the interval. However, subject is not informed about individual values but her own.

There are two identical items offered for "sale". The auction experiment therefore functions in two stages. In the first stage first "item" is sold, and the second item is sold in the second stage. We also assume that each bidder has only unit demand. The bidding is therefore done only in prices, not quantities and the winner of the auction for the first item is not allowed to participate in bidding for the second item. The values assigned to participants are same in both periods, as well as the composition of the groups. The exception of course is the winning bidder from the first auction, so the second auction only consists of either 2 or 5 bidders.

Treatments A and B

After receiving her private value, each subject has to make an offer. In treatment A she is informed that she and 2 other randomly selected subjects are participating in the auction, whereas in treatment B subject participates with 5 other randomly selected subjects. She is equally informed as in treatment A, and all other conditions are the same. Subjects have to place a bid, however bid can be 0, which is also the lowest input limit. All bids are also sealed; the bidder does not see other participants' bids. After all the bids are collected, the highest and the second highest offer are recorded. The highest bid determines the winner, the second highest determines the price paid.

The only information that participants receive after first of the two units is sold is whether they won the auction or not, and consequentially, whether they are eligible to participate in the auction for the second good. Assuming that participants have common and monotonically increasing bidding function, disclosure of winner's bid would also disclose the highest bidder's valuation (type), whereas disclosed selling price would inform of the second highest bidder type. Since the later would also participate in the second bidding, the solution of the auction would be trivial. It can be easily shown that, knowing the highest type, the participants with lower valuations would have no incentive to place their bids.

In the second auction the participants remain of the same type and the identical object is offered. The only difference is, that they now place bids for auction with 2 or 5 bidders depending on a treatment. After the bids are placed, again the highest bid determines the winner and the second highest bid, the price paid. This time, all are informed whether they won an auction or not, what was the price paid by the winner in each auction and their own profits. The profit is determined as:

$$\Pi_i = \begin{cases} x_i - p, & \text{if } i \text{ is a winner} \\ 0, & \text{if } i \text{ is not a winner} \end{cases}$$

The actual payoff in euros was calculated as $0.5\Pi_i$. Both, the interval boundaries and the conversion rate were calculated with intention that the average payoff of each participants in both treatments (30 stages) is 5 euros. Including the 3 euros participation fee this makes each subjects average earnings of 8 euros. The two auction stages of treatments A and B are repeated 30 times.

Treatment C

Treatment C is in most aspects the same as treatments A and B. The distribution function of values remains the same, the interval as well. Here however, are bidders unsure of the real number of auction participants. Beside personal value, their only information is that there are either 3 or 6 bidders participating in the auction, both with 50% probability. 12 participants are divided to 4 or 2 groups with random matching. After finishing the first stage of each auction here as well, subjects are only informed about whether they won or not, but not also the winning bid. For the second stage the same rule applies, just that subjects are informed that there are either 2 or 5 participants in this stage. The payoff function remains the same. This treatment is also repeated through 30 stages. Appendix C contains an English version of the instructions used in the experiment. Figure C.1 depicts the experimental interface.

3.5 Results

Given the theoretical framework discussed above, we test several predictions of the theory. First, we expect the second unit bids on average to be higher than first unit bids. Second, we expect that prices will remain constant on average for all treatments with known number of bidders as well as in the treatment with uncertain number of bidders. Third, the same strategic behavior should lead the subjects to have the same bid/value ratio on average for the second unit in all implemented treatments.

3.5.1 First- versus second-unit bids

It is interesting to first have a look at the average bidding strategies for the first and the second unit. Regardless the treatment it is optimal to bid one's value for the second unit, and *less* than one's value for the first unit. We run t-statistics for every subject to verify whether the first unit bid is truly lower than the second unit bid on average. See the Table C.3 in Appendix C for details. There is a large proportion of people where we cannot demonstrate this.

A brief statistics is presented in the Table 3.1. Only 61% of the subjects in a 3-bidder treatment and 58% of them in a 6-bidder treatment increase their bids on average for the second unit. Only the cases where subject hasn't won the first unit are used for the comparison. This demonstrates that subjects were not really sure what is the best strategy to optimize their revenues.

Treatment	Obs.	2. bid $-$1. bid	St.Dev.	Increases
$n = 3$	360	2.504	5.300	61%
$n = 6$	400	1.444	4.077	58%
$n =$ undisclosed	348	1.540	2.818	83%

Table 3.1: Average differences between first and second unit bids. "Increases" represents the proportion of subjects where the second bid is on average significantly higher than the first bid. Based on Table C.3.

Another surprise that one may find is much higher proportion, 83%, of the subjects that increase their bids in the treatment with undisclosed number of bidders. In theory, the optimal strategy is harder to determine here, yet subjects behave more consistently with the "basic logic" represented by our hypothesis. We demonstrate below, that resulting prices are actually not more optimal in the treatment with undisclosed number of bidders.

3.5.2 First- versus second-unit prices

Calculating for the mean auction prices we find our results in line with previous research. Prices on average decline, and the decline is understandably fiercer when the number of bidders is low. Table 3.2 displays a basic summary of all three treatments:

			Observed prices		RNNE prices	
Treatment	Obs.		Unit 1	Unit 2	Unit 1	Unit 2
$n = 3$	180	mean	7.41	4.56	4.93	4.83
		std. dev.	(4.26)	(3.86)	(2.35)	(3.98)
$n = 6$	80	mean	12.41	10.87	11.27	11.29
		std. dev.	(3.00)	(3.37)	(2.71)	(3.54)
$n = undisclosed$	182	mean	8.76	5.81	7.83	7.12
		std. dev.	(2.54)	(2.81)	(2.82)	(3.20)

Table 3.2: Summary statistics of observed and realized equilibrium prices for all treatments. In the uncertainty case weighted according to realized β_3 and β_6.

The predictions made for prices in given experiment were 4.93 and 4.83 for the first and the second unit respectively in the treatment with 3 bidders initially, 11.27 and 11.29 for treatment with 6 bidders initially. For the treatment with uncertainty the prediction was 7.83 for the first unit and 7.12 for the second unit in the treatment with uncertainty. Experimental results show, that prices decline in all performed treatments. In a treatment with 3 bidders for 2.85, in a treatment with 6 bidders for 1.54 and in a treatment with uncertainty for 2.95. The higher decrease in a treatment with 3 bidders, compared to the one with 6 is not unanticipated. Optimal bid in the first was $\frac{1}{2}$ of the value, compared to $\frac{4}{5}$ in the later.

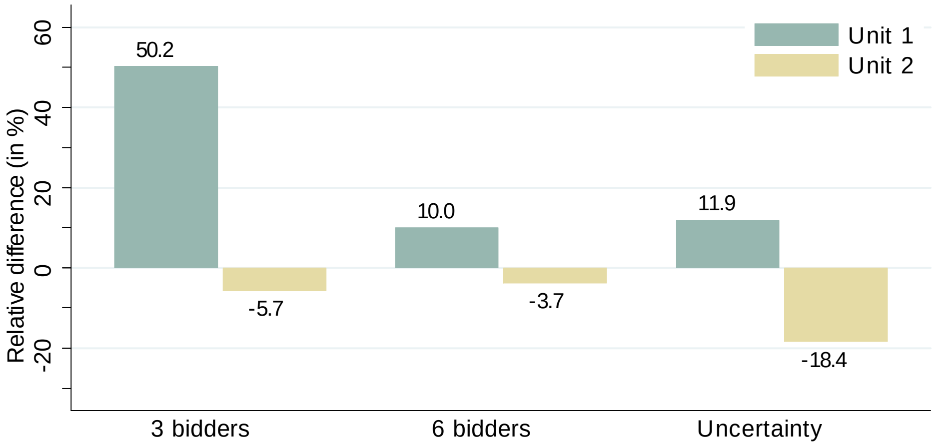

Figure 3.3: Difference between observed auction prices and expected RNNE prices expressed as percentage of RNNE price.

Subjects are apparently resilient to decrease bids (with respect to their personal values), and that manifests more when expected decrease is higher. In experiment this results in higher prices and corresponding larger divergence from risk neutral Nash equilibrium for the treatment with 3 bidders.

We observe that prices fall below equilibrium the most in a treatment with uncertainty, however we cannot say that the uncertainty is a sole reason for price declines. There are apparently several reasons why bidders' strategies are suboptimal and uncertainty about the number of bidders is most likely just one of them. To investigate the prices further, we display the difference between the mean prices for all units in all treatments and the corresponding risk neutral equilibrium predictions for these means in Figure 3.3.

As expected from the Table 3.2, Figure 3.3 shows that subjects "overbid" for the first unit, and "underbid" for the second. We see that bidders overbid approximately to the same amount in the treatment with 6 bidders as they do in the treatment with uncertainty. More interesting is to see the second unit bids. While decreases from RNNE bids (one's value) are minimal for the certainty case, they are quite substantial for the case of uncertainty. Figure 3.4 below clearly displays an increased proportion of sub-value bids for second unit in treatment with uncertainty compared to treatments A and B. These results are further confirmed by the Table C.6 in Appendix C, which counts the number of price decreases and increases. We see that price decline in 82% of the cases in treatment A, whereas only 60% of the cases in treatment B. Average decrease in the treatment with 3 bidders is 2.85 while only 1.54 in the treatment with 6 bidders initially. An interesting observation regarding the treatment with uncertainty is that people will more likely behave as if less bidders are present than equally distributing chances between 3 and 6. Table C.3 shows that price declines 83% of the cases, practically as frequently as in the treatment with 3 bidders. Average difference between the first and the second price also confirms that.

t-test	3 bidders	6 bidders	Uncertainty
$H_0 : q_1 = q_2$	9.70	5.43	12.24
	(0.000)	(0.000)	(0.000)
$H_0 : q_1 = q_1^*$	7.80	3.37	2.48
	(0.000)	(0.001)	(0.014)
$H_0 : q_2 = q_2^*$	-0.95	1.12	-6.51
	(0.341)	(0.267)	(0.000)

Table 3.3: Hypothesis tests on auction prices. For each hypothesis there is t-statistics in the first row and the corresponding two-side p-value in the second row.

Table 3.3 provides a verification of some of our hypotheses. We cannot confirm the null hypotheses about the equality of the first and the second unit prices for none of the three treatments. Equally we cannot confirm the null hypothesis that the average price for the first unit should not differ from the risk neutral equilibrium prediction. When comparing the second unit average prices and RNNE predictions the results are expectedly different. Here the null hypothesis of equality is not rejected for the treatments with 3 and 6 bidders, whereas for the case of uncertainty the hypothesis is rejected. This is just a formal proof for what is displayed in Figure 3.3. Uncertainty about the number of competitors apparently puzzles subjects and impedes their intuition about the optimal bidding strategy.

We also perform nonparametric tests of our hypotheses. Wilcoxon signed-rank test confirms the difference for treatment A with $z = 12.723$; treatment B with $z = 10.735$; and treatment C (uncertainty) with $z = 13.175$. All probabilities are 0.0000. This confirms the results obtained previously with t-statistics.

3.5.3 Individual bids

The above analysis tells us about general patterns in each treatment, it does not however, reveal much about the reasons for the observed phenomenon. In this part we address the issue of individual differences that might lead to differences in mean prices and suboptimal bidding.

The actual bidders' behavior is displayed in Figure 3.4. There are some notable differences between the treatments. We observe that the price declines are very similar on average in treatments with 3 bidders and in treatment with uncertainty. Above figures show that bidding nevertheless differs and bidders likely have different strategies. Table C.3 in Appendix C provides summary statistics for each subjects behavior including the t-statistics for the equality of mean bid for the first and the second unit. Around 40%

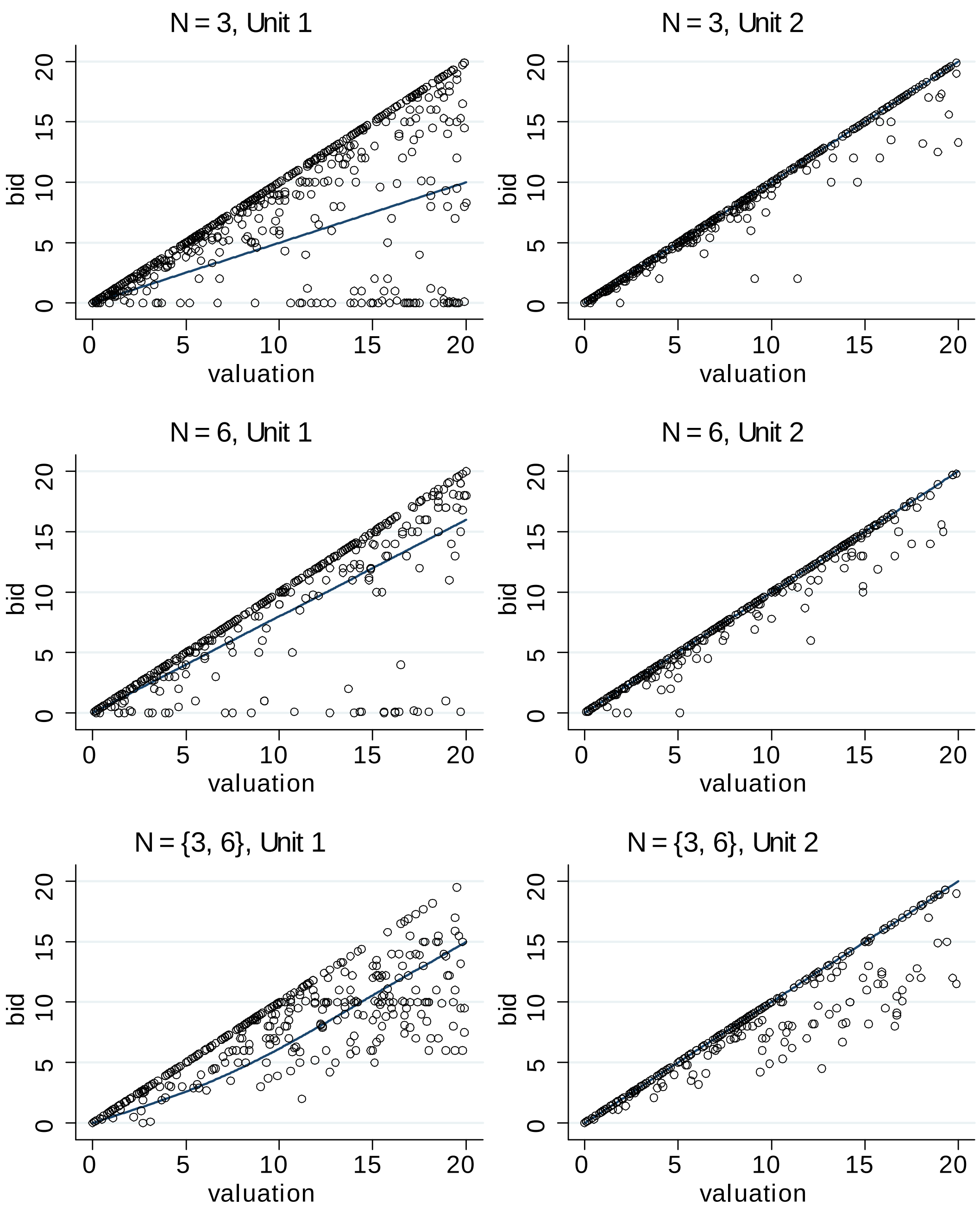

Figure 3.4: Individual bids for given personal values. Lines represent RNEE bid function.

of the subjects in the treatments with 3 and 6 bidders do not make significant different bids for the first and the second unit. And given the fact that large majority of subjects bid their value for the second unit, it is probably these 40% of the subjects who contribute to the high means for the first unit. Appendix C also contains scatter plots of selected subjects' bids. We can see in Figures C.2 to C.4 that there is a great diversity of bidding strategies.

One interesting fact that we observe is that some subjects tend to change their strategy, depending on their valuation being low or high. We define "low" values as those equal to or below 10 and "high" those above 10, 10 being the expected value of valuation support. This behavior is not so unreasonable, given that valuation 10 has only 25% chance of being the highest in the auction with 3 bidders, and 3.13% chance in the auction with 6 bidders competing.Mann–Whitney test on bid/equilibrium bid ratios of each subject

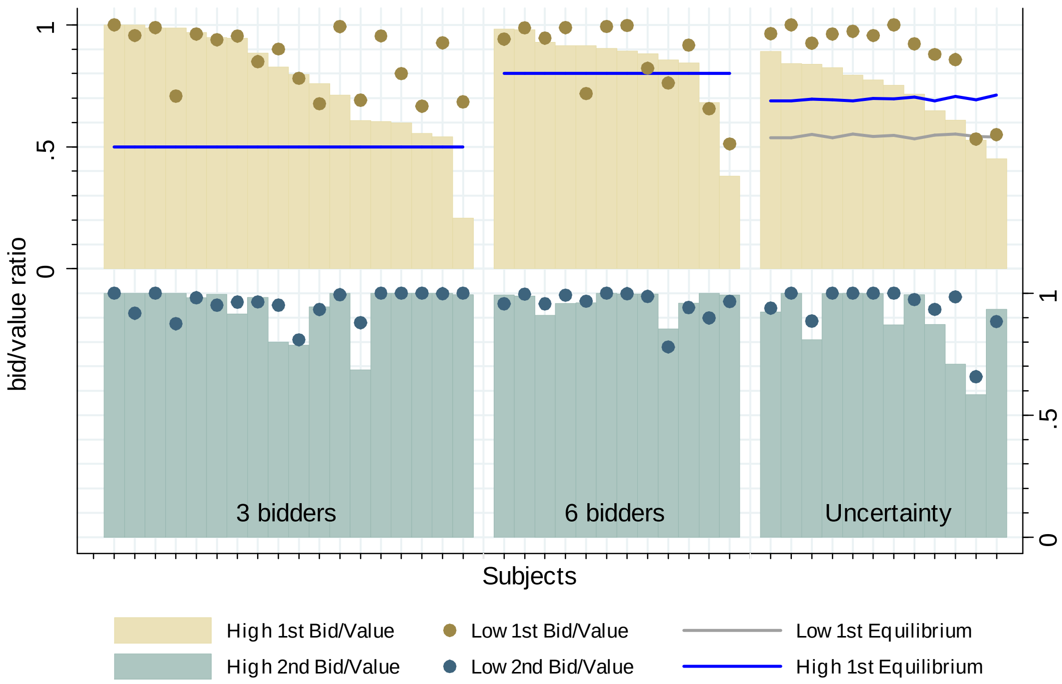

Figure 3.5: Average bid/value ratios per bidder. *1st Bid/Value* represents first unit's bid ratios, *2nd Bid/Value* second unit's bid ratios. *High* represents average ratios based on valuations greater than 10, *Low* represents the rest. Subjects ordered by *High 1st Bid/Value.*

shows that these differences indeed are important. 22% of subjects had significantly different low and high valuation bids for the first unit in a treatment with 3 bidders, 33% in the treatment with 6 bidders, and 100% in the treatment with uncertainty. In the last case, when we normalize bids with valuation rather than with equilibrium bid, tests still appear significant 75% of the times. We demonstrate these differences in Figure 3.5. Tests how well does subjects' bidding correspond to RNEE prediction are presented in Table C.4 and Table C.5 in Appendix C.

The most interesting are the ratios for the first unit in the uncertainty treatment. Here, it seems, subjects (correctly) believe that when they have high value bidding should be different than then when their value is low. Surprisingly however, subjects systematically bid higher proportion of their value when the latter is low, than when it is high. Equilibrium strategy suggests the opposite. A low highest value, say 5, is much more likely to result from an auction of 3 bidders (98.5% chance) than from the one of 6. In theory the bidding strategy should then be very similar to the one in auction with 3 bidders and bid/value ratio should be low. With higher values, chances of auction with 6 bidders increase (up to 50%) and corresponding bid/value ratio should increase as well. Subjects with low values apparently prefer to maximize the probability of winning the current unit, than to maximize the expected profit.

We can also observe that majority of subjects resort to a very simple strategy: bidding their valuation already in the first round. We also noted that almost 60% of the subjects had the same bidding strategy for both units in treatments with 3 and 6 bidders, whereas only 40% of subjects maintained their strategy in the treatment with uncertainty. There, also substantially more subjects had some form of contingency bidding strategy compared to the other two treatments.

One of the consequences of the heterogeneity of observed strategies is that auctions were not very efficient. First unit bidding resulted in 73.3%, 75% and 65% efficiency in the treatments with 3, 6 and uncertain number of bidders, respectively. On the contrary, second unit auctions were relatively efficient, in all treatments more than 90%. This goes along with the above results. Table C.1 in Appendix C describes the efficiency more in detail. The other viewpoint of mechanism design is seller's revenues. As depicted in Table C.2 uncertainty of bidders does not result in higher revenues for a seller, contrary to McAfee and McMillan (1987) conjecture. Revenues are significantly higher than expected by RNNE in the treatment with 3 bidders, whereas in the uncertainty treatment revenues per unit are even a bit lower than those predicted by RNNE.

3.6 Conclusion

Understanding bidding behavior is an important issue in modern auction design. Spread of modern auctions backed with all-electronic platforms has increased the number of people who exchange goods or services through this mechanism. Popularity of auctions and ease of participation has raised new issues for the auction science as well. For example, internet auction houses are most frequently characterized with: large number of items offered (and bided), various auction mechanisms, variable and unpredictable demand, minimal involvement of the auction house in the intermediation, feedback reputation mechanisms, etc. This all goes in contrast with traditional auctions with higher participation costs, limited and known number of expert bidders and active involvement of the auction house. We can claim that behavioral characteristics of auction participants are more important nowadays than in the past and uncertainty about the number of competitive bidders is a good example of that.

The present chapter discusses price and bidding behavior in the multiunit auctions with single unit demand. By running several laboratory experiments we test the theoretical prediction that prices should stay constant in an auction for 2 units sold in a sequence with a given number of bidders. We contrast this with a situation where bidders are uncertain about the size of their competition and compare the differences in price variation. Also in the latter case prices should remain constant on average. The intuition for martingale prices is based on desire of bidders to participate in the bidding for the second unit as well. There the competition will be less fierce, and price will be potentially lower. To increase the chances to participate in the bidding for the later unit(s) bidders will decrease their bids for the first unit, but only to extent where their joint expected profits are maximal. In the case of uncertainty these strategies are harder to calculate and bidders may be inclined to prefer the first unit to the later ones.

Our results show that neither with 3 or 6 bidders, nor with auctions with uncertain number of bidders, prices remain the same for the two units sold in a sequence. Prices for the first unit are significantly higher than for the second. In our experiment the differences are 2.85, 1.54 and 2.95 for 3 bidders, 6 bidders and uncertainty respectively. This goes in line with empirical papers such as of Ginsburgh (1998), or experimental works like Keser and Olson (1996) or Wang (2006). In our experiment prices fall regardless the treatment, although the effect is stronger in the case of uncertainty. We also find that subjects on average never bid according to the risk neutral Nash equilibrium, except for the second unit if they are certain about the number of participants. Their bid for the first unit is usually much higher than optimal. In the treatment with uncertainty many subjects tend to bid lower than their value for the second unit, which comes as a surprise. We also analyzed each subjects bidding strategies separately. Results show that high proportion of subjects bidding their value for the first unit, 50%, 67% and 25% for treatments with 3, 6 and uncertain number of bidders respectively, is the main reason for average prices being above the risk neutral Nash equilibrium. Interestingly, the uncertainty of the number of bidders caused that less people were bidding their value for the first unit. In this treatment also only 17% of the subjects bid the same for the two units, whereas around 40% of the subjects were not changing their bid significantly in the other two treatments. Despite the observed pattern in the treatment with uncertainty seems more consistent with the equilibrium for the first unit sold, it also displays an unexpected proportion of bids below one's value in the bidding for a second unit. This contributed to higher price decline than the other two treatments.

References

Adam, K. (2007). Experimental evidence on the persistence of output and inflation. *Economic Journal*, 117(520):603–636.

Anufriev, M. and Hommes, C. (2008). Evolutionary selection of individual expectations and aggregate outcomes. Mimeo, University of Amsterdam.

Arifovic, J. and Sargent, T. J. (2003). Laboratory experiments with an expectational phillips curve. In Altig, D. E. and Smith, B. D., editors, *Evolution and Procedures in Central Banking*. Cambridge University Press, Cambridge.

Ashenfelter, O. (1989). How auctions work for wine and art. *Journal of Economic Perspectives*, 3:23–36.

Ashenfelter, O. and Graddy, K. (2003). Auctions and the price of art. *Journal of Economic Literature*, 41:763–787.

Assenza, T., Heemeijer, P., Hommes, C., and Massaro, D. (2011). Individual expectations and aggregate macro behavior. CeNDEF Working Papers 11-01, Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance.

Baak, S. J. (1999). Tests for bounded rationality with a linear dynamic model distorted by heterogeneous expectations. *Journal of Economic Dynamics and Control*, 23(9-10):1517–1543.

Bakhshi, H. and Yates, A. (1998). Are u.k. inflation expectations rational? Technical Report 81, Bank of England Working Paper.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Batchelor, R. and Dua, P. (1996). Empirical measures of inflation uncertainty: A cautionary note. *Applied Economics*, 28(3):333–41.

Beggs, A. and Graddy, K. (1997). Declining values and the afternoon effect: Evidence from art auctions. *Rand Journal of Economics*, 28:3:544–65.

Ben-David, I., Graham, J. R., and Harvey, C. R. (2010). Managerial miscalibration. NBER Working Papers 16215, National Bureau of Economic Research, Inc.

Berlemann, M. and Nelson, F. (2005). Forecasting inflation via experimental stock markets some results from pilot markets. Ifo Working Paper Series Ifo Working Paper No. 10, Ifo Institute for Economic Research at the University of Munich.

Bernanke, B. (2007). Inflation expectations and inflation forecasting. Technical report, Speech at National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts.

Bernasconi, M. and Kirchkamp, O. (2000). Why do monetary policies matter? an experimental study of saving and inflation in an overlapping generations model. *Journal of Monetary Economics*, 46(2):315–343.

Bernhardt, D. and Scoones, D. (1994). A note on sequential auctions. *American Economic Review 84:3*, 84:3:653–57.

Biais, B., Hilton, D., Mazurier, K., and Pouget, S. (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *Review of Economic Studies*, 72(2):287–312.

Black, J. and de Meza, D. (1992). Systematic price differences between successive auctions are no anomaly. *Journal of Economics and Management Strategy*, 1:607–628.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3):623–685.

Bloom, N., Floetotto, M., and Jaimovich, N. (2010). Really uncertain business cycles. Mimeo, Stanford University.

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

Boero, G., Smith, J., and Wallis, K. F. (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 Devetag, G. (2005). Expectations structure in asset pricing experiments. In Lux, T., Reitz, S., and Samanidou, E., editors, *Nonlinear Dynamics and Heterogeneous Interacting Agents*, Lecture Notes in Economics and Mathematical Systems. Springer Verlag, Berlin, 550 edition.

Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal*, 114(497):592–621.

Branch, W. A. (2007). Sticky information and model uncertainty in survey data on inflation expectations. *Journal of Economic Dynamics and Control*, 31(1):245–276.

Branch, W. A. and Evans, G. W. (2006). Intrinsic heterogeneity in expectation formation. *Journal of Economic Theory*, 127(1):264–295.

Branch, W. A. and McGough, B. (2008). Replicator dynamics in a cobweb model with rationally heterogeneous expectations. *Journal of Economic Behavior and Organization*, 65(2):224–244.

Branch, W. A. and McGough, B. (2009). A new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control*, 33(5):1036–1051.

Brock, W. A. and Hommes, C. H. (1997). A rational route to randomness. *Econometrica*, 65(5):1059–1096.

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

Bryan, M. F. and Palmqvist, S. (2005). Testing near-rationality using detailed survey data. Working Paper Series 183, Sveriges Riksbank.

Burns, P. (1985). Experience and decision making: A comparison of students and businessmen in a simulated progressive auction. *Research in Experimental Economics: A Research Annual. Vol. 3. Vernon Smith, ed. Greenwich*, pages 139–57.

Carceles-Poveda, E. and Giannitsarou, C. (2007). Adaptive learning in practice. *Journal of Economic Dynamics and Control*, 31(8):2659–2697.

Carroll, C. D. (2003a). Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics*, 118(1):269–298.

Carroll, C. D. (2003b). The epidemiology of macroeconomic expectations. NBER Working Papers 8695, National Bureau of Economic Research, Inc.

Chanel, O., Gérard-Varet, L.-A., and Vincent, S. (1996). *Auction Theory and Practice: Evidence from the Market for Jewellery*. Elsevier, Amsterdam. in Economics of the Arts: Selected Essays eddited by Victor Ginsburgh and Pierre-Michel Menger.

Chavas, J.-P. (2000). On information and market dynamics: The case of the u.s. beef market. *Journal of Economic Dynamics and Control*, 24(5-7):833–853.

Conover, W. J. (1999). *Practical Nonparameteric Statistics*. Wiley, 3rd edition.

Cooley, T. F. and Prescott, E. C. (1995). Economic growth and business cycles. In Cooley, T. F., editor, *Frontiers of Business Cycle Research*. Princeton University Press, Princeton.

Cox, J. C., Smith, V. L., and Walker, J. M. (1988). Theory and individual behavior in first-price auctions. *Journal of Risk and Uncertainty*, 1:61–99.

Curtin, R. (2005). Inflation expectations: Theoretical models and empirical tests. Mimeo, University of Michigan.

D'Amico, S. and Orphanides, A. (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 Budescu, D. V. (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.

Dyer, D., Kagel, J. H., and Levin, D. (1989). Resolving uncertainty about the number of bidders in independent private-value auctions: An experimental analysis. *RAND Journal of Economics*, 20:268–79.

Engelberg, J., Manski, C. F., and Williams, J. (2009). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business & Economic Statistics*, 27:30–41.

Engelbrecht-Wiggans, R. (1994). Sequential auctions of stochastically equivalent objects. *Economic Letters*, 44:1-2:87–90.

Engelmann, D. and Grimm, V. (2006). Bidding behavior in multi-unit auctions Ű an experimental investigation. Working Paper, 24.

Engle-Warnick, J. and Turdaliev, N. (2010). An experimental test of taylor-type rules with inexperienced central bankers. *Experimental Economics*, 13(2):146–166.

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

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

Evans, G. W., Honkapohja, S., and Williams, N. (2010). Generalized stochastic gradient learning. *International Economic Review*, 51(1):237–262.

Evans, M. and Wachtel, P. (1993). Inflation regimes and the sources of inflation uncertainty. *Journal of Money, Credit and Banking*, 25(3):475–511.

Fehr, E. and Tyran, J.-R. (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.

Friedman, M. (1977). Nobel lecture: Inflation and unemployment. *Journal of Political Economy*, 85(3):451–72.

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

Ginsburgh, V. (1998). Absentee bidders and the declining price anomaly in wine auctions. *Journal of Political Economy*, 106:6:1302–19.

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

Giordani, P. and Söderlind, P. (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.

Harstad, R. M., Kagel, J. H., and Levin, D. (1990). Equilibrium bid functions for auctions with an uncertain number of bidders. *Economics Letters*, 33:35–40.

Haruvy, E., Lahav, Y., and Noussair, C. N. (2007). Traders' expectations in asset markets: Experimental evidence. *American Economic Review*, 97(5):1901–1920.

Hazelett, D. and Kernen, A. (2002). Hyperinflation and seigniorage in an experimental overlapping generations economy. Mimeo, Whitman College.

Heemeijer, P., Hommes, C., Sonnemans, J., and Tuinstra, J. (2009). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic Dynamics and Control*, 33(5):1052–1072.

Hoffrage, U. (2004). Overconfidence. In Pohl, R. F., editor, *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*, pages 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., Huang, H., and Wang, D. (2005a). A robust rational route to randomness in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 29(6):1043–1072.

Hommes, C., Sonnemans, J., Tuinstra, J., and van de Velden, H. (2005b). 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.

Kagel, J. H. and Levin, D. (2001). Behavior in multi-unit demand auctions: Experiments with uniform price and dynamic auctions. *Econometrica*, 69:2:413–454.

Kelley, H. and Friedman, D. (2008). Learning to forecast rationally. In Plott, C. R. and Smith, V. L., editors, *Handbook of Experimental Economics Results*, volume 1, chapter 35, pages 303–310. Elsevier.

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

Keser, C. and Olson, M. (1996). *Experimental Examination of the Declining price Anomaly*. Elsevier, Amsterdam. in Economics of the Arts: Selected Essays eddited by Victor Ginsburgh and Pierre-Michel Menger.

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

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

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

Lei, V., Noussair, C. N., and Plott, C. R. (2001). Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica*, 69(4):831–59.

Levi, M. D. and Makin, J. H. (1980). Inflation uncertainty and the phillips curve: Some empirical evidence. *American Economic Review*, 70(5):1022–27.

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

List, J. A. and Lucking-Reiley, D. (2000). Demand reduction in multiunit auctions: Evidence from a sportscard field experiment. *American Economic Review*, 90:4:961–972.

Liu, F. and Lahiri, K. (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. *Review of Economic Studies*, 77(1):305–338.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4):1295–1328.

Mankiw, N. G., Reis, R., and Wolfers, J. (2004). Disagreement about inflation expectations. *NBER Macroeconomics Annual 2003*, 18:209–248.

Marcet, A. and Sargent, T. J. (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., Spear, S. E., and Sunder, S. (1993). Expectationally driven market volatility: An experimental study. *Journal of Economic Theory*, 61(1):74–103.

Marimon, R. and Sunder, S. (1993). Indeterminacy of equilibria in a hyperinflationary world: Experimental evidence. *Econometrica*, 61(5):1073–107.

Marimon, R. and Sunder, S. (1995). Deoes a constant money growth rule help stabilize inflation: Experimental evidence. *Carnegie - Rochester Conference Series on Public Policy*, 45:111–156.

Matthews, S. (1987). Comparing auctions for risk averse buyers: A buyerŠs point of view. *Econometrica*, 55:633–646.

McAfee, R. P. and McMillan, J. (1987). Auctions with a stochastic number of bidders. *Journal of Economic Theory*, 43:1–19.

McAfee, R. P. and Vincent, D. (1993). The declining price anomaly. *Journal of Economic Theory*, 60:1:191–212.

McCallum, B. T. and Nelson, E. (2004). Timeless perspective vs. discretionary monetary policy in forward-looking models. *Federal Reserve Bank of St. Louis Review*, 86(2):43–56.

Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics*, 54(7):2065–2082.

Milgrom, P. R. and Weber, R. J. (2000). *A theory of auctions and competitive bidding II*. Edward Elgar, Cheltenham, U.K. in The Economic Theory of Auctions eddited by P. Klemperer.

Mishkin, F. S. (2008). Whither federal reserve communications. *Speech at the Peterson Institute for International Economics*. Washington, D.C., July 28, 2008.

Molnár, K. (2007). Learning with expert advice. *Journal of the European Economic Association*, 5(2-3):420–432.

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

Muto, I. (2011). Monetary policy and learning from the central bank's forecast. *Journal of Economic Dynamics and Control*, 35(1):52–66.

Neugebauer, T. and Pezanis-Christou, P. (2007). Bidding behavior at sequential first-price auctions with(out) supply uncertainty: A laboratory analysis. *Journal of Economic Behavior and Organization*, 63(1):55–72.

Noussair, C. N., Pfajfar, D., and Zsiros, J. (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.

Nunes, R. (2009). Learning the inflation target. *Macroeconomic Dynamics*, 13(02):167–188.

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

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

Orphanides, A. and Williams, J. C. (2005a). The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations. *Journal of Economic Dynamics and Control*, 29(11):1927–1950.

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

Pekec, A. and Tsetlin, I. (2008). Revenue ranking of discriminatory and uniform auctions with an unknown number of bidders. *Management Science*, 54:9:1610–1623.

Pesaran, M. H. (1987). *The Limits to Rational Expectations*. Basil Blackwell, Oxford, reprinted with corrections 1989 edition.

Pfajfar, D. (2008). Formation of rationally heterogeneous expectations. Mimeo, University of Tilburg.

Pfajfar, D. and Santoro, E. (2008). Credit market distortions, asset prices and monetary policy. Cambridge Working Papers in Economics 0825, Faculty of Economics, University of Cambridge.

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

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

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

Rich, R. and Tracy, J. (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 Butler, J. S. (1998). Disagreement as a measure of uncertainty: A comment on bomberg. *Journal of Money, Credit and Banking*, 30(3):pp. 411–419.

Roos, M. W. and Luhan, W. J. (2008). Are expectations formed by the anchoring-and-adjustment heuristic? an experimental investigation. Ruhr Economic Papers 0054, Rheinisch-Westfälisches Institut für Wirtschaftsforschung.

Sargent, T. J. (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press.

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

Schmalensee, R. (1976). An experimental study of expectation formation. *Econometrica*, 44(1):17–41.

Smith, V. L., Suchanek, G. L., and Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, 56(5):1119–51.

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 edition.

Wang, J. T.-y. (2006). The ebay market as sequential second price auctions-theory and experiments. Working Paper.

Weber, R. (1983). *Multiple object auctions*. New York University Press, New York. in Auctions, Bidding, and Contracting: Uses and Theory, eddited by R. Engelbrecht-Wiggans, M. Shubik and R.M. Stark.

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.

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

Yaniv, I. and Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General*, 124(4):424 – 432.

Zarnowitz, V. and Lambros, L. A. (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.

Appendix A

Additional information for Chapter 1

A.1 Tables and figures

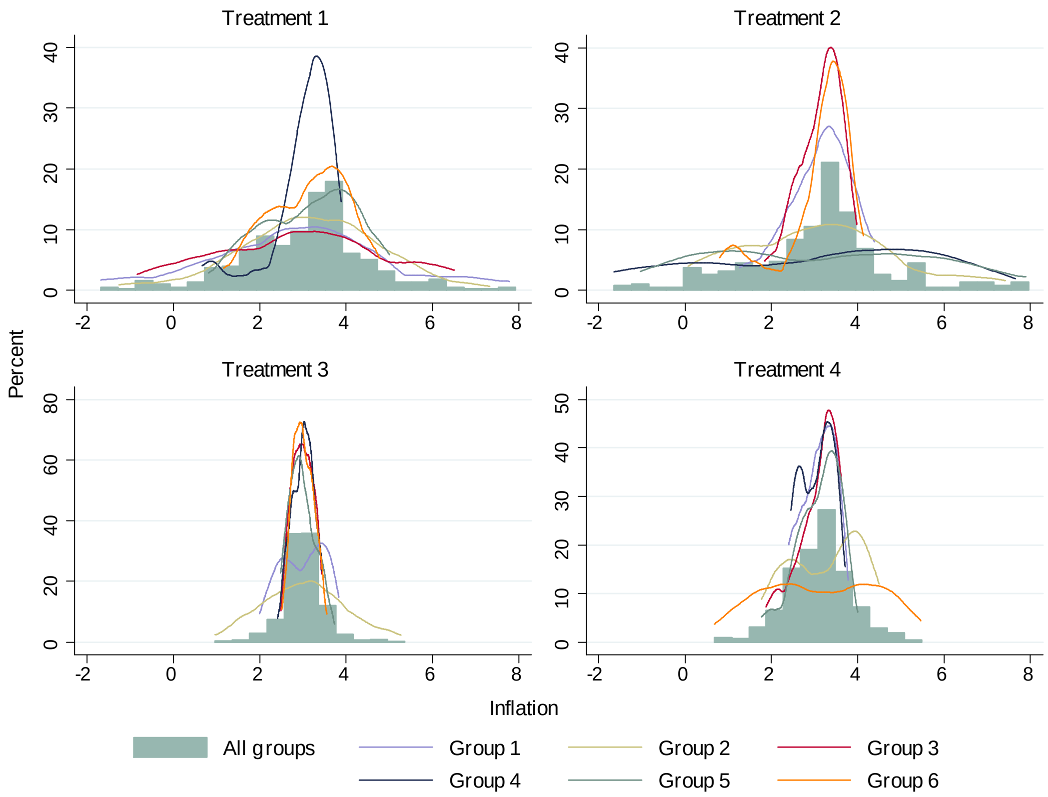

Figure A.1: Histogram of individual inflation forecasts by treatment and kernel density functions of each independent group in a given treatment.

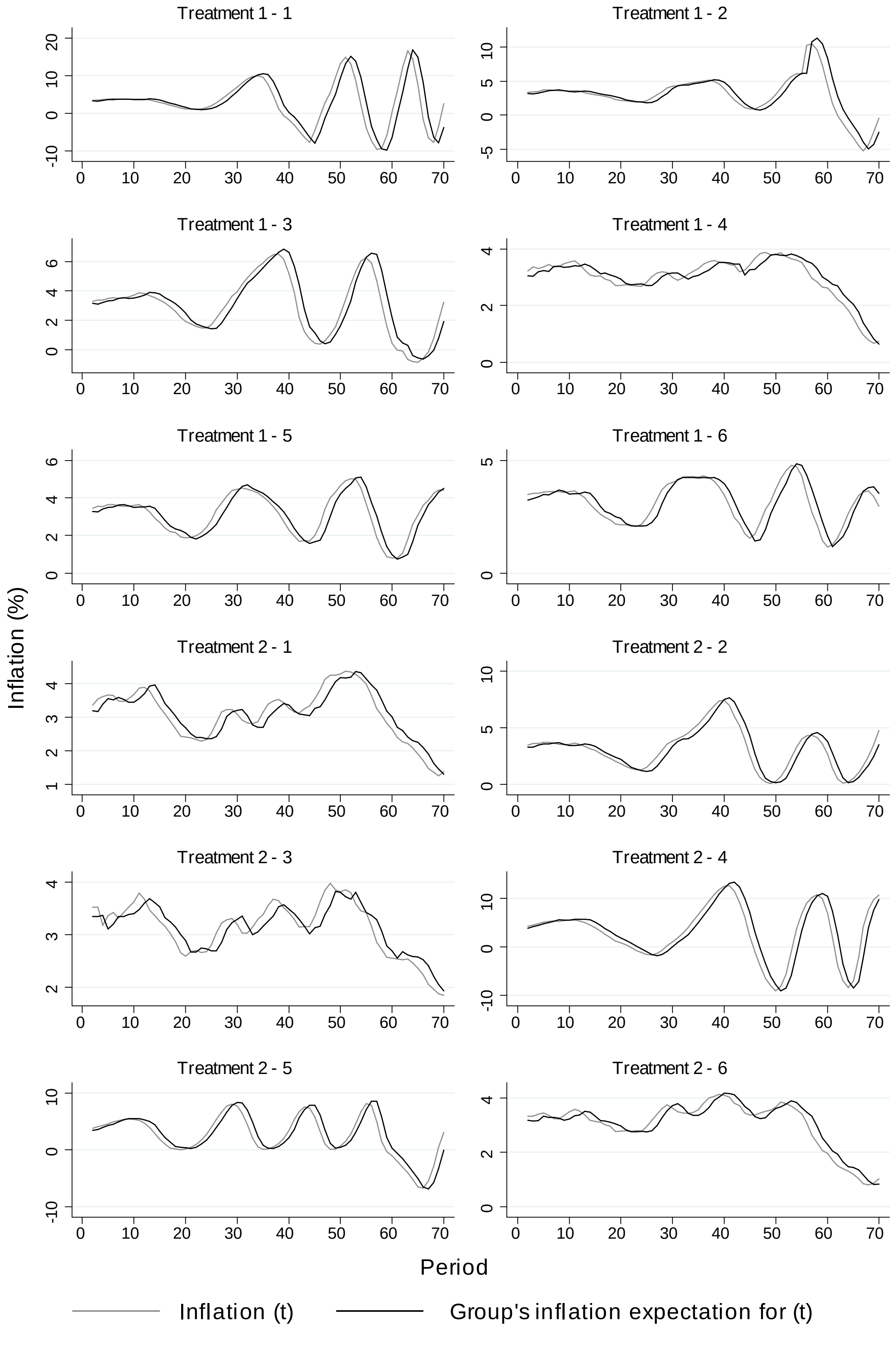

Figure A.2: Inflation and inflation expectations per group, Part 1

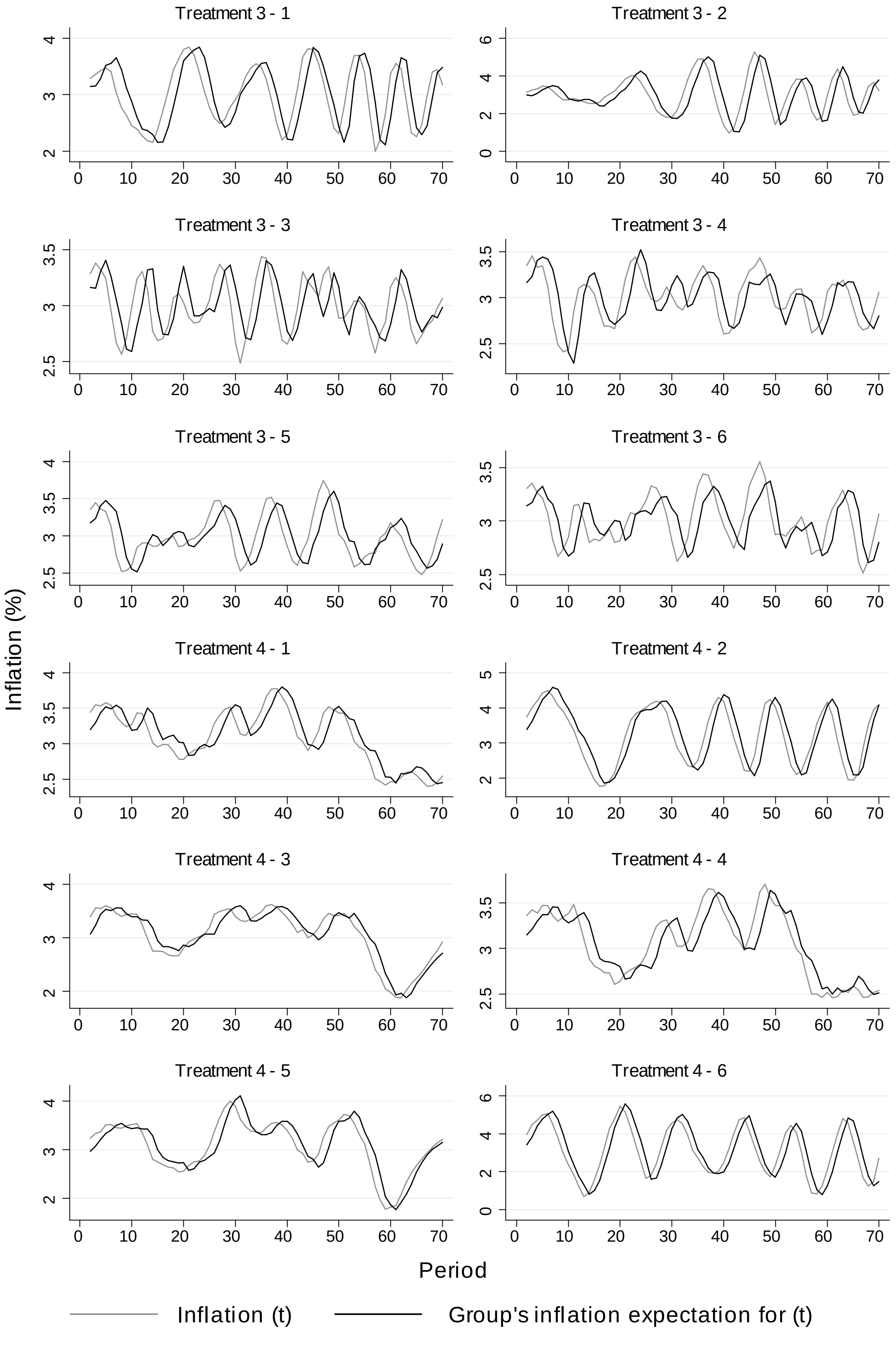

Figure A.3: Inflation and inflation expectations per group, Part 2

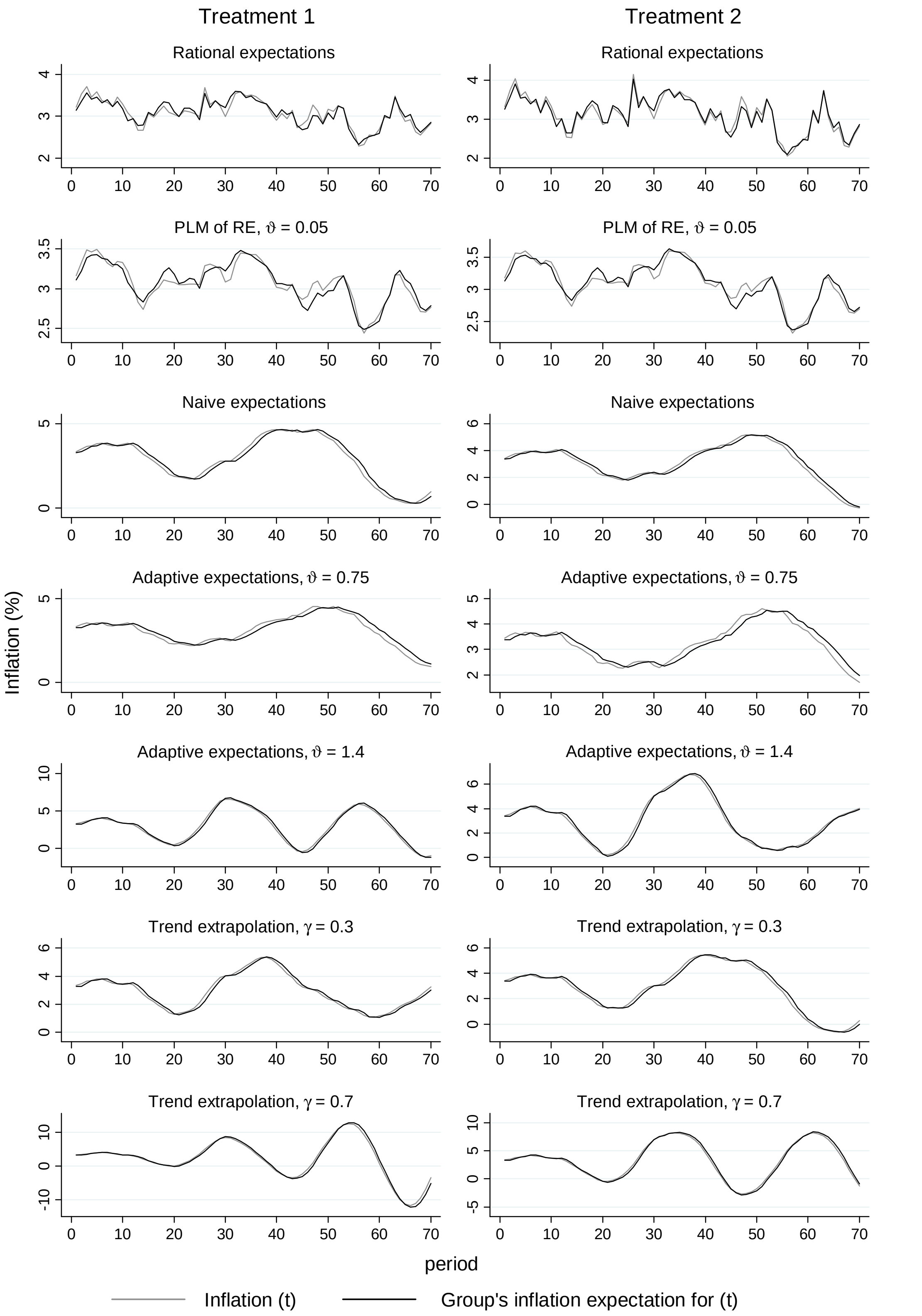

Figure A.4: Alternative expectation formation rules (treatments 1 and 2).

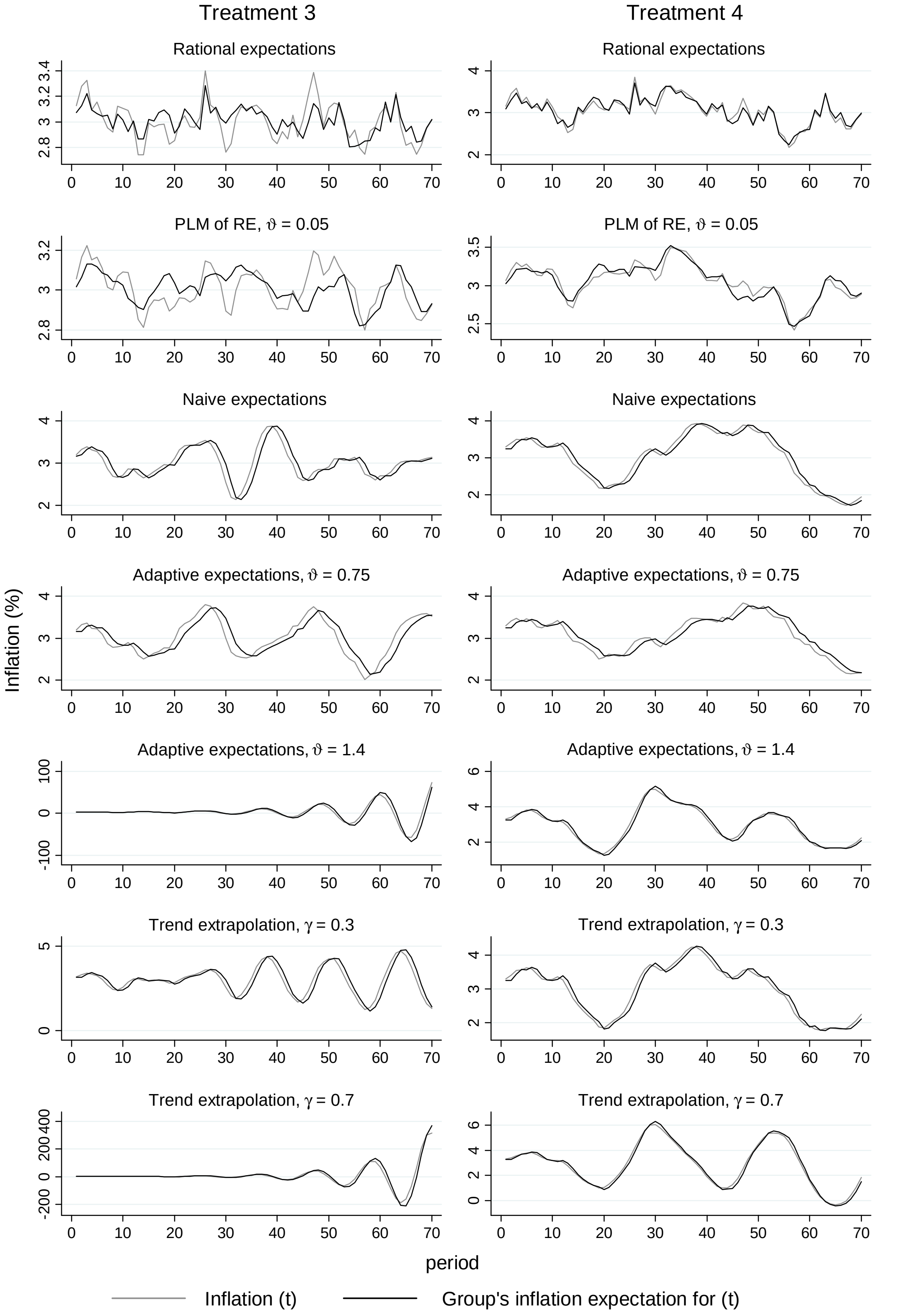

Figure A.5: Alternative expectation formation rules (treatments 3 and 4).

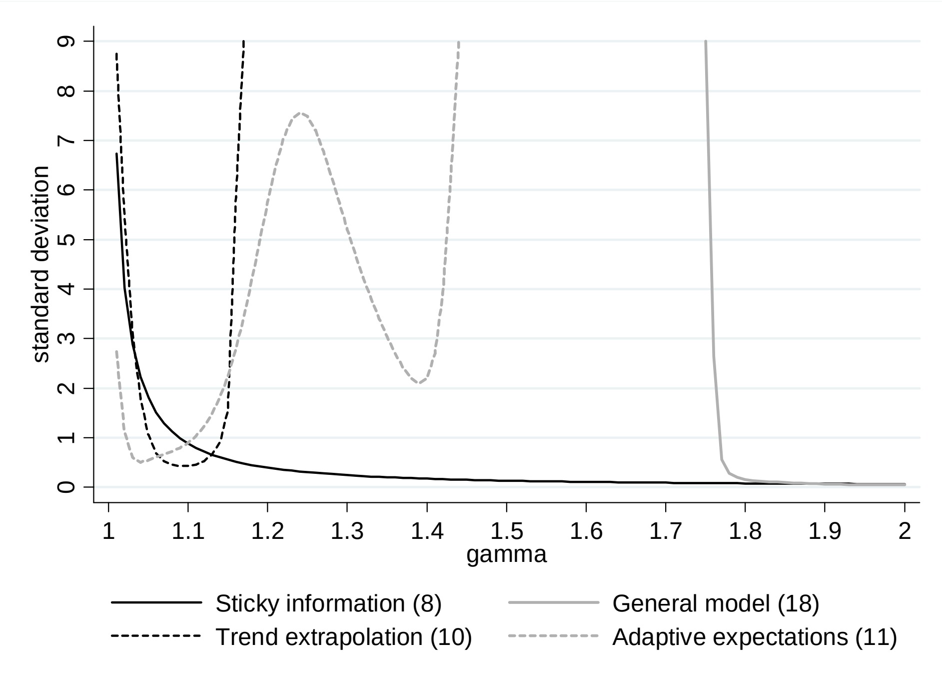

Figure A.6: Variability of inflation and alternative expectation formation rules (Simulated values for inflation forecast targeting).

model (eq.)	1	2	3	4	All
Rational expectations (1.7)	35.2	48.1	5.6	25.9	28.7
AR(1) process (1.19)	0.0	0.0	0.0	1.9	0.5
Sticky information type (1.8)	3.7	1.9	16.7	3.7	6.5
Adaptive expectations (1.11)	9.3	3.7	7.4	9.3	7.4
Trend extrapolation (1.10)	35.2	25.9	25.9	33.3	30.1
Recursive - lagged inflation (1.13)	3.7	5.6	24.1	13.0	11.6
Recursive - REE (1.14)	0.0	1.9	9.3	0.0	2.8
Recursive - trend extrapolation (1.16)	0.0	0.0	0.0	1.9	0.5
Recursive - AR(1) process (1.15)	13.0	13.0	11.1	11.1	12.0

Table A.1: Inflation expectation formation (percent of subjects, Comparison 1)

A.2 Instructions used in experiment

Thank you for participating in this experiment, a project in economic investigation. Your earnings depend on your decisions and the decisions of the other participants. There is a show-up fee of 4 Euros assured. From now on until the end of the experiment you are not allowed to communicate with each other. If you have a question raise your hand and one of the instructors will answer the question in private. Please do not ask aloud.[1]

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 profit depends 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 the general rise in prices in the economy. In each period it depends on the inflation expectations of the agents in the economy (you and the other 8 participants in this experiment), the output gap and small random shocks.
- The **output gap** measures by how much (in %) 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 potential level. In each period it depends on the inflation expectations of the agents in the economy, the past output gap, the interest rate and small random shocks.which have equal probability of having a positive or negative effect on inflation and are normally distributed.

[1] Instructions to participants in experiment at Universitat Pompeu Fabra are originally in Spanish. In experimental sessions, they were accompanied with the screenshot of the experimental interface (Figure A.7) and the profit table with earnings for various combinations of *estimation error* and *confidence interval*.

- The **interest rate** is (in this experiment) the price of borrowing money (in %) for one period. The interest rate is set by the monetary authority. Their decision mostly depends on the inflation expectations of the agents in economy.

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

- You enter the economy in period 1. In this period you will be given computer-generated past values of inflation, the output gap and the interest rate for 10 periods back (Called: -9, -8, ... -1, 0)
- In period 2 you will be given all the 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 the 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 the 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 payoff will depend on the accuracy of your prediction of inflation in the future period. In each period your prediction will consist of two parts:

1. *Expected inflation*, (in %) that you expect in the NEXT period ($Exp.Inf.$)
2. The *Confidence interval* ($Conf.Int.$) around your prediction for which you think there is 95% probability that the actual inflation will fall into. The interval is determined as the number of percentage points for which the actual inflation can be higher or lower.

Example 2 *Let's say you think that inflation in the next period will be* 3.7%. *And you also think it is most likely (*95% *probability) that the actual inflation will not differ from that value by more than* 0.7 *percentage points. Therefore, you expect that there is* 95% *probability that actual inflation in the next period will be between* 3.0% *and* 4.4% *(*$3.7\% \pm 0.7\%$*). Your inputs in the experiment will be* 3.7 *under 1) and* 0.7 *under 2).*

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

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

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

$$Exp.Inf.-Conf.Int. \leq Inflation \leq Exp.Inf.+Conf.Int.$$

and 0 otherwise.

This expression tells you, that x will be 1, if actual inflation falls between $Exp.Inf.-Conf.Int.$ (3.0% in our example) and $Exp.Inf.+Conf.Int.$ (4.4% in our example).

The *first part* of the payoff function states that you will receive some payoff if the actual value in the next period differs 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. 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 no larger than ± 4 percentage points. The more certain of the actual value you are, the smaller interval you give, and the higher your payoff will be if the actual inflation indeed is in the given interval, but there will also be a greater chance that the actual value falls outside your interval. In our example this interval is ± 0.7 percentage points. If the actual inflation falls in this interval you receive 38.8 units ($100/(1+0.7)-20$) in addition to the payoff from the first part of the payoff function. If the actual value is outside your interval, your receive 0.

On the attached sheet you can find a table which shows various combinations of *forecast error* and *confidence interval* needed to earn a given number of points. See also the figure on the next page.

Information after each period

Your payoff depends on your predictions for the next period and the actual realization in the next period. Because the actual inflation will only be known 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 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 is in your expected interval and what your earnings for this period are.

The units in the experiment are fictitious. Your actual payoff will be the sum of profits from all the periods converted to euros in 1/500 conversion.

If you have any questions please ask them now!

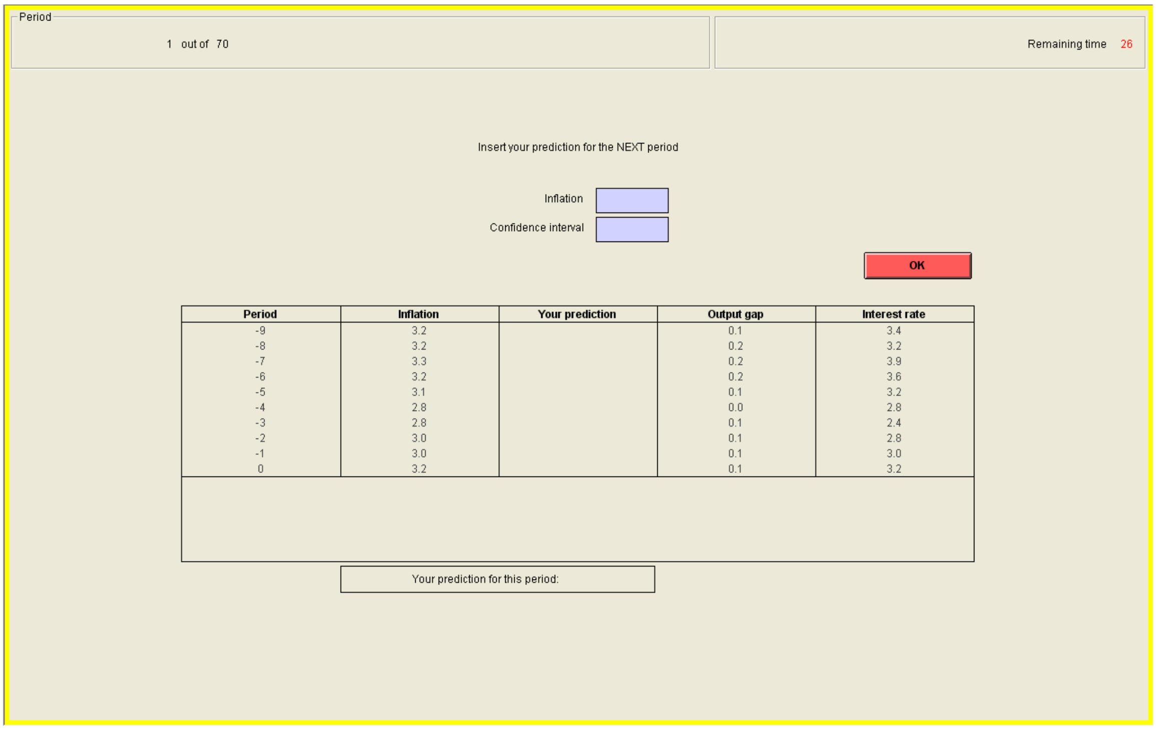

Figure A.7: Screenshot of the experimental interface

Questionnaire[2]

1. If you believe that inflation in the next period will be _4.2%_, and you are quite sure that it will be higher than _3.5%_ and lower than _4.9%_, you will type:
   Under (1) _ _ _ _ _ _ _ _ _ for inflation, and
   Under (2) _ _ _ _ _ _ _ _ _ for confidence interval.

2. You are now in period _ _ _ 10 _ _ _. You have information about past inflation, the output gap and the interest rate up to period _ _ _ _ _ _ _ _ _ and you have to predict the inflation for period _ _ _ _ _ _ _ _ _.

[2]Options (1) and (2) point to the different fields on the screenshot of the experimental interface.

Appendix B

Additional information for Chapter 2

B.1 Tables and figures

In Table B.9 we estimate the following regression using the system GMM estimator of Blundell and Bond (1998) for dynamic panel data:

$$r^k_{t+1} = \alpha + \beta r^k_t + \gamma sip^k_{t+1|t} + u^{em}_t.$$

Tables B.10 and B.11 report results of the following regressions:

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

where $D_7 = 1$ if the upper interval (C_U) has exactly the same width as the lower interval (C_L) 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 B.10 displays the results of logit estimations, while Table B.11 presents the results of Poisson estimations.

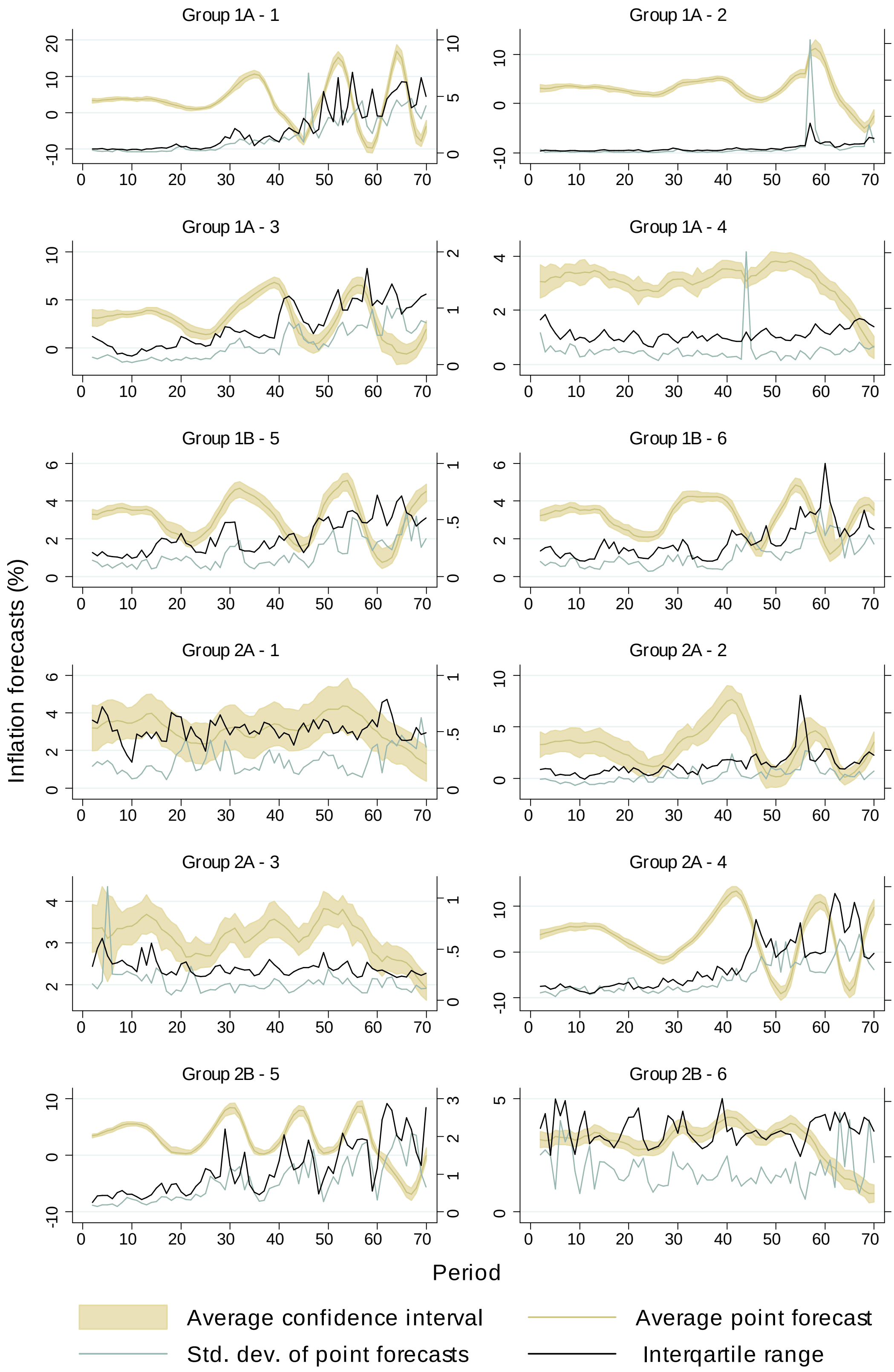

Figure B.1: Average inflation forecasts and average confidence intervals (left axis) and disagreement and uncertainty measures (right axis) per group. Interquartile range is calculated from the aggregate expectation distribution as described in Subsection 2.4.1.

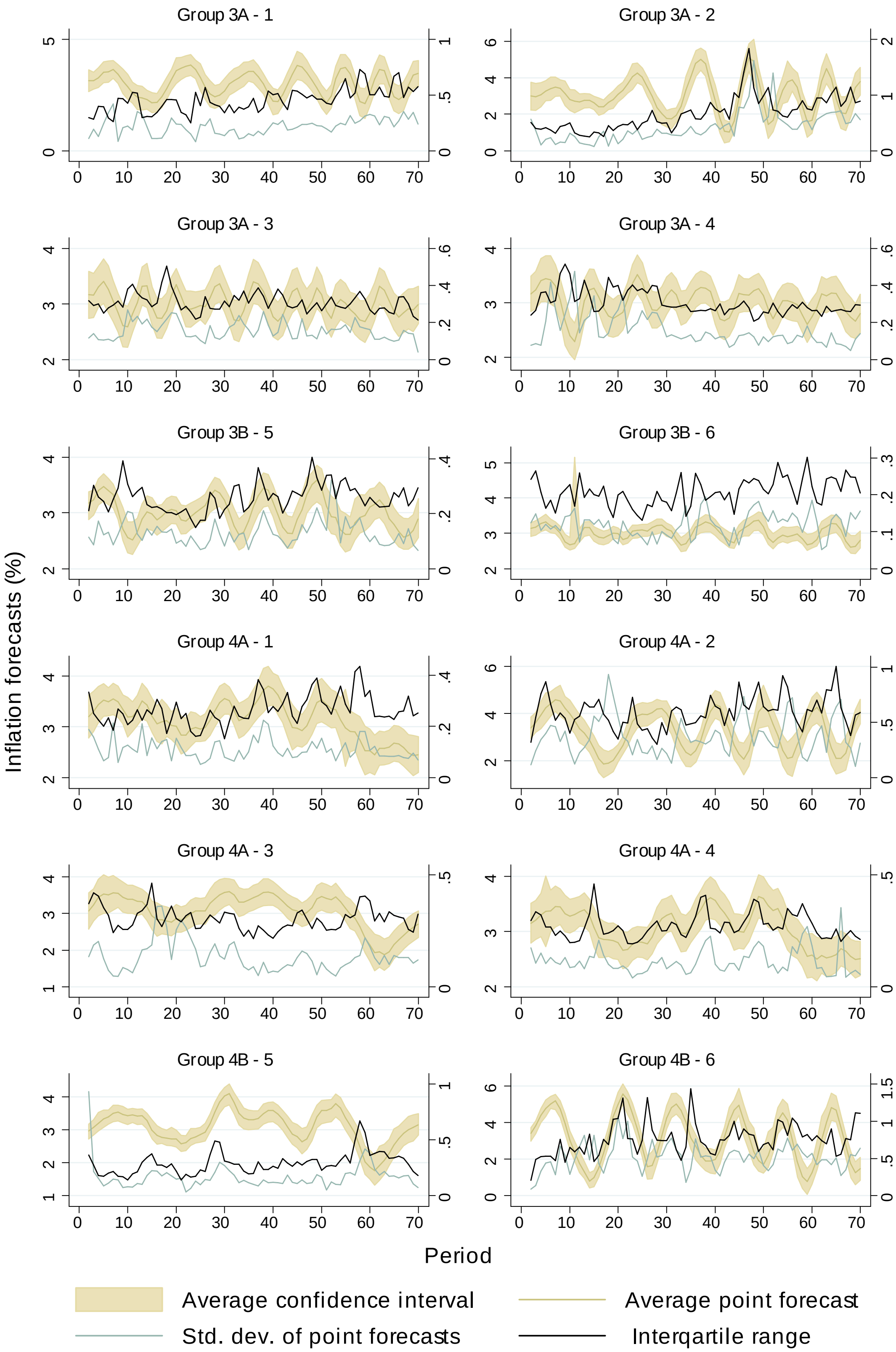

Figure B.2: Average inflation forecasts and average confidence intervals (left axis) and disagreement and uncertainty measures (right axis) per group. Interquartile range is calculated from the aggregate expectation distribution as described in Subsection 2.4.1.

		Confidence bound								
		Symmetric			Lower			Upper		
Treat	Group	$<$	$\approx$	$>$	$<$	$\approx$	$>$	$<$	$\approx$	$>$
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 B.1: Percentage of subjects by group with underprediction/overprediction of confidence interval. Note: the benchmark confidence level is $1.96 * sd^k_{t-1}$. $<$ ($>$) 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			treatments Ap			treatments Bp		
Inflation	$\uparrow$	$\downarrow$	$\sim$	$\uparrow$	$\downarrow$	$\sim$	$\uparrow$	$\downarrow$	$\sim$
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 B.2: Interval correctness depending on the phase of the inflation cycle (% of decisions). $\uparrow$ denotes cases when inflation increases for at least the last 2 periods, and $\downarrow$ denotes cases when it decreases for at least the last 2 periods. $\sim$ represents all other cases. Subjects "underpredict" when the actual inflation is larger than their predicted upper confidence bound; and "overpredict" when the actual inflation is lower than their predicted lower confidence bound.

x_t^k :	all	treat.Ap	treat.Bp
$sip_{t\|t-1}^k$	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)
$D_L \lvert\pi_{t-1}\rvert$	0.1164***	0.0893	0.1184***
	(0.0256)	(0.1862)	(0.0369)
$D_H \lvert\pi_{t-1}\rvert$	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 B.3: Forecasting accuracy and confidence intervals. Note: table is based on equation (2.5). Coefficients are based on fixed effects Poisson estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

x_t^k :	all	treat.Ap	treat.Bp
$sip_{t\|t-1}^k$	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)
α	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)
σ_u	0.6835	0.6804	0.7284
	(0.0728)	(0.1406)	(0.0951)
ρ^*	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 B.4: Forecasting accuracy and confidence intervals. Note: table is based on equation (2.5). Coefficients are based on random effects logit estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

x_t^k :	all	treat.Ap	treat.Bp
$sip_{t\|t-1}^k$	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)
α	-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)
α^*	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 B.5: Forecasting accuracy and confidence intervals. Note: table is based on equation (2.5). Coefficients are based on random effects Poisson estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

$sip^k_{t+1\|t}$:	*all*	*treat.Ap*	*treat.Bp−L*	*treat.Bp−U*
$sip^k_{t\|t-1}$	0.4472***	0.5530***	0.4468***	0.0991
	(0.1058)	(0.0817)	(0.0418)	(0.1038)
sdv^j_{t-1}	0.1119**	0.0993***	0.1472***	0.2929***
	(0.0448)	(0.0373)	(0.0322)	(0.0788)
α	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 B.6: Confidence intervals and standard deviation of point forecasts. Note: table is based on the equation: $sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma sdv^j_{t-1} + u^{em}_t$. 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 the 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)
α	0.3402*
	(0.2976)
N	14904
Wald $\chi^2_{(6)}$	107.5

Table B.7: 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 the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

$sip^k_{t+1|t}$:	*all*	*treat.Ap*	*treat.Bp − L*	*treat.Bp − 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)
α	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 B.8: Confidence intervals and the effect of forecast errors. Note: table is based on the 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 the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

$r^k_{t+1|t}$:	*all*	*treat.Ap*	*treat.Bp − L*	*treat.Bp − U*
$r^k_{t|t-1}$	0.6970***	0.6757***	0.8521***	0.8524***
	(0.1376)	(0.1596)	(0.0250)	(0.0254)
$sip^k_{t+1|t}$	0.0559	0.0812	0.6387***	-0.1139*
	(0.0928)	(0.1102)	(0.1980)	(0.0619)
α	-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 B.9: 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 the 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)
$D_1 y_{t-1}$	0.2836	0.4825**	-0.0122	-0.0028
	(0.4162)	(0.2245)	(0.2913)	(0.2846)
$D_2 y_{t-1}$	-0.3912*	0.1116	-0.0490	-0.0605
	(0.2058)	(0.2339)	(0.1847)	(0.1859)
$D_3 y_{t-1}$	-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)
α	-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)	
σ_u	0.7920	0.5744	0.6879	-
	(0.0996)	(0.0762)	(0.0840)	
ρ^*	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 B.10: 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 the 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)
α	-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)	
α^*	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 B.11: 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 the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

B.2 Instructions used in experiment

Thank you for participating in this experiment, a project in 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.[1]

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 the general rise in prices in the economy. In each period it depends on the inflation expectations of the agents in the economy (you and the other 8 participants in this experiment), the output gap and random shocks which have equal probability of having a positive or negative effect on inflation and are normally distributed.

- The **output gap** measures by how much (in percentage) 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. In each period it depends on the inflation expectations of the agents in the economy, the past output gap, the interest rate and random shocks which have equal probability of having a positive or negative effect on inflation and are normally distributed.

[1] The instructions to participants in the experiment were accompanied with the screenshot of the experimental interface (Figure B.3) and the profit table with earnings for various combinations of *estimation error* and *confidence interval*.

- The **interest rate** is (in this experiment) the price of borrowing money (in percentage) for one period. The interest rate is set by the monetary authority. Their decision mostly depends on the inflation expectations of the agents in economy.

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

- You enter the economy in period 1. In this period you will be given computer-generated past values of inflation, the output gap and the interest rate for 10 periods back (Called: -9, -8, ... -1, 0)
- In period 2 you will be given all the 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 the 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 the 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 inflation in the future period. In each period your prediction will consist of two parts:

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

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

c) *Upper bound* (in percentage) of your prediction. You must be almost sure that the actual inflation will be lower 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 3 *Let's say you think that inflation in the next period will be* 3.7%. *And you also think it 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 a 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 differs 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 no 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 your payoff will be if the actual inflation is indeed in the given interval, but there will also be a greater chance that the actual value falls outside your interval. In our example this interval was 0.8 percentage points. If the actual inflation falls in this interval you 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 value is outside your interval, your receive 0.

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

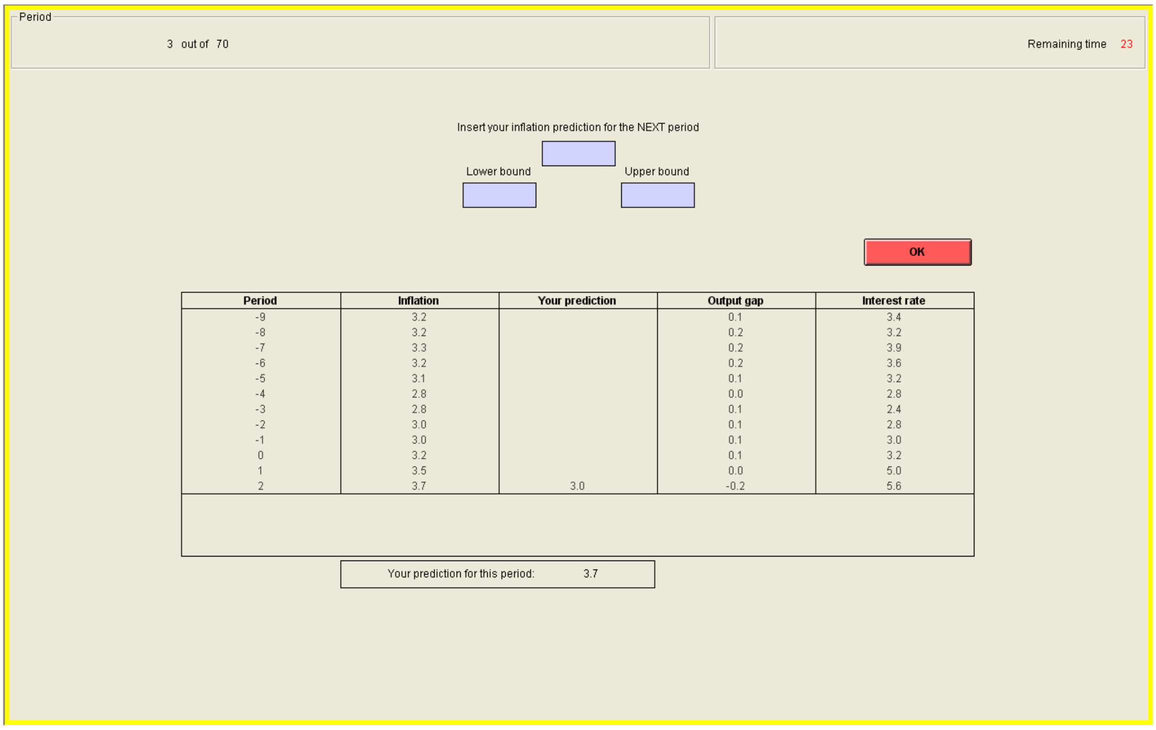

Figure B.3: Screenshot of the experimental interface for treatments Bp. Figure A.7 depicts a screenshot for the treatment with symmetric confidence intervals.

Information after each period

Your payoff depends on your predictions for the next period and the actual realization in the next period. Because the actual inflation will only be known 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 period 1.

- In any other period, you will receive 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 for this period are.

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[2]

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 by more than _0.4_ nor up by 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. You are now in period _ _ _15_ _ _. You have information about past inflation, the output gap and the interest rate up to period _ _ _ _ _ _ _ _ and you have to predict the inflation for period _ _ _ _ _ _ _ _.

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

Appendix C

Additional information for Chapter 3

C.1 Tables

Treatment	No. Units	Efficiency (%)	
		Unit 1	Unit 2
n = 3	180	73.3	99.4
n = 6	80	75.0	96.3
n = undisclosed	132	65.0	91.9

Table C.1: Efficiency of auctions (% of highest value bidders that won corresponding auctions)

Treatment	No. Bidders	No. Units		Observed revenue	RNNE revenue
n = 3	540	360	per unit	5.9844	4.8833
			per bidder	3.9896	3.2556
n = 6	480	160	per unit	11.6381	11.2833
			per bidder	3.8794	3.7611
n = undisclosed	480	264	per unit	7.2850	7.4716
			per bidder	3.7215	3.7056

Table C.2: Observed and expected average revenues of the seller for both units. In the unicertainty case weighted according to realised β_3 and β_6.

Treatment	Subj. ID	No. Obs.	Mean 1st unit	Mean 2nd unit	t-test	p value
3 bidders	1	18	5.644	5.706	0.68	0.253
3 bidders	2*	20	5.260	7.995	2.62	0.008
3 bidders	3*	24	4.038	9.746	3.39	0.001
3 bidders	4	12	7.583	7.583	0.00	0.500
3 bidders	5*	23	6.052	6.835	6.24	0.000
3 bidders	6	23	6.643	7.026	1.57	0.066
3 bidders	7*	20	5.240	5.885	2.02	0.029
3 bidders	8*	27	2.193	11.456	6.36	0.000
3 bidders	9*	17	5.165	7.388	4.30	0.000
3 bidders	10*	19	4.179	5.721	4.35	0.000
3 bidders	11	20	5.960	6.735	0.95	0.176
3 bidders	12*	17	5.253	5.435	1.78	0.047
3 bidders	13	20	6.160	6.160	0.00	0.500
3 bidders	14*	20	3.755	9.655	3.99	0.000
3 bidders	15	17	6.088	6.224	1.53	0.072
3 bidders	16*	21	2.595	9.933	3.93	0.000
3 bidders	17	16	6.088	6.094	1.00	0.167
3 bidders	18*	26	6.035	8.296	2.12	0.022
6 bidders	19*	36	3.367	10.831	6.35	0.000
6 bidders	20	29	6.783	6.972	1.60	0.061
6 bidders	21	31	7.932	7.994	0.61	0.272
6 bidders	22	36	5.853	5.889	0.19	0.425
6 bidders	23	31	7.745	8.906	1.47	0.076
6 bidders	24	36	7.514	7.617	0.60	0.278
6 bidders	25*	31	7.887	8.735	2.01	0.027
6 bidders	26*	33	7.482	8.552	1.95	0.030
6 bidders	27*	33	8.245	8.336	2.32	0.014
6 bidders	28*	34	6.950	7.815	2.03	0.025
6 bidders	29*	34	6.721	8.341	2.25	0.016
6 bidders	30*	36	4.878	7.967	3.16	0.002
Uncertainty	31*	30	7.223	8.803	3.39	0.001
Uncertainty	32	21	7.119	6.795	-0.95	0.824
Uncertainty	33*	31	6.242	8.003	4.01	0.000
Uncertainty	34*	33	4.770	5.709	7.66	0.000
Uncertainty	35*	22	7.041	8.077	2.61	0.008
Uncertainty	36*	37	4.378	8.700	5.70	0.000
Uncertainty	37*	26	7.981	9.223	2.92	0.004
Uncertainty	38*	29	6.114	7.038	2.54	0.008
Uncertainty	39*	31	5.448	7.319	3.39	0.001
Uncertainty	40*	29	5.852	7.217	2.30	0.015
Uncertainty	41*	32	6.531	7.759	5.75	0.000
Uncertainty	42	27	5.052	5.519	1.08	0.145

Table C.3: Comparison of mean subject bids for the first and the second unit by subject. Null hypothesis of the t-test assumes equality of means, p value is one-sided. * denotes subjects where alternative hypothesis *Mean 2nd unit* $>$ *Mean 1st unit* is signifficant at 0.05.

Treatment	Subj. ID	No. Obs.	Mean Subj. 1st unit	Mean RNNE 1st unit	t-test	p value
3 bidders	1*	30	9.203	4.835	9.07	0.000
3 bidders	2*	30	6.800	4.842	3.41	0.002
3 bidders	3	30	6.043	5.313	0.66	0.514
3 bidders	4*	30	11.527	5.763	10.87	0.000
3 bidders	5*	30	7.850	4.498	7.39	0.000
3 bidders	6*	30	8.027	4.760	9.20	0.000
3 bidders	7*	30	7.560	4.032	6.18	0.000
3 bidders	8*	30	3.167	5.785	-2.52	0.017
3 bidders	9*	30	7.960	5.273	5.40	0.000
3 bidders	10*	30	6.090	4.892	4.16	0.000
3 bidders	11*	30	8.447	5.137	4.99	0.000
3 bidders	12*	30	8.003	4.277	7.88	0.000
3 bidders	13*	30	8.620	4.355	7.16	0.000
3 bidders	14	30	7.803	5.895	1.68	0.103
3 bidders	15*	30	9.213	4.732	9.19	0.000
3 bidders	16	30	6.220	5.678	0.45	0.658
3 bidders	17*	30	9.157	4.612	8.17	0.000
3 bidders	18*	30	7.353	4.662	3.33	0.002
6 bidders	19*	40	4.685	9.260	-4.47	0.000
6 bidders	20*	40	9.135	7.888	7.96	0.000
6 bidders	21*	40	9.873	8.480	7.33	0.000
6 bidders	22*	40	6.973	5.770	5.70	0.000
6 bidders	23	40	9.620	8.530	1.90	0.065
6 bidders	24*	40	8.273	7.948	2.19	0.034
6 bidders	25*	40	9.453	8.288	3.87	0.000
6 bidders	26*	40	9.183	8.226	2.17	0.036
6 bidders	27*	40	9.618	7.842	10.47	0.000
6 bidders	28	40	8.170	7.618	1.53	0.135
6 bidders	29	40	8.215	7.702	0.87	0.389
6 bidders	30	40	6.070	7.176	-1.35	0.185
Uncertainty	31*	40	8.183	6.417	5.21	0.000
Uncertainty	32*	40	9.898	7.881	8.84	0.000
Uncertainty	33	40	7.103	6.869	0.53	0.597
Uncertainty	34*	40	5.695	7.260	-5.38	0.000
Uncertainty	35*	40	8.753	6.995	6.80	0.000
Uncertainty	36*	40	4.828	6.451	-3.22	0.003
Uncertainty	37*	40	9.298	8.121	3.01	0.005
Uncertainty	38*	40	7.793	5.750	7.82	0.000
Uncertainty	39*	40	7.020	6.081	2.70	0.010
Uncertainty	40*	40	7.788	6.188	3.12	0.003
Uncertainty	41	40	7.048	6.959	0.22	0.830
Uncertainty	42*	40	7.793	5.409	7.00	0.000

Table C.4: Comparison of subject bids for the first unit to corresponding RNNE prediction by subject. Null hypothesis of t-test assumes equality of means, p value is one-sided. * denotes subjects where alternative hypothesis *Mean Subj. 1st unit* > *Mean RNNE 1st unit* is signifficant at 0.05.

Treatment	Subj. ID	No. Obs.	Mean Subj. 2nd unit	Mean RNNE 2nd unit	t-test	p value
3 bidders	1*	18	5.706	5.972	-3.14	0.003
3 bidders	2	20	7.995	7.995	0.00	1.000
3 bidders	3*	24	9.746	9.758	-1.81	0.041
3 bidders	4	12	7.583	7.583	0.00	1.000
3 bidders	5*	23	6.835	7.017	-3.34	0.001
3 bidders	6*	23	7.026	8.009	-3.37	0.001
3 bidders	7	20	5.885	5.995	-1.15	0.132
3 bidders	8	27	11.456	11.530	-1.00	0.163
3 bidders	9*	17	7.388	7.835	-2.06	0.028
3 bidders	10*	19	5.721	7.132	-2.89	0.005
3 bidders	11*	20	6.735	8.320	-2.89	0.005
3 bidders	12*	17	5.435	5.694	-2.63	0.009
3 bidders	13*	20	6.160	6.295	-2.30	0.016
3 bidders	14	20	9.655	9.655	0.00	1.000
3 bidders	15	17	6.224	6.312	-1.65	0.059
3 bidders	16	21	9.933	9.933	0.00	1.000
3 bidders	17	16	6.094	6.094	0.00	1.000
3 bidders	18	26	8.296	8.308	-1.00	0.163
6 bidders	19*	36	10.831	11.003	-2.66	0.006
6 bidders	20*	29	6.972	7.217	-1.78	0.043
6 bidders	21*	31	7.994	8.610	-2.80	0.004
6 bidders	22	36	5.889	6.064	-1.22	0.115
6 bidders	23	31	8.906	8.932	-1.28	0.106
6 bidders	24*	36	7.617	9.117	-7.55	0.000
6 bidders	25	31	8.735	8.735	0.00	1.000
6 bidders	26	33	8.552	8.770	-1.18	0.123
6 bidders	27	33	8.336	8.409	-1.51	0.071
6 bidders	28*	34	7.815	8.171	-2.46	0.010
6 bidders	29*	34	8.341	8.382	-2.60	0.007
6 bidders	30*	36	7.967	8.100	-1.69	0.050
Uncertainty	31	30	8.803	8.803	0.00	1.000
Uncertainty	32*	21	6.795	8.014	-3.23	0.002
Uncertainty	33*	31	8.003	9.094	-4.71	0.000
Uncertainty	34*	33	5.709	9.500	-8.97	0.000
Uncertainty	35	22	8.077	8.086	-1.45	0.081
Uncertainty	36*	37	8.700	9.319	-2.32	0.013
Uncertainty	37*	26	9.223	10.338	-2.74	0.006
Uncertainty	38	29	7.038	7.038	0.00	1.000
Uncertainty	39*	31	7.319	7.445	-2.08	0.023
Uncertainty	40	29	7.217	7.221	-1.00	0.163
Uncertainty	41*	32	7.759	9.816	-4.08	0.000
Uncertainty	42*	27	5.519	5.844	-2.10	0.023

Table C.5: Comparison of subject bids for the second unit to corresponding RNNE prediction by subject. Null hypothesis of t-test assumes equality of means, p value is one-sided. * denotes subjects where alternative hypothesis *Mean Subj. 2nd unit* > *Mean RNNE 2nd unit* is signifficant at 0.05.

	3 bidders	6 bidders	Uncertainty
Increases	7%	16%	7%
Equalities	11%	24%	10%
Decreases	82%	60%	83%

Table C.6: Comparison of subject bids for the first unit to corresponding RNNE prediction by subject. Null hypothesis of t-test assumes equality of means, p value is one-sided. * denotes subjects where alternative hypothesis *Mean Subj. 1st unit* $>$ *Mean RNNE 1st unit* is signifficant at 0.05.

C.2 Instructions used in experiment

You are about to participate in an experiment. After experiment starts, you are not allowed to communicate with any other participant in the room. Any questions should be addressed to the experimenter.

You will participate in a fictitious auction as a bidder. You will have to post offers to buy certain artificial commodities. All participants receive exactly the same instructions.

Structure of the experiment
Experiment will consist of 30 periods. Each period will represent a different auction. There will be TWO units of commodities offered in EACH auction. In a given period, the value of the first unit is exactly the same as the value of the second unit offered. However, the values of each pair of commodities will be different in each period.

The two units will be offered sequentially: after the first one is sold, the second will be offered. You are interested to buy only ONE unit of a commodity in each period. If you won the first item of a given period, you will not be allowed to bid for the second.

What is auctioned?
In each of the 30 periods, two artificial commodities will be offered. Each of the two will be worth an amount X for you. You will be given your valuation at the beginning of each period. The number X will be randomly drawn from the uniform distribution between 0 and 20. Each value between these two numbers is equally likely to be chosen. In a given period every participant will receive a different personal value of the item.

How does the bidding procedure look like?
In each period (auction) you will receive your new personal value. Then you will be asked to put a single bid for the offered commodity. You will have 30 seconds to make this decision. You will NOT see the offers of the other bidders, nor will you be allowed to post offers of less than 0 and greater to your personal value. If you do not post an offer in 30 seconds, your bid will be recorded as 0.

First stage
After all participants in the auction have posted their offer for the first good, the highest bidder will be determined. You will be informed if you have won the auction or not. If you won the auction, you will not participate in bidding for the second commodity in given period. If you lost the auction you will proceed to bidding for the second commodity. However, you will not receive the information of the winning amount yet.

Second stage
In the second stage, commodity with the same value will be offered, and you will have

another 30 seconds to make a bid. There will be the same participants in this stage, except the winner of the first stage, so there will be one bidder less. After all bids are posted, the winner will be determined in the same way as in the first stage. You will then be informed of the results of both stages: whether you have won any of the two goods or not, the price paid by the winner in the first auction, the price paid by the winner in the second auction, the number of bidders in the auction, and your profits. After that the new auction period will start.

Treatment A specific instructions:

How many people participate in the auction?
There will be 3 participants in each auction. In every period you will be randomly matched with two other participants. All three will participate in the bidding for the first unit. The person who obtained the first unit will not participate in the bidding for the second unit; therefore only 2 bidders will participate in the second stage. **End of treatment A specific instructions.**

Treatment B specific instructions:

How many people participate in the auction?
There will be 6 participants in each auction. In every period you will be randomly matched with two other participants. All three will participate in the bidding for the first unit. The person who obtained the first unit will not participate in the bidding for the second unit; therefore only 5 bidders will participate in the second stage. **End of treatment B specific instructions.**

Treatment C specific instructions:

How many people participate in the auction?
In each auction there will be an unknown number of bidders. However, you know that there are only two possible states with equal (50%) probability to occur. Either the number of participants is 3 or 6. Therefore in the bidding for the first unit, there will be either 3 or 6 participants. The person who obtained the first unit will not participate in the bidding for the second unit; therefore only 2 or 5 bidders will participate in the second stage.

You will not receive the information of the correct number of bidders until both of the two stages are finished. The group members and their number will be different in each period. **End of treatment C specific instructions.**

Who is the winner and what is his payoff?

The winner will be the bidder with the highest offer made. If there will be more than one bidder with the highest offer, the single winner will be randomly chosen among them.

However, the winner will only pay the price equal to the SECOND HIGHEST offer. For instance, if the bidders a, b, and c offered 10, 15 and 8, the highest bidder will be b, who offered 15, but will pay only 10 for the commodity.

The payoff of the winner will be the value of the commodity minus the price he paid, ($Profit = X - p$).

Your total payoff will be calculated as a sum of your gains in all 30 periods with the conversion rate of €1 = 20 monetary units in the experiment.

If you have any questions, you should ask them now.

Figure C.1: Screenshot of the experimental interface for treatment C, auction with bidders uncertain between 3 or 6 participants.

C.3 Individual bidding behavior

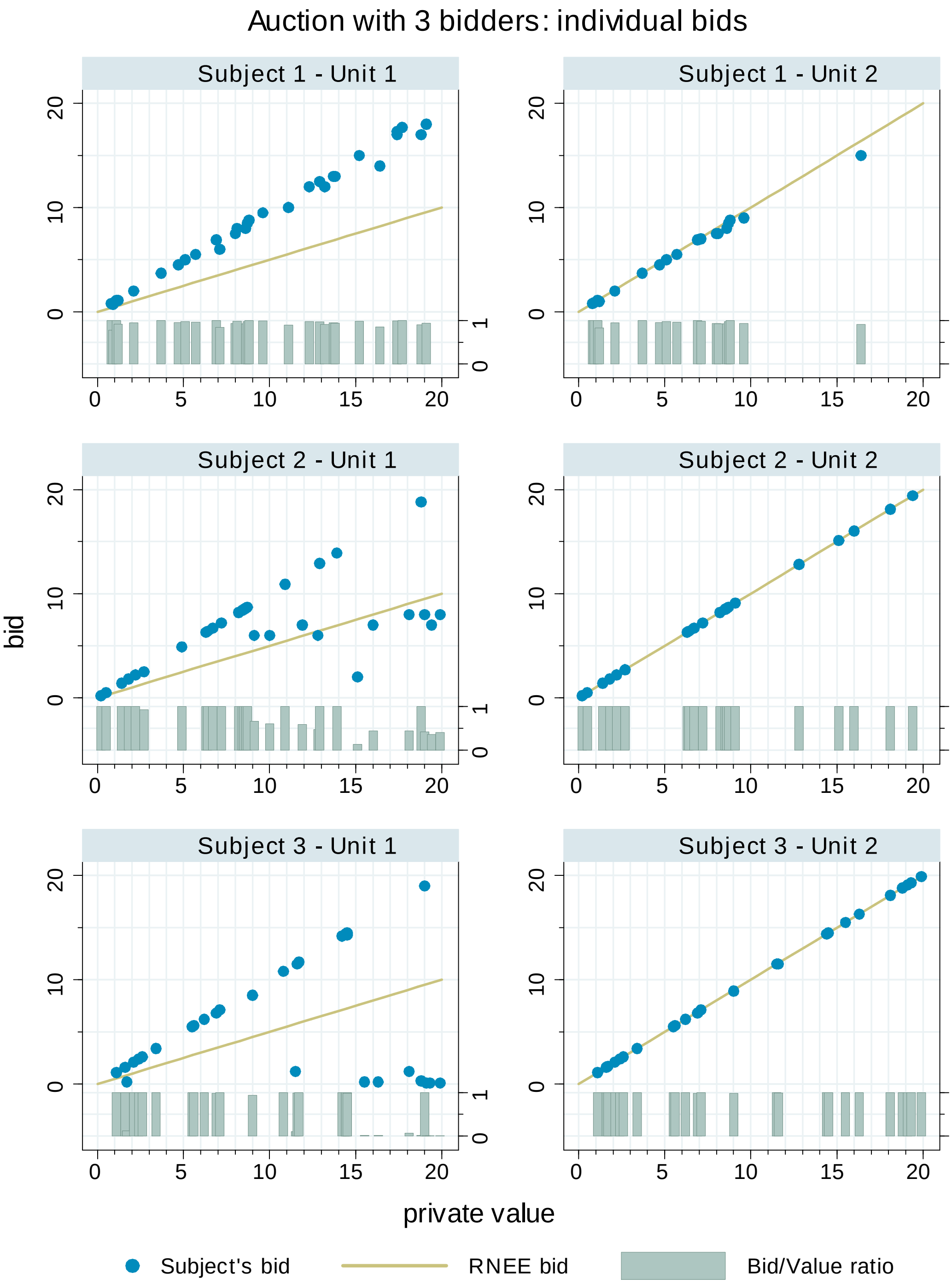

Figure C.2: Auctions with 3 bidders - bids vs. private values and corresponding bid/value ratios by subject. Selected examples.

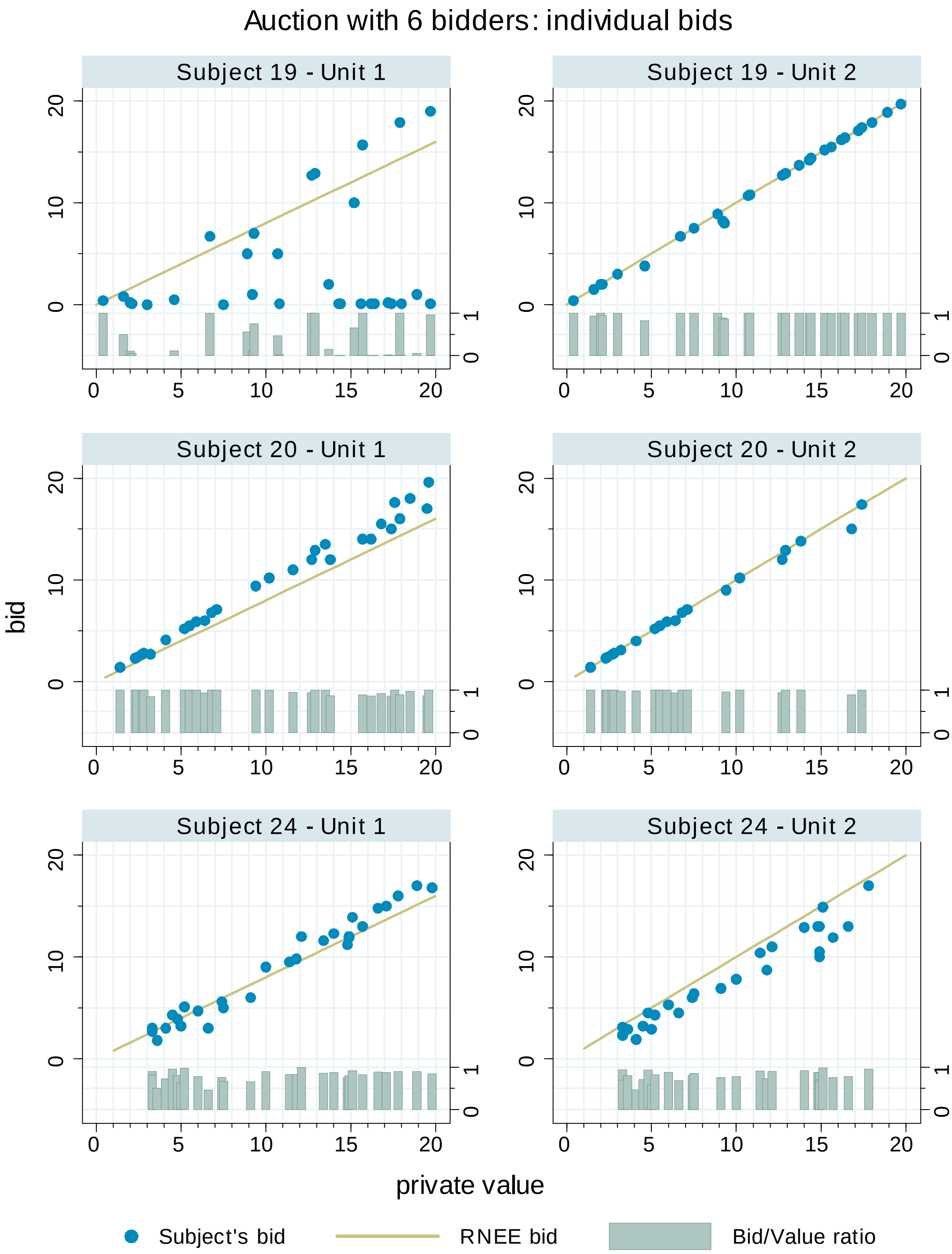

Figure C.3: Auctions with 6 bidders - bids vs. private values and corresponding bid/value ratios by subject. Selected examples.

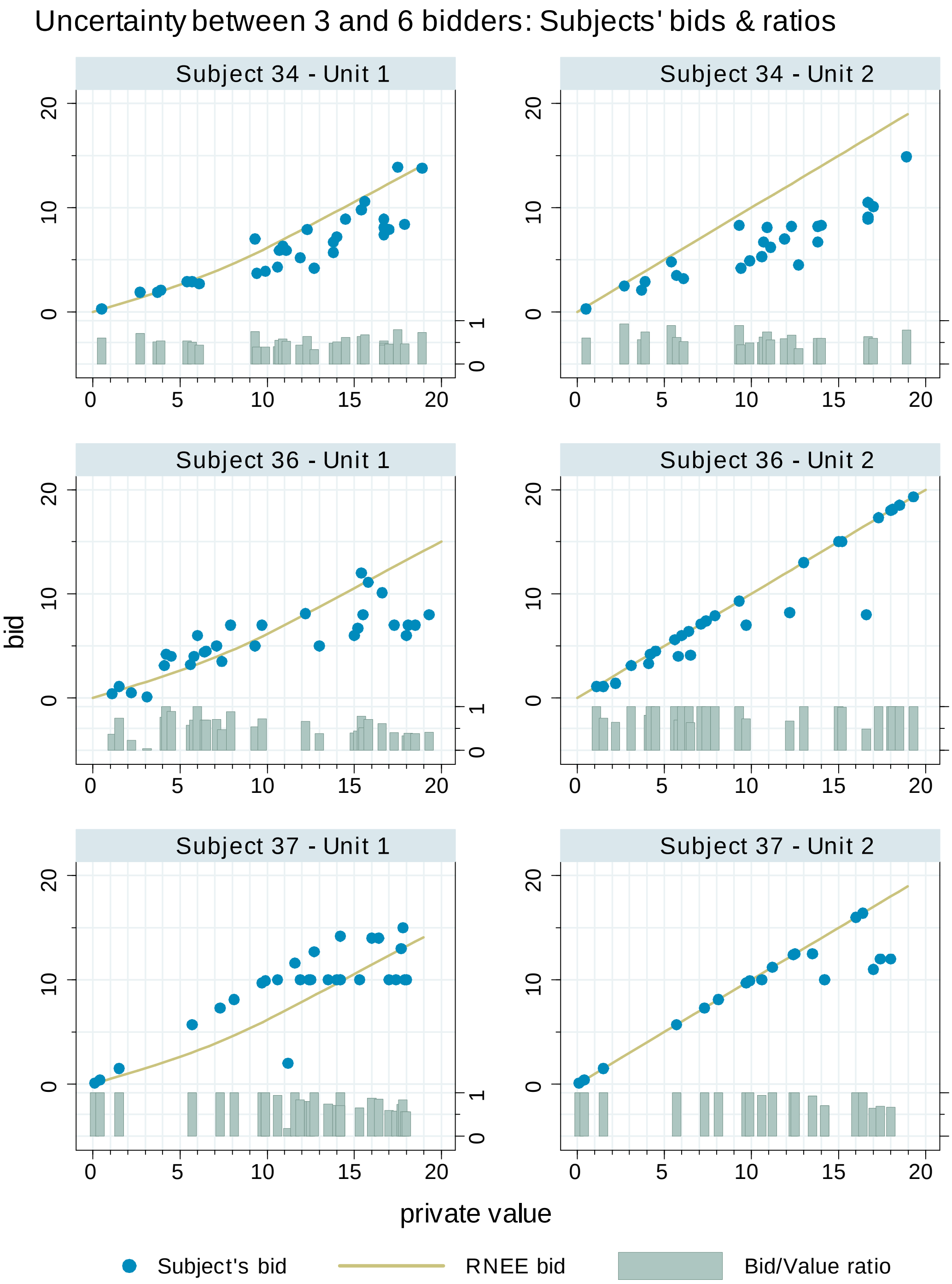

Figure C.4: Auctions with bidders uncertain between 3 or 6 participants - bids vs. private values and corresponding bid/value ratios by subject. Selected examples.