Macroeconomic Literacy and Expectations*

Michael Mirdamadi — University of Western Ontario
Luba Petersen — Simon Fraser University

September 2018

Abstract

We explore the effects of macroeconomic literacy on expectation formation in an experimental economy where participants' aggregated expectations endogenously influence macroeconomic variables. We systematically vary the information participants receive about the economy's data-generating process (no information, qualitative information, and quantitative information) and the central bank's targets. Our experimental results suggest there are many advantages to providing precise quantitative training about the macroeconomy. Compared to an environment where forecasters have no initial information about the structure of the economy, quantitative information about the underlying data-generating process consistently reduces inflation forecast errors, reduce disagreements about inflation, and encourages a larger reaction to past forecast errors. Inflation variability is on average lower with quantitative information. Qualitative information, by contrast, is inconsistently effective at influencing forecasting behavior. Providing information about the central bank's targets increases inflation forecast errors and disagreement about inflation as it introduces an additional piece of information into subjects' information sets to coordinate on.

JEL classifications: C9, D84, E52, E43, G12, G14
Keywords: expectations, macroeconomics, monetary policy, literacy, laboratory experiment

*Corresponding author: Luba Petersen (lubap@sfu.ca, Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada.) We gratefully acknowledge the Social Science and Humanities Research Council of Canada for generous financial support and Sherry Luo for valuable research assistance.

1 Introduction

Do households need to understand the complex inner-workings of the macroeconomy for monetary policy to be effective? The long-standing view in policy circles had been no. So long as financial markets participants understand the effects that monetary and fiscal policy would have on the economy, policy is thought to trickle down to households through changes in market prices and employment opportunities. Since the 2008 Financial Crisis, however, central banks have grappled with understanding the relatively weak effect conventional and unconventional monetary policy has had on public expectations. Poor macroeconomic literacy on the part of households has been one of many explanations for the sluggish response of expectations in response to extraordinarily expansionary monetary policy.

This paper examines how macroeconomic knowledge influences the expectation formation process. This is an important avenue of study as many central banks including the Federal Reserve, Bank of England and the ECB continue to invest time and resources in designing and implementing their own education campaigns. Households' expectations about the economy largely affect their decisions to consume, save, and invest, which in turn affects macroeconomic outcomes. An economically literate populace would likely better understand the forces driving aggregate activity and form more rational expectations in response to policy, implying a better understanding of monetary policy and its effects on the economy would lead to greater efficacy of monetary policy.

It is empirically challenging to disentangle the effects of central bank education on the expectation-formation process and macroeconomic conditions. First, a central bank is more likely to communicate its targets and its understanding of the economy when they are more certain that the public has a similar, common understanding. Central bank efforts over the past three decades to improve the public's macroeconomic literacy have been influenced by their past communication successes and failures.

To tackle these questions, we design a laboratory experiment to study the effects that macroeconomic literacy has on individual and aggregate expectations and aggregate dynamics. As our focus is entirely on expectation formation, the experimental task concentrates solely on repeatedly soliciting incentivized forecasts from participants. Exogenous, known shocks to aggregate demand and subject-supplied forecasts directly influence the decisions of automated households and firms, and thus, the experimental economy.

In a between-subject design, we vary the nature of information provided to independent groups of subjects. The different levels of information are meant to resemble the different levels of macroeconomic literacy the public may have, ranging from absolutely no understanding of the relationship between macroeconomic variables to having the access to the central bank's forecasting model. In our baseline NoInfo treatment, participants are provided with no information about the relationship between macroeconomic variables. This environment can be thought of as our completely economically illiterate population who must learn the economy's data-generating process through

experience. In our QualInfo treatment, subjects receive instructions explaining the qualitative relationship between all macroeconomic variables. This treatment mimics the qualitative education that central banks are increasingly providing to the public. In two additional treatments, we supplement the NoInfo and QualInfo treatments with information about the central bank's quantitative inflation and output gap targets. The QualInfo with Targets treatment is most consistent with the information provided to the public by inflation-targeting central banks like the Bank of Canada and the ECB. Finally, in our QuantInfo treatment, subjects' receive instructions explaining in full detail the quantitative data-generating process, as one may read in a central bank's forecasting model.

We observe in all treatments a range of forecasting heuristics despite participants receiving common information about the economy's DGP. Heterogeneity in inflation forecasts is noticeably lower when participants receive precise quantitative information. Information about the central bank's targets and qualitative information do not, however, meaningfully impact the coordination of expectations.

Macroeconomic training significantly influences a number of dimensions of expectation formation. Forecast accuracy improves with different types of macroeconomic training. Inflation expectations are significantly more accurate when subjects receive precise quantitative information about the underlying DGP. Inflation forecast errors are nearly 30% lower in the QuantInfo treatment than in the NoInfo treatment while output forecasts errors are 10% lower in the QualInfo treatment after subjects gain experience.

We also find that disagreement about inflation is significantly lower in the QuantInfo treatment, and output forecasts are better coordinated when subjects face large shocks. Qualitative information, in contrast, increases disagreement among forecasters.

Communicating the central bank's quantitative inflation and output gap targets has mixed effects. Compared to the NoInfo treatment, providing targets to participants in the NoInfoTarget treatment increases inflation forecast errors and disagreement about both output and inflation. We only observe a reduction in output gap forecast errors. The increase in disagreement and inflation forecast errors in the NoInfoTarget treatment is likely due to the introduction of a new focal point, the targets, in addition to the aggregate demand shocks. The upside of providing the targets is that there is a noticeable increase in ex-ante rational types and a reduction in the frequency of backward-looking heuristics as the targets appear to encourage forecasting toward the steady-state. There is also a sizeable increase in the frequency of expectations anchored on the central bank's targets. Together, this shift in the distribution of forecasts encourages more stable aggregate expectations and aggregate dynamics. Likewise, communicating the central bank's targets when participants already have a qualitative understanding of the economy's DGP is less effective at managing expectations. Compared to the QualInfo treatment, providing targets in the QualInfoTarget treatment increases errors and disagreement about inflation. We observed increased

heterogeneity in forecasting heuristics as subjects have additional information to coordinate their forecasts on.

Overall, our experimental results indicate that providing precise quantitative information about the economy's data-generating process is more effective at managing expectations than providing qualitative information or quantitative targets. We speculate that the improvements in forecasting inflation in the QuantInfo treatment come from precise information of the slope of the Phillips Curve. Subjects in this treatment have a better sense of the relative importance of aggregate shocks and expectations in driving inflation.

The paper is organized as follows. Section 2 discusses related empirical and experimental literature on expectation formation and financial literacy. Section 3 lays out our experimental design and hypotheses. Section 4 presents our experimental results and Section 5 concludes with a discussion and policy recommendations.

2 Literature

Full-information rational expectations (FIRE) has become the default approach to modelling agents' expectations in macroeconomics. This assumption is very important as it significantly strengthens the predicted potency of monetary policy. Specifically, it assumes that agents fully incorporate all future monetary policy decisions and their effect on the economy into their forecasts. In turn, this expectations channel of monetary policy significantly stabilizes concurrent economic activity in response to demand shocks.

Monetary policy is considerably less effective when agents are assumed to form certain types of non-rational expectations. Kryvtsov and Petersen (2015) show that the strength of the expectations channel is significantly weakened when agents incorporate irrelevant historical information into their forecasts. In this case, the central bank's policy rule must respond more aggressively to inflation and output to achieve the same level of macroeconomic stability as would be obtained under rational expectations.

The assumption of rationality has been strongly rejected by both empirical and experimental data. Coibion et al. (2017) provide an extensive survey of empirical work studying surveyed expectations. FIRE has been rejected extensively within the Livingston Survey and Michigan Survey of Consumers and the Survey of Professional Forecasters (Roberts 1998, Mankiw et al. 2003, Croushore 1997, Coibion and Gorodnichenko 2015, Malmandier and Nagel 2016). Likewise, in laboratory settings, participants do not exhibit rational expectations despite having full knowledge of the economy's data-generating process (Kryvtsov and Petersen 2015, Mokhtarzadeh and Petersen 2017).

Recursive least squares learning is an alternative approach to modeling expectations. Agents are assumed to update their forecasts based on their past forecast accuracy. This is typically done through an updating of parameters of a perceived law-of-motion. Constant gain least squares

(CGLS) learning models assign greater weight to more recent observations. CGLS modeling has been incorporated into RBC and medium-scale macroeconomic models (Slobodyan and Wouters 2012, Williams 2003, Orphanides and Williams 2005). Empirically, constant-gain learning models have been shown to perform better than Kalman Filter and Recursive Least Squares learning at fitting surveyed expectations (Branch and Evans 2006). Eusepi and Preston (2010) consider the effects of central bank communication in an environment where agents are learning the parameters of an economy's data generating process (DGP). With complete quantitative information about the central bank's policy rule or simply a knowledge of the variables upon which the nominal interest rate decisions are conditioned, the rational expectations equilibrium (REE) is stable for all parameter values under learning. In short, it is sufficient to communicate the variables in the central bank's reaction function so long as the policy maker is satisfying the Taylor principle by raising rates more than one-for-one with inflation. Communicating only the inflation target can generate expectations-driven instability, especially in environments with persistent shocks.

Our paper contributes to a large literature studying the individual benefits of economic literacy. Numerous surveys from the OECD, PISA, StatsCan, and IMF have shown a positive correlation between economic education and literacy and income. Jappelli (2010) finds that economic literacy is positively correlated with GDP per capita. Economic literacy is higher among recipients of better math test scores, from countries with higher GDP growth and higher social security contributions. Bruine de Bruin et al. (2010) observe a strong negative correlation between inflation expectations and wealth. Older, single, poorer, less-educated Americans expect higher inflation. They also observe a significant positive relationship between being financially constrained and expecting inflation to be above 5%.

There has been a significant effort by central banks over the last decade to improve both economic and financial literacy. Over 30 central banks provide educational material in partnership with policy makers and economic institutions (Fluch, 2007). Primary target groups in the education system are schoolchildren, post-secondary students, and secondary school teachers. Central bank outreach has primarily focused on topics such as interest rates and compound interest, inflation, debt, mortgage finance, and insurance. These institutions also share their knowledge with the public through publication of their own research and projection models. The effects of these education initiatives on macroeconomic expectations is not yet clear. Our paper provides causal evidence on the efficacy of macroeconomic training.

Drager et al. (2016) compare professional forecasts from the Survey of Professional Forecasters to consumer forecasts from the University of Michigan Survey of Consumers in terms of their ability to forecast according to the Fisher equation, Phillips Curve, and Taylor principle. They find that expectations are most theory-consistent with regards to the Fisher equation (84% of professionals, 50% of consumers). Roughly 50% of professionals and consumers form forecasts consistent with the Taylor principle. Consumer expectations appear to be the least consistent with the Phillips

curve, with only 34% of households consistently expecting unemployment and inflation to move in opposite directions. By contrast, 51% of professional forecasters form expectations consistent with the Phillips curve. The vast majority of macroeconomic models used by central banks assume that agents understand these relationships. With only 6% of consumers and 31% of professional forecasters able to form theory-consistent expectations, there is significant room and critical need to improve the public's macroeconomic literacy.

In a related study, Burke and Manz (2014) surveyed experimental subjects' financial literacy and in an incentivized experiment asked them to predict inflation. Economic literacy was positively correlated with forecast accuracy, and economic literacy accounted for most of the variation across demographics. Our experimental design differs from Burke and Manz (2014) in several ways. Our interest is in the causal effect of teaching an economy's law of motion, in particular the structural equations and parameters which describe output, inflation, and the interest rate. While Burke and Manz use financial literacy surveys as a measure of subjects' knowledge, we control for macroeconomic literacy with the use of different types of information. We explore the effect of providing more precise information about an economy's structure. Moreover, our subjects' expectations have an endogenous impact on the experimental economy. This allows us to additionally evaluate the aggregate and endogenous impact of macroeconomic training.

This paper also makes important contributions to the design of learning-to-forecast (LTF) experiments, pioneered by Marimon and Sunder (1993). In these experiments, groups of subjects submit forecasts about variables that are endogenously determined by the submitted aggregate forecasts. As noted by Hommes and Lux (2013), a key advantage of LTF experiments is the ability to obtain clean data on expectations while controlling for all other market conditions. By the very wording of its name, LTF experiments focus on understanding how people learn to form beliefs in their environment. The majority of experiments in this literature provide subjects with a qualitative description of the underlying DGP. This is for at least a couple of reasons. First, much of the literature is interested is studying how subjects come to learn the structure of the underlying market or economy. Second, it is often thought that a multi-equation DGP is too complex for subjects to learn. The advantage of using qualitative instructions is that they are less technical and easier to convey to participants. However, they come at the cost of a lack of full information about the DGP. This makes it impossible to test the predictions of models where agents are assumed to have a complete understanding of their economy's DGP. A small number of experiments do provide subjects with a fully quantitative DGP. Table 1 categorizes the experimental LTF literature by the type of instructions about the DGP provided to subjects.

A large branch of the literature focuses on asset price expectations. In these experiments, participants forecast in a single-equation environment where asset prices endogenously respond to expectations of future asset prices. Sonnemans and Tuinstra (2010) find that quantitative and qualitative information about the DGP lead to similar forecasting heuristics. To the best of our

knowledge, this is the first laboratory experiment to systematically compare forecasting behavior under different information sets in a more complex, multi-equation environment.

3 Experimental Design and Hypotheses

The purpose of our experiment is to identify how people form expectations of economic conditions under different information sets. Specifically, we seek to understand how qualitative and quantitative information about an economy's data-generating process influences forecasting heuristics.

Our experiment economy follows the design of Mokhtarzadeh and Petersen (2017) and is modeled as a linearized system of equations based on a reduced-form version of the New Keynesian framework. The New Keynesian framework is commonly employed by central banks in their own forecasting models. We choose a simple version as a starting point, but leave a more complex and nonlinear data-generating process for future research.

The economy evolved according to a system of four equations:

$$x_t = \mathbb{E}_t x_{t+1} - \sigma^{-1}(i_t - \mathbb{E}_t \pi_{t+1} - r_t^n) \tag{1}$$

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t \tag{2}$$

$$i_t = \phi_\pi \pi_t + \phi_x x_t \tag{3}$$

$$r_t^n = \rho_r r_{t-1}^n + \epsilon_{rt} \tag{4}$$

Equation (1) is the Investment-Saving (IS) curve and describes the level of the output gap x_t, or aggregate demand relative to the steady state, which changes in response to the aggregate expected future output gap $\mathbb{E}_t x_{t+1}$, and the difference between the real interest rate, $i_t - \mathbb{E}_t \pi_{t+1}$ and the natural rate of interest, r_t^n. The difference between the real interest rate and natural interest rate is sensitive to households' intertemporal elasticity of substitution, σ^{-1}.

Equation (2) is the New Keynesian Phillips Curve which describes the level of inflation, π_t in response to changes in aggregate expected future inflation, $\mathbb{E}_t \pi_{t+1}^*$ and current aggregate demand, x_t. In this equation, β is a parameter that represents the economy's discount rate and κ is a function of parameters that represents the sensitivity of firms' pricing behaviour in response to aggregate demand.

Equation (3) is the Taylor rule equation which represents the central bank's policy reaction function. The nominal interest rate changes based on deviations of output gap and inflation from their respective steady states. The parameters ϕ_π and ϕ_x dictate the central bank's reaction to inflation and output gap, with the goal of keeping aggregate demand and inflation stable at their steady-state levels of 0.

Equation (4) describes the natural rate of interest which is subject to random perturbations.

For the purposes of this study, r_t^n is understood as a *shock* to the economy to aggregate demand while ϵ_{rt} represents an *i.i.d. N(0,*σ_ϵ*)* innovation.

The parameters of the data-generating process were chosen in accordance of moments of Canadian data following Kryvtsov and Petersen (2013); $\sigma = 1$, $\beta = 0.989$, $\kappa = 0.13$, $\phi_\pi = 1.5$, $\phi_x = 0.5$, $\rho_r = 0.57$, $\sigma_\epsilon = 1.13$, and steady-state values $\pi^* = x^* = i^* = 0$.

Each session of the experiment began with participants receiving a detailed common set of instructions about the task that they would be participating in, followed by four unpaid practice rounds with the computer interface. Thereafter, participants repeatedly submitted forecasts over two sequential 30-period horizons. During the first 10 periods of the first sequence, participants had 65 seconds to submit their forecast, followed by 50 seconds thereafter. After all subjects had submitted their forecasts for the subsequent period, or time ran out, the median submitted forecasts for output and inflation for each group were selected as the aggregate expectations and directly used in the calculation of output, inflation, and the nominal interest rate. We selected the median forecast rather than the average to avoid significantly biased aggregate expectations driven by a small number of subjects.

Participants were incentivized to submit accurate forecasts. Subject i's score in each period t was calculated as:

$$Score_{i,t} = 0.3(2^{-0.01|E^*_{i,t-1}\pi_t - \pi_t|} + 2^{-0.01|E^*_{i,t-1}x_{i,t} - x_t|}) \tag{5}$$

where $E^*_{i,t-1}\pi_t - \pi_t$ and $E^*_{i,t-1}x_{i,t} - x_t$ are subject i's inflation and output gap forecast errors submitted in period *t-1* forecasting period t, respectively. This scoring continuously decreased subjects' payoffs as forecast errors grew larger. A subject would earn a maximum score per period of 0.6 for a forecast error of 0. For every 100 basis points error made for each variable, a subject's score would decrease by 50%. At the end of the two repetitions, the subjects' total scores were tallied, converted into dollars and paid out in cash in addition to a show-up fee of $7. Finally, all subjects received common information about historical data to be presented on their screen and details about how their payoffs would be computed.

3.1 Treatments

Our experiment involves five treatments that differ based on the information provided to the participants about the economy's data-generating process. The treatments and their differences are summarized in Table 2. In all five treatments participants were provided basic definitions about the key variables they would observe during the experiment (the output gap, inflation, nominal interest rate, and demand shocks). They were also informed that all values would be expressed and inputted in basis points.

Participants in the *NoInfo* treatment received no additional information about the data-generating process. They were not informed about the relationship between different macroeconomic variables or that their aggregated expectations would have an endogenous effect on the economy. Moreover,

participants were not informed about the reaction function or objectives of the central bank.

The *NoInfo-Targets* treatment provided the same basic information as the NoInfo treatment as well as precise information about the central bank's inflation and output targets of zero.

The *QualInfo* treatment involved presenting participants with a qualitative description of the data-generating process. The instructions described qualitatively the reaction of each variable to the other macroeconomic variables. The qualitative instructions also conveyed that aggregate expectations would have an effect on the economy. The central bank's objectives and reaction function were explained qualitatively.

The *QualInfo-Targets* treatment extended the QualInfo treatment by additionally providing subjects with the central bank's precise quantitative inflation and output targets.

Finally, the *QuantInfo* treatment provided participants with a detailed quantitative description of the economy's data-generating process. Participants learned the economy's system of equations, including the values of all parameters and targets, i.e. β, σ, κ, ρ, σ_ϵ, ϕ_π, ϕ_x, π^* and x^*.

3.2 Testable Hypotheses

In this subsection we lay out testable hypotheses associated with our five information treatments. A standard assumption in macroeconomic theory is that the representative agent forms rational expectations over their stochastic environment. Under this assumption, agents' forecast errors would be on average zero and not systematically biased. This standard modelling assumption is the basis of our first hypothesis.

Hypothesis 1: Expectations are rational, ie. are not systematically influenced by current or lagged innovations.

Extensive experimental and survey evidence has demonstrated that individuals significantly over-rely on historical information to formulate their forecasts. Without detailed information about the economy's data-generating process, participants in the NoInfo and QualInfo treatments would likely have difficulty forming accurate forecasts. Thus, we form an alternative set of hypotheses that increasing the precision of information about the data-generating process reduces participants' forecast errors and disagreement.

Alternative Hypothesis 1: Absolute forecast errors are lowest in the QuantInfo treatment, medium in the QualInfo treatments, and largest in the NoInfo treatments.

Alternative Hypothesis 2: Quantitative information about the central bank's targets reduces absolute forecast errors.

Alternative Hypothesis 3: Disagreement among forecasters - measured as the standard deviation of forecasts in a given round - is lowest in the QuantInfo treatment, medium in the QualInfo treatments, and largest in the NoInfo treatments.

Alternative Hypothesis 4: Quantitative information about the central bank's targets reduces disagreement.

3.3 Procedures

The experiments were conducted at the CRABE laboratory at Simon Fraser University from September 2016 to July 2018. A total of 208 undergraduate students were recruited on a first-come first-serve basis from a subject pool of over 3000 subjects from various disciplines. No subject had experience in a learning-to-forecast experiment. Each session, consisting of 7 participants[1], lasted approximately 90 minutes including 30 minutes for instructions and practice. We conducted 6 sessions per treatment, with two repetitions each consisting of 30 periods. We collected a total of 12,480 observations on individual forecasts, of which after eliminating a single outlier subject we had 7,440 observations.

Within a treatment, sessions differed based on the randomly generated shock sequence. Shock sequences were pre-selected for the sake of consistency across treatments. The shocks were drawn from a normal distribution where two-thirds of the time, the shock took on a value between -138 and +138 basis points, and 95% of the time took on a value between -276 and +276 basis points. The standard deviations of the individual shock sequences ranged from 125 to 155 basis points. This was an additional dimension of exogenous variation that we are able to exploit in our analysis.

We programmed the experiment in *Redwood* (Pettit et al. 2014). Figure 1 presents a sample screenshot from our NoInfo and QualInfo treatments. Participants submitted numerical predictions for the subsequent period's inflation and output. They were told to submit their forecasts as integers in basis points. The user interface contained graphical time series of historical interest rates, shocks (including the current period shock), output, inflation and personal expectations over time. The screen also included a timer indicating the amount of time remaining in a period and their accumulating score. Subjects had access to their paper instructions and a computer calculator during the entire duration of the experiment.

Participants were instructed not to communicate with each other during the experiment. They were allowed to ask clarifying questions, but were not allowed to ask questions regarding strategy. The average payment, including a CDN$7 show-up fee was CDN$19.75 and ranged from $17 to $26.

[1] NoInfo Session 6 and NoInfo-Target Session 3 only contained 6 participants

4 Experimental Results

4.1 Individual Results

Forecast Errors and Disagreement

We begin by evaluating how subjects' forecast errors are systematically influenced by the availability of information about the economy's data-generating process. Figure 2 presents the empirical cumulative distribution functions for absolute output and inflation forecast errors, by repetition, for all subjects and periods in the five information treatments. There are no stark differences across treatments for output gap forecasts, with median absolute forecasts errors around 75 basis points across the treatments . There is more distinct differences across treatments in terms of inflation forecasts. Inflation forecasts are the lowest across the entire distribution in the QuantInfo treatment. NoInfo and QualInfo inflation forecast errors are slightly higher at the median and much higher for more extreme forecast errors. Interestingly, absolute forecast errors increase further when participants have precise quantitative information about the central bank's targets in the NoInfo-Targets and QualInfo-Targets treatments.

Figure 4 presents the distribution of forecast disagreement, by treatment and repetition, for all periods. While there does not appear to be any significant differences in output disagreement, we do see

We construct a series of mixed-effects regressions that estimate the effects of the information treatments on forecast accuracy and forecast disagreement. Our initial specifications estimates the effect of qualitative and quantitative information about the DGP and targets on subjects' ex-ante one-period ahead absolute forecast errors and the standard deviation in forecasts. Specifically, we regress

$$\begin{aligned} log|E_{i,t}z_{t+1} - z_{t+1}| &= \alpha + \beta_1 NoInfoTarget + \beta_2 QualInfo + \beta_3 QualInfoTarget \\ &\quad + \beta_4 QuantInfo + \eta_{i,t}, \\ log(s.d.E_t z_{t+1}) &= \alpha + \beta_1 NoInfoTarget + \beta_2 QualInfo + \beta_3 QualInfoTarget \\ &\quad + \beta_4 QuantInfo + \eta_{i,t} \end{aligned}$$

where z refers either to output gap or inflation. NoInfoTarget, QualInfo, QualInfoTarget and QuantInfo are dummy variables that takes the value of 1 for observations from the respective treatments and zero otherwise. Thus, $\hat{\alpha}$ is the estimates of the average log absolute forecast error or standard deviation in in forecasts observed in the NoInfo treatment while $\hat{\alpha}+\hat{\beta}_i$ for $i \in \{1, 2, 3, 4\}$ are the estimates observed in the QualInfo, QualInfoTarget and QuantInfo treatments respectively. Forecast error results are presented in the first four columns of Table 3 while standard deviation results are presented in the remaining columns.

First, $\hat{\alpha}$ is positive and significantly different from zero in all specifications, indicating that NoInfo subjects make significant forecast errors and exhibit significant disagreement. Neither the errors nor the disagreement are declining with experience.

Compared to the NoInfo treatment, output gap forecast errors are not significantly different when subjects receive additional information when participants are inexperienced. With experience, average absolute forecast errors are lower in all the information treatments except the QuantInfo treatment. Mean Inflation forecast errors are made larger in the QualInfo-Target treatment and smaller in the QuantInfo treatment, irrespective of experience. We also observe experienced subjects in the NoInfo-Target treatment making larger inflation forecast errors.

We next break down our analysis by pairwise treatment comparisons. Table 4 presents estimated effects of different treatment variations on output gap and inflation forecast errors. Beginning with the NoInfo treatments in Panel A, we observe that providing details about the central bank's targets when subjects know nothing about the economy's DGP results in significantly smaller output gap forecast errors and larger inflation forecast errors when subjects are experienced. The additional target information also leads to increased disagreement about output and inflatexperienced subjects to disagree more about output and inflation.

Comparing the NoInfo-Target and QualInfo-Target in Panel B allows us to evaluate the effect of qualitative information when subjects already know the central bank's targets. Subjects who receive supplementary qualitative information about the DGP disagree significantly less about future output (when inexperienced) but significantly more about future inflation. Inflation forecast errors are significantly larger greater when presented with qualitative information about the DGP.

Panel C considers the effect of providing targets to subjects with a qualitative understanding of the economy. Providing details about the targets results in significantly larger inflation forecast errors and inflation disagreement. The key advantage of the additional target information is that it reduces output gap disagreement.

Panel D compares the QualInfo-Target and QuantInfo treatments. Supplementing the QualInfo-Target subjects with further details of the parameterization of the DGP in the QuantInfo treatment significantly reduces inflation disagreement and inflation forecast errors, but also generates increased disagreement about the output gap.

Overall, it appears that information about the central bank's targets appears to have limited benefits. Target information when subjects have no understanding of the economy improves output gap accuracy when subjects are experienced, but overall leads to more disagreement. With more experience, target information only serves to reduce output gap disagreement. Additional target information and qualitative information about the DGP does not appear to improve inflation forecast accuracy or disagreement, and in some cases, appears to worsen it. By contrast, providing quantitative information about the relationship between macroeconomic variables appears to significantly improve inflation forecast accuracy.

Observation 1: Qualitative information about the DGP reduces (increases) inexperienced subjects' output gap (inflation) forecast errors and disagreement.

Observation 2: Quantitative information about the DGP significantly reduces inflation forecast errors and disagreement and increases output gap disagreement

Observation 3: Supplementary information about targets when subjects have incomplete information about the DGP increases inflation forecast errors and disagreement.

Forecasting Heuristics

We next explore the heterogeneity in forecasting heuristics. We classify participants into a particular type by treatment and experiment. Our approach is to calculate, for each subject, hypothetical expectations according to various classes of forecasting heuristics. We then compute the root-mean squared error (RMSE) of their actual forecasts from the various hypothetical forecasts to identify the best fitting (lowest RMSE) heuristic.

Table 5 presents the five general classes of heuristics that we consider. The heuristics are either widely-regarded assumptions or found to well-describe expectations in other learning-to-forecast experiments. M1 describes an ex-ante rational forecaster who, knowing the DGP, forms the correct forecasts for the subsequent period's output and inflation. An ex-ante rational forecaster will employ the current innovation and past shocks, as well as parameters of the model, to formulate their forecast. M2 describes an ex-post rational forecaster who is simply forms a highly correct forecast. M3 describes a naive forecaster who expects next period's outcomes to simply be what she observed in the previous period. M4 considers the possibility that forecasters use past forecast errors to update their past forecast errors, while M5 assumes subjects extrapolate the past trends in output and inflation to formulate their forecasts. For models M4 and M5, we consider a variety of parameterizations. The distribution of forecasting heuristics, by treatment and repetition, are presented in Figure 5.

The first key takeaway is that there is considerable heterogeneity in subjects' forecasting heuristics, even when they have complete information about the economy's DGP in the QuantInfo treatment. We observe all five heuristics being represented in nearly all treatments and levels of experience. We follow Allaj (2018) in measuring heterogeneity observed in the treatment as one minus the square root of the sum of the squares of the relative frequencies of the different categories. Table 6 presents the heterogeneity scores of each treatment from lowest to highest. Across the board, QuantInfo leads to the least heterogeneity in forecasting heuristics. The reduction in heterogeneity of heuristics is more pronounced when subjects are inexperienced. With experience, QuantInfo are only marginally more homogeneous in how they forecast than the next categories. While QuantInfo

subjects become more heterogeneous with experience, subjects in other treatments appear to learn to coordinate on a fewer set of heuristics.

The vast majority of subjects do not form ex-ante rational (M1) forecasts. Across treatments 0-5% of subjects forecasting output and 7-22% of subjects forecasting inflation are best described by the M1 model. The QuantInfo treatment has the highest proportion of subjects who forecast the output gap according to the M1 model, but the difference from other treatments is very modest.

Intriguingly, inexperienced NoInfo subjects are the most likely to formulate ex-ante rational forecasts. This may be due to the fact that, with limited information about the economy's data-generating process, participants have only the current shock and historical information to focus on. However, with experience this proportion decreases by roughly one-half.

In contrast, we observe the highest proportion of naive forecasting heuristics in the QuantInfo treatment. For inexperienced subjects, 71% of subjects forecasting output and 55% of subjects forecasting inflation are using the M3 model. These proportions fall to 60% and 48% with experience. By contrast, we observe relatively low proportions of NoInfoTarget and QualInfoTarget subjects rely on historical information to forecast (the proportions range from 20-40%). We suspect this may be because the central bank's target is made relatively more focal and encourages a reversion of expectations to the target.

Likewise, the trend-chasing heuristic (M4) is the most prevalent in the QuantInfo treatment. Among inexperienced QuantInfo subjects, 5% using M4 to forecast the output gap and 12% use it to forecast inflation. Both these proportions increase with experience to 12% and 19%, respectively. By contrast, with experience, the QualInfoTarget information is the most reliable at reducing the prevalence of trend-chasing expectations.

Finally, we turn our attention to a class of models of constant gain learning (M5) to describe our subjects' behavior. Versions of this type of adaptive learning model are widely used in the learning-in-macroeconomics literature and have found strong support in empirical survey and experimental data (Milani 2011, Malmandier and Nagel 2015, Pfajfar and Zakelj 2016). We observe a small proportion of subjects employing M5 to forecast output and inflation. Once experienced, NoInfo and NoInfoTarget participants are the most willing to update their forecasts based on their past forecast errors (roughly 25-45% of subjects in these treatments are best described by the M5 model). QuantInfo subjects are the least likely to employ the heuristic, with roughly 10-15% employing it for their different forecasts. In fact, among experienced subjects, the ordering is nearly perfect: subjects with less information are more willing to update subsequent forecasts based on their past forecast errors.

Observation 4: Less information about the DGP encourages more subjects to adopt a constant-gain learning heuristic and update their forecasts based on their past errors.

Our next specification considers how information about the DGP alters subjects' reliance on current vs historical information. We estimate ex-ante forecast errors as functions of the history of innovations to the natural rate of interest, ϵ_{t-s}, where we interact these innovations with treatment dummies:

$$\begin{aligned} E_{i,t}z_{t+1} &= \alpha_1 + \alpha_2 TreatmentDummy + \beta 1\epsilon_{r,t} + \beta 2\epsilon_{r,t} \times TreatmentDummy... \\ &\quad + \beta M\epsilon_{r,t-T} + \beta N\epsilon_{r,t-T} \times TreatmentDummy + \eta_{i,t} \end{aligned}$$

where $T = 6$. This very general specification provides insight into how subjects over- and under-react to current and historical events that are purely exogenous to theirs and others' forecasting heuristics. Under rational expectations, ex-ante forecast errors should be uncorrelated with shock innovations at any lag. In other words, all regressors should have an estimated coefficient that is not statistically different from zero. Under a backward-looking heuristic, ex-ante forecast errors should significantly under-react to current innovations (e.g. $\hat{\beta}_1 < 0$) and over-react to lagged innovations (e.g. $\hat{\beta}_M > 0$).

We begin by comparing forecast errors in the NoInfo and NoInfoTarget treatment in Panel A of Table 7. NoInfo participants exhibit an upward bias in their errors ($\hat{\alpha_1} > 0$), and significantly over react to current and lagged innovations, suggesting a considerable over-reliance on historical information. Providing additional information about the targets reduces the upward bias in inflation. It does, however, make subjects over-reactive to current innovations when forecasting inflation. Otherwise, it does not appear to significantly alter how subjects forecast.

The effects of providing target information in the QualInfo treatment is presented in Panel B. In both the QualInfo and QualInfo treatment, we observe a significant upward bias in output gap forecasts and inexperienced inflation forecasts. QualInfo subjects tend to under-react to current innovations and over-react to one- to four-period lagged innovations. Providing information about the central bank's targets encourages more responsiveness to current innovations but has mixed effects for lagged innovations. Output gap forecasts become less reliant on two-, three-, and four-period lagged innovations, while inflation forecasts become more reliant on one- and three-period lagged innovations.

We next the effects of providing qualitative information about the DGP when subjects already know the central bank's target (NoInfoTarget vs QualInfoTarget). The results are presented in Panel C of Table 8. When subjects only know the central bank's targets (NoInfoTarget), there is a significant upward bias in forecast errors. Output gap forecasts and inexperienced inflation forecasts tend to over-react to current and most lagged innovations. Providing qualitative information about the DGP does not have an effect on the bias, except for increasing the error on inflation forecasts.

Qualitative information does not have a consistent effect on how participants react to current or lagged innovations.

Finally, we evaluate the effects of providing precise quantitative information about the DGP when participants already have a qualitative understanding of the economy (QualInfoTarget vs. QuantInfo). In the QualInfoTarget treatment, forecasts are biased upward, especially when subjects are inexperienced. Subjects significantly over-react to most lagged innovations when forming their forecast, and over-react to current innovations when experienced. Providing a precise quantitative DGP to participants significantly increases the upward-bias in output gap forecasts and lowers the bias in inflation forecasts. Moreover, it reduces subjects' reaction to current innovations when forecasting inflation, but makes them more sensitive to lagged innovations when forecasting output gap. In other words, it encourages more rational forecasting for inflation and more backward-looking forecasting for output.

Observation 5: Expectations are not ex-post rational regardless of the information available about the DGP.

Observation 6: Quantitative information reduces the subject's reliance on historical information when forecasting inflation at the cost of more reliance when forecasting output gap.

4.2 Aggregate Results

What impact does information about the DGP have on macroeconomic stability? To answer this, we begin by plotting the time series of output and inflation for all sequences and repetitions in Figure 6. The time series of output is very consistent across treatments. This provides further evidence that subjects are not randomizing. If they were, the time series of a single sequence would differ considerably across treatments. Second, it suggests that providing additional details about the macroeconomy does not appear to have an important effect on median output forecasts.

We come to a different conclusion when we look at the time series of inflation. Instructions appear to play a role in shaping median inflation expectations. The paths of inflation differs more noticeably across the five treatments. Sequences 2, 4, and 6 in Repetition 2 have shock sequences that are relatively more variable (the standard deviation of the shocks exceeds 145 basis points). In these sequences, inflation is most stable in the QuantInfo treatment and most volatile in the NoInfo treatment. In the low variability shock sequences, sequences 1, 3, and 5 in Repetition 2, inflation is relatively more volatile in the QuantInfo treatment.

Importantly, we observe no explosive dynamics in the NoInfo and QualInfo treatments despite subjects not knowing the targets of the central bank or the central bank's quantitative response to inflation and output. This said, the economies are clearly not converging to the rational expectations equilibrium in any of these sessions.

Measures of the standard deviation of output and inflation relative to the REE prediction at the session-repetition level are presented as box plots in Figure 7. On average, output is between 1 to 1.8 times more volatile than in an environment rational expectations. Relative to REE, the output gap is most stable in the QualInfoTarget treatment and most volatile in the QualInfo in Repetition 1 and NoInfo in repetition 2. Inflation is 1.4 to 2 times more volatile than in an environment with rational expectations. We observe the lowest levels of inflation volatility in the QuantInfo treatment and the highest levels in the NoInfoTarget and QuantInfoTarget treatments. Wilcoxon rank-sum tests indicate that the differences in relative volatilities across treatments is statistically insignificant in most treatment comparisons. The only exception is that inflation volatility is significantly lower in the QuantInfo treatment than the QualInfoTarget treatment in repetition 2 (N=6 for each treatment, $p = 0.026$).

In the QuantInfo treatment, subjects are informed that the shock sequences have a standard deviation of 138 basis points. In low variability shock sequences like sequences 1, 3, and 5, subjects with knowledge of the DGP (QuantInfo) tend to generate more volatile median forecasts than the sequence would suggest, while their median forecasts in high variability shock sequences are less volatile. This suggests an important limitation to providing precise quantitative information about a DGP to the public. Emphasizing particular quantitative relationships and assumption that may frequently lead to biased forecasts.

For each session and repetition, we compute the central bank's relative welfare loss ratio as $\frac{var(\pi_t)+\omega var(x_t)}{var(\pi_t^{REE})+\omega var(x_t^{REE})}$ where $\omega = \frac{\kappa}{\mu} = 0.013$ is a function of the price elasticity of demand parameterized to be $\mu = 10$. The distributions of welfare losses, relative to the REE predicted welfare loss, are presented as box plots in Figure 8. We observe significantly lower welfare losses in the QuantInfo treatment than in the QualInfoTarget treatment, consistent with our above finding that inflation volatility is also lower in the QuantInfo treatment. That is, providing participants with the quantitative DGP significantly reduces inflation volatility and the welfare loss associated with it.

Observation 7: Quantitative information reduces inflation volatility and improves aggregate welfare.

5 Discussion

Many countries have introduced national education initiatives to improve financial literacy. Particularly progressive countries in this regard, with established programs targeting the education system exist in Japan, Czech Republic, Brazil, India, Israel, and Indonesia. These financial education initiatives are often run by central banks or a government ministry associated with finance or financial services. While another 28 countries are planning a national strategy, few countries claim to employ any significant economic literacy component (OECD, 2015). There is little significant planning regarding macroeconomic literacy, which we believe has its own distinct purpose and place

alongside teaching financial literacy.

The results from our simple experiment suggest that providing people more information about the economy that they are interacting in does not necessarily improve their ability to forecast. Qualitative information reduces inexperienced output gap errors but increases inexperienced inflation forecast errors. For experienced subjects, it does not significantly improve forecast accuracy. Providing the central bank's targets when subjects have no information about the DGP improves experienced output forecast accuracy but worsens inflation accuracy. With already a qualitative understanding of the economy, providing information about targets worsens forecast accuracy.

Quantitative information does, however, significantly improve inflation forecasting and reduces disagreement about inflation. Updating in response to past forecast errors is significantly greater when subjects have precise information about the economy. This, in turn, generates less volatility in inflation expectations and, consequently, inflation. We find that subjects receiving quantitative information over-react less to current and lagged innovations than those receiving no or qualitative information. We speculate that this is a consequence of QuantInfo subjects better understanding the stabilizing effects of monetary policy on inflation. In other words, the expectations channel of monetary policy is stronger when subjects have precise quantitative information about their environment.

There are numerous challenges to providing quantitative macroeconomic training. The public's education and interest are obvious factors in the efficacy of this endeavor. The choice of which of many central bank models of the economy to be taught to the public requires attention. There is a clear trade off between accuracy and comprehensibility when presenting models to the public. More research is needed about how the general public would digest quantitative macroeconomic education. Nonetheless our experiment sheds light on the potential advantages of more precise training for the public.

Methodologically, we do not take a stand on the *right* way to provide instructions in learning-to-forecast experiments. If the research question is about how people learn to forecast in an environment where they do not know the parameters of the model, then it makes sense to only provide qualitative instructions. If, however, the experiment is interested in testing a theory where agents are assumed to have full information about the underlying data-generating process, then it is appropriate to give subjects such quantitative information. Our work shows that quantitative information is still insufficient to generate the rational expectations assumed in most macroeconomic models. In a related project, Mokhtarzadeh and Petersen (2017) study expectation formation in an environment where subjects have quantitative information about the economy's data-generating process. They find that expectations are significantly nudged to the rational expectations equilibrium by providing simple-to-use macroeconomic projections of output and inflation. While macroeconomic education can improve how people forecast, simpler-to-use information like projections are consistently more effective at managing expectations.

6 References

1. Adam, K. (2007). Experimental evidence on the persistence of output and inflation. The Economic Journal, 117(520), 603-636.

2. Ahrens, S., Lustenhouwer, J., and Tettamanzi, M. (2017). The Stabilizing Role of Forward Guidance: A Macro Experiment.

3. Arifovic, J., Hommes, C. H., and Salle, I. (2016). Learning to believe in simple equilibria in a complex OLG economy-evidence from the lab (No. 16-06). Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance.

4. Arifovic, J. and Petersen, L. (2017). Stabilizing expectations at the zero lower bound: Experimental evidence. Journal of Economic Dynamics and Control, 82, 21-43.

5. Assenza, T., Heemeijer, P., Hommes, C. H., and Massaro, D. (2013). Individual expectations and aggregate macro behavior.

6. Burke, M. A., and Manz, M. (2014). Economic literacy and inflation expectations: evidence from a laboratory experiment. Journal of Money, Credit and Banking, 46(7), 1421-1456.

7. Bruine de Bruin, W., Vanderklaauw, W., Downs, J. S., Fischhoff, B., Topa, G., and Armantier, O. (2010). Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy. Journal of Consumer Affairs, 44(2), 381-402.

8. Bao, T., Duffy, J., and Hommes, C. (2013). Learning, forecasting and optimizing: An experimental study. European Economic Review, 61, 186-204.

9. Bao, T. and Duffy, J. (2016). Adaptive versus eductive learning: Theory and evidence. European Economic Review, 83, 64-89.

10. Branch, W. A., and Evans, G. W. (2006). A simple recursive forecasting model. Economics Letters, 91(2), 158-166.

11. Coibion, O., and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review, 105(8), 2644-2678.

12. Coibion, O., Gorodnichenko, Y., and Kamdar, R. (2017). The Formation of Expectations, Inflation and the Phillips Curve (No. w23304). National Bureau of Economic Research.

13. Colasante, A., Palestrini, A., Russo, A., and Gallegati, M. (2017). Adaptive expectations versus rational expectations: Evidence from the lab. International Journal of Forecasting, 33(4), 988-1006.

14. Colasante, A., Alfarano, S., Camacho, E., and Gallegati, M. (2018). Long-run expectations in a learning-to-forecast experiment. Applied Economics Letters, 25(10), 681-687.

15. Cornand, C., and M'baye, C. K. (2016). Does inflation targeting matter? An experimental investigation. Macroeconomic Dynamics, 1-40.

16. Cornand, C., and M?baye, C. K. (2016). Band or point inflation targeting? An experimental approach. Journal of Economic Interaction and Coordination, 1-27.

17. Croushore, D. D. (1997). The Livingston Survey: still useful after all these years. Business Review-Federal Reserve Bank of Philadelphia, 2, 1.

18. Eusepi, S. and Preston, B. (2010). Central bank communication and expectations stabilization. American Economic Journal: Macroeconomics, 2(3), 235-271.

19. Fluch, M. (2007). Selected Central Banks? Economic and Financial Literacy Programs. Monetary Policy and the Economy, Q 3, 85-103.

20. 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.

21. Hennequin, M., and Hommes, C. (2018). Managing bubbles in experimental asset markets with monetary policy. Working paper.

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

23. Hommes, C., and Lux, T. (2013). Individual expectations and aggregate behavior in learning-to-forecast experiments. Macroeconomic Dynamics, 17(2), 373-401.

24. Hommes, C., and Makarewicz, T. (2017). Price level versus inflation targeting under heterogeneous expectations: a laboratory experiment. Working paper.

25. Hommes, C., Massaro, D., and Salle, I. (2015). Monetary and Fiscal Policy Design at the Zero Lower Bound: Evidence from the Lab. CeNDEF working paper, (15-11).

26. Hommes, C., Sonnemans, J., Tuinstra, J., and van de Velden, H. (2005). A strategy experiment in dynamic asset pricing. Journal of Economic Dynamics and Control, 29(4), 823-843.

27. Hommes, C., Sonnemans, J., Tuinstra, J., and Van De Velden, H. (2007). Learning in cobweb experiments. Macroeconomic Dynamics, 11(S1), 8-33.

28. Jappelli, T. (2010). Economic literacy: An international comparison. The Economic Journal, 120(548), F429-451.

29. Kryvtsov, O., and Petersen, L. (2013). Expectations and monetary policy: experimental evidence (No. 2013-44). Bank of Canada Working Paper.

30. Lim, S., Prescott, E., and Sunder, S. (1994). Stationary Solution to the Overlapping Generations Model of Fiat Money: Experimental Evidence. Empirical Economics, 19(2), 255-277.

31. Malmendier, U. and Nagel, S. (2015). Learning from inflation experiences. The Quarterly Journal of Economics, 131(1), 53-87.

32. Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. NBER macroeconomics annual, 18, 209-248.

33. Marimon, R. and Sunder, S. (1993). Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence. Econometrica, 61(5), 1073-1108.

34. Marimon, R. Spear, S., and Sunder, S. (1993). Expectationally-Driven Market Volatility: An Experimental Study. Journal of Economic Theory, 61(1), 74-103.

35. Marimon, R. and Sunder, S. (1994). Expectations and Learning Under Alternative Monetary Regimes: An Experimental Approach. Economic Theory, 4, 131-162.

36. Marimon, R. and Sunder, S. (1995). Does a Constant Money Growth Rule Help Stabilize Inflation?: Experimental Evidence. Carnegie-Rochester Conference Series on Public Policy, 43, 111-156.

37. Mauersberger, F. (2017). The Curse of a Simplified Macroeconomic Model - An Experimental Study.

38. Mokhtarzadeh, F., and Petersen, L. (2017). Coordinating expectations through central bank projections.

39. Petersen, L. (2014). Forecast error information and heterogeneous expectations in learning-to-forecast macroeconomic experiments. In Experiments in Macroeconomics (pp. 109-137). Emerald Group Publishing Limited.

40. Pettit, J., Kephart, C., Friedman, D. and Oprea, R. (2014). Software for Continuous Game Experiments. Experimental Economics, 17, 631-648

41. Pfajfar, D. and Žakelj, B. (2014). Experimental evidence on inflation expectation formation. Journal of Economic Dynamics and Control, 44, 147-168.

42. Pfajfar, D.and Žakelj, B. (2016a). Uncertainty in forecasting inflation and monetary policy design: Evidence from the laboratory. International Journal of Forecasting, 32(3), 849-864.

43. Pfajfar, D. and Žakelj, B. (2016b). Inflation expectations and monetary policy design: Evidence from the laboratory. Macroeconomic Dynamics, 1-41.

44. Roberts, J. M. (1998). Inflation expectations and the transmission of monetary policy. Divisions of Research and Statistics and Monetary Affairs, Federal Reserve Board.

45. Slobodyan, S. and Wouters, R. (2012). Learning in an estimated medium-scale DSGE model. Journal of Economic Dynamics and Control, 36(1), 26-46.

46. Sonnemans, J. and Tuinstra, J. (2010). Positive expectations feedback experiments and number guessing games as models of financial markets. Journal of Economic Psychology, 31, 964-984.

47. Organization for Economic Co-operation and Development (2015). Comparative Tables on the Implementation of National Strategies for Financial Education.

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

7 Tables and Figures

Table 1: Instruction content about experimental data-generating process across learning-to-forecast experiments[I]

Type	No Instructions	Qualitative Instructions	Quantitative Instructions
New Keynesian	Adam (2007)	Pfajfar and Žakelj (2014) Cornand and M'Baye (2015, 2017) Assenza et al. (2015) Hommes et al. (2015) Pfajfar and Žakelj (2016a,b) Ahrens et al. (2017) Mauersberger (2017) Hommes and Makarewicz, T. (2017)	Petersen (2014)* Kryvtsov and Petersen (2015)* Arifovic and Petersen (2017) Mokhtarzadeh and Petersen (2017)
Other		Hommes et al. (2005, 2007) Heemeijer et al. (2009) Sonnemans and Tuinstra (2010) Bao et al. (2013) Arifovic et al. (2016) Bao et al. (2017) Colasante et al. (2017) Colasante et al. (2018) Hennequin and Hommes (2018)	Marimon and Sunder (1993, 1994, 1995) Marimon et al. (1993) Lim et al. (1994) Nagel (1995) Bao and Duffy (2016)

(I) This table presents a non-exhaustive list of learning-to-forecast experiments where subjects are incentivized to forecast accurately. *Subjects were provided qualitative instructions during the instruction phase and able to look up quantitative instructions during the experiment.

Table 2: Information presented to subjects across treatments

Treatment	Number of sessions	Qualitative relationship between variables	Quantitative relationship between variables	CB targets
NoInfo	6	no	no	no
NoInfo-Target	6	no	no	yes
QualInfo	6	yes	no	no
QualInfo-Target	6	yes	no	yes
QuantInfo	6	yes	yes	yes

Table 3: Effects of information on log absolute forecast errors and disagreement

	$log\|FEx_{i,t+1}\|$		$log\|FE\pi_{i,t+1}\|$		$log(s.d.E_t x_{t+1})$		$log(s.d.E_t \pi_{t+1})$	
Repetition	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2
NoInfo-Target	0.011	-0.106**	-0.010	0.200***	0.175***	0.125**	-0.030	0.188***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)
QualInfo	0.025	-0.119***	-0.016	-0.005	0.169**	0.014	0.100	0.116
	(0.05)	(0.04)	(0.05)	(0.05)	(0.07)	(0.06)	(0.07)	(0.07)
QualInfo-Target	0.009	-0.077*	0.169***	0.265***	0.076	0.083	0.260***	0.361***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)
QuantInfo	0.046	-0.030	-0.410***	-0.316***	0.389***	0.549***	-0.427***	-0.328***
	(0.05)	(0.05)	(0.04)	(0.05)	(0.11)	(0.10)	(0.07)	(0.07)
α	4.199***	4.297***	3.419***	3.481***	3.831***	3.845***	3.365***	3.340***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)
N	5935	5965	5911	5937	900	900	900	900
χ^2	1.113	10.14	201.2	223.9	18.34	37.47	122.8	151.0

(I) This table presents results from a series of mixed effects panel regressions. NoInfoTarget, QualInfo, QualInfoTarget, and QuantInfo are treatment-specific dummies. α denotes the estimated constant which can be interpreted as the mean value of the dependent variable in the NoInfo treatment. Robust standard errors are employed. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 4: Effects of information on log absolute forecast errors and disagreement

Panel A	$log\|FEx_{i,t+1}\|$		$log\|FE\pi_{i,t+1}\|$		$log(s.d.E_t x_{t+1})$		$log(s.d.E_t \pi_{t+1})$	
Treatment	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2
NoInfo$^{\alpha}$	4.199***	4.297***	3.419***	3.481***	3.831***	3.845***	3.365***	3.340***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)
NoInfo×Target†	-0.002	-0.093**	0.047	0.156***	0.175***	0.125**	-0.030	0.188***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)
N	2325	2331	2310	2324	360	360	360	360
χ^2	0.00167	4.307	1.062	10.98	7.418	4.329	0.172	7.542

Panel B	$log\|FEx_{i,t+1}\|$		$log\|FE\pi_{i,t+1}\|$		$log(s.d.E_t x_{t+1})$		$log(s.d.E_t \pi_{t+1})$	
Treatment	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2
NoInfo × Target$^{\alpha}$	4.197***	4.204***	3.466***	3.637***	4.006***	3.971***	3.335***	3.528***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)
QualInfo × Target†	0.011	0.016	0.122***	0.109**	-0.099*	-0.043	0.290***	0.174***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)
N	2373	2389	2368	2379	360	360	360	360
χ^2	0.0612	0.129	7.213	5.844	3.125	0.609	20.00	9.823

Panel C	$log\|FEx_{i,t+1}\|$		$log\|FE\pi_{i,t+1}\|$		$log(s.d.E_t x_{t+1})$		$log(s.d.E_t \pi_{t+1})$	
Treatment	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2
QualInfo$^{\alpha}$	4.246***	4.267***	3.403***	3.475***	4.220***	4.394***	3.465***	3.456***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.09)	(0.08)	(0.05)	(0.05)
QualInfo × Target†	-0.037	-0.047	0.185***	0.271***	-0.313***	-0.466***	0.160**	0.245***
	(0.05)	(0.05)	(0.05)	(0.04)	(0.10)	(0.09)	(0.07)	(0.06)
N	2409	2425	2402	2413	360	360	360	360
χ^2	0.615	1.046	16.16	37.21	9.567	24.98	5.689	15.90

Panel D	$log\|FEx_{i,t+1}\|$		$log\|FE\pi_{i,t+1}\|$		$log(s.d.E_t x_{t+1})$		$log(s.d.E_t \pi_{t+1})$	
Treatment	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2
QualInfo × Target$^{\alpha}$	4.209***	4.220***	3.588***	3.746***	3.907***	3.928***	3.625***	3.701***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
QuantInfo	0.037	0.047	-0.579***	-0.581***	0.313***	0.466***	-0.687***	-0.689***
	(0.05)	(0.05)	(0.04)	(0.04)	(0.10)	(0.09)	(0.06)	(0.06)
N	2409	2425	2401	2409	360	360	360	360
χ^2	0.615	1.046	177.6	180.6	9.567	24.98	115.2	139.7

(I) This table presents results from a series of mixed effects panel regressions. α denotes the estimated constant (estimated mean of the treatment) and † is the additional estimated mean effect of the treatment. Robust standard errors are employed. *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table 5: Forecasting Heuristics

Model Class	Heuristic Name	Model	Parameterization
M1	Ex-ante rational	$E_{i,t}x_{t+1} = 0.47r_t^n + 0.83\epsilon_t$	
		$E_{i,t}\pi_{t+1} = 0.14r_t^n + 0.25\epsilon_t$	
M2	Ex-post rational	$E_{i,t}x_{t+1} = x_{t+1}$	
		$E_{i,t}\pi_{t+1} = \pi_{t+1}$	
M3	Naive	$E_{i,t}x_{t+1} = x_{t-1}$	
		$E_{i,t}\pi_{t+1} = \pi_{t-1}$	
M4	Trend Chasing	$E_{i,t}x_{t+1} = x_{t-1} + \tau(x_{t-1} - x_{t-2})$	$\tau = 0.1, ..., 1.5$
		$E_{i,t}\pi_{t+1} = \pi_{t-1} + \tau(\pi_{t-1} - \pi_{t-2})$	$\tau = 0.1, ..., 1.5$
M5	Constant Gain	$E_{i,t}x_{t+1} = x_{t-1} + \gamma(E_{i,t-2}x_{t-1} - x_{t-1})$	$\gamma = 0.1, .., 1.5$
		$E_{i,t}\pi_{t+1} = \pi_{t-1} + \gamma(E_{i,t-2}\pi_{t-1} - \pi_{t-1})$	$\gamma = 0.1, .., 1.5$

Table 6: Heterogeneity in Heuristics

	Output Gap		Inflation	
Repetition	Treatment	Score	Treatment	Score
Repetition 1	QuantInfo	0.27	QuantInfo	0.40
	QualInfo	0.39	NoInfo	0.49
	NoInfoTarget	0.43	QualInfo	0.51
	NoInfo	0.43	QualInfoTarget	0.52
	QualInfoTarget	0.45	NoInfoTarget	0.52
Repetition 2	QuantInfo	0.37	QuantInfo	0.45
	QualInfoTarget	0.40	NoInfoTarget	0.45
	QualInfo	0.41	QualInfo	0.50
	NoInfo	0.42	QualInfoTarget	0.51
	NoInfoTarget	0.43	NoInfo	0.53

Table 7: Rationality

	Panel A					Panel B			
	NoInfo vs. NoInfoTarget					QualInfo vs. QualInfoTarget			
	FEx_{t+1}		$FE\pi_{t+1}$			FEx_{t+1}		$FE\pi_{t+1}$	
	Rep 1	Rep 2	Rep 1	Rep 2		Rep 1	Rep 2	Rep 1	Rep 2
ϵ_t	-0.042	0.144***	-0.059***	0.054**	ϵ_t	-0.178***	-0.036	-0.016	0.016
	(0.04)	(0.04)	(0.02)	(0.03)		(0.04)	(0.04)	(0.03)	(0.02)
$\epsilon_t \times$NoInfoTarget	0.060	0.062	0.099***	0.093**	$\epsilon_t \times$QualInfoTarget	0.137**	0.207***	0.052	0.164***
	(0.06)	(0.06)	(0.03)	(0.04)		(0.06)	(0.06)	(0.04)	(0.03)
ϵ_{t-1}	0.197***	0.108***	0.041*	-0.005	ϵ_{t-1}	0.193***	0.131***	-0.035	-0.022
	(0.04)	(0.04)	(0.02)	(0.03)		(0.04)	(0.03)	(0.02)	(0.02)
$\epsilon_{t-1} \times$NoInfoTarget	-0.000	-0.004	0.025	0.055	$\epsilon_{t-1} \times$QualInfoTarget	0.038	-0.045	0.127***	0.092***
	(0.05)	(0.05)	(0.03)	(0.04)		(0.05)	(0.05)	(0.03)	(0.03)
ϵ_{t-2}	0.251***	0.173***	0.012	0.018	ϵ_{t-2}	0.288***	0.259***	0.072***	0.069***
	(0.03)	(0.03)	(0.02)	(0.02)		(0.03)	(0.03)	(0.02)	(0.02)
$\epsilon_{t-2} \times$NoInfoTarget	-0.036	-0.027	0.034	0.053**	$\epsilon_{t-2} \times$QualInfoTarget	-0.039	-0.130***	0.015	-0.002
	(0.05)	(0.05)	(0.03)	(0.03)		(0.05)	(0.05)	(0.03)	(0.03)
ϵ_{t-3}	0.107***	0.187***	-0.021	0.071***	ϵ_{t-3}	0.068**	0.195***	-0.040**	0.037**
	(0.03)	(0.03)	(0.02)	(0.02)		(0.03)	(0.03)	(0.02)	(0.01)
$\epsilon_{t-3} \times$NoInfoTarget	-0.031	-0.047	0.066**	-0.023	$\epsilon_{t-3} \times$QualInfoTarget	-0.015	-0.096**	0.051*	0.010
	(0.05)	(0.04)	(0.03)	(0.03)		(0.04)	(0.04)	(0.03)	(0.02)
ϵ_{t-4}	0.089***	0.134***	-0.001	0.060***	ϵ_{t-4}	0.094***	0.181***	0.006	0.051***
	(0.03)	(0.03)	(0.02)	(0.02)		(0.03)	(0.03)	(0.02)	(0.02)
$\epsilon_{t-4} \times$NoInfoTarget	-0.093**	0.036	-0.008	0.017	$\epsilon_{t-4} \times$QualInfoTarget	-0.061	-0.082*	-0.010	-0.017
	(0.04)	(0.04)	(0.03)	(0.03)		(0.04)	(0.05)	(0.03)	(0.03)
NoInfoTarget	0.325	-2.271	1.837	-6.499*	QualInfoTarget	-6.200	-0.991	2.381	4.033
	(6.18)	(6.00)	(3.30)	(3.75)		(6.19)	(5.79)	(3.87)	(3.52)
α	25.022***	5.738	10.531***	2.788	α	34.232***	6.610*	7.453**	-0.341
	(4.29)	(4.34)	(2.29)	(2.70)		(4.44)	(3.93)	(2.94)	(2.05)
N	1852	1851	1852	1851	N	1915	1927	1915	1927
χ^2	177.1	242.8	58.94	122.7	χ^2	274.1	247.0	88.69	165.5

(I) This table presents results from a series of mixed effects panel regressions. NoInfoTarget, QualInfo, QualInfoTarget and QuantInfo are treatment-specific dummies. α denotes the estimated constant. Robust standard errors are employed. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 8: Rationality

	Panel C					Panel D			
	NoInfoTarget vs. QualInfoTarget					QualInfoTarget vs. QuantInfo			
	FEx_{t+1}		$FE\pi_{t+1}$			FEx_{t+1}		$FE\pi_{t+1}$	
	Rep 1	Rep 2	Rep 1	Rep 2		Rep 1	Rep 2	Rep 1	Rep 2
ϵ_t	0.018	0.206***	0.040*	0.147***	ϵ_t	-0.042	0.170***	0.035	0.180***
	(0.04)	(0.04)	(0.02)	(0.02)		(0.04)	(0.04)	(0.03)	(0.03)
$\epsilon_t \times$QualInfoTarget	-0.059	-0.036	-0.005	0.033	$\epsilon_t \times$QuantInfo	-0.578***	-0.318**	-0.156***	-0.278***
	(0.06)	(0.06)	(0.03)	(0.04)		(0.12)	(0.14)	(0.03)	(0.03)
ϵ_{t-1}	0.197***	0.104***	0.066***	0.050**	ϵ_{t-1}	0.231***	0.086**	0.092***	0.070***
	(0.04)	(0.04)	(0.02)	(0.02)		(0.04)	(0.04)	(0.03)	(0.03)
$\epsilon_{t-1}\times$ QualInfoTarget	0.034	-0.019	0.025	0.020	$\epsilon_{t-1}\times$QuantInfo	-0.010	0.117	-0.184***	-0.149***
	(0.05)	(0.05)	(0.03)	(0.03)		(0.12)	(0.15)	(0.03)	(0.03)
ϵ_{t-2}	0.215***	0.147***	0.046***	0.072***	ϵ_{t-2}	0.250***	0.128***	0.087***	0.067***
	(0.04)	(0.03)	(0.02)	(0.02)		(0.03)	(0.03)	(0.02)	(0.02)
$\epsilon_{t-2}\times$QualInfoTarget	0.035	-0.018	0.041	-0.004	$\epsilon_{t-2}\times$QuantInfo	0.396**	0.233	-0.124***	-0.096***
	(0.05)	(0.05)	(0.03)	(0.03)		(0.19)	(0.15)	(0.02)	(0.02)
ϵ_{t-3}	0.076**	0.140***	0.045**	0.049***	ϵ_{t-3}	0.053*	0.099***	0.011	0.047**
	(0.04)	(0.03)	(0.02)	(0.02)		(0.03)	(0.03)	(0.02)	(0.02)
$\epsilon_{t-3}\times$QualInfoTarget	-0.023	-0.041	-0.034	-0.002	$\epsilon_{t-3}\times$QuantInfo	0.632***	0.369**	-0.029	-0.040*
	(0.05)	(0.04)	(0.03)	(0.03)		(0.25)	(0.16)	(0.02)	(0.02)
ϵ_{t-4}	-0.004	0.170***	-0.010	0.077***	ϵ_{t-4}	0.033	0.099***	-0.005	0.034
	(0.03)	(0.03)	(0.02)	(0.02)		(0.03)	(0.03)	(0.02)	(0.02)
$\epsilon_{t-4}\times$QualInfoTarget	0.037	-0.071	0.005	-0.043	$\epsilon_{t-4}\times$QuantInfo	0.566**	0.158	-0.000	-0.022
	(0.04)	(0.04)	(0.03)	(0.03)		(0.27)	(0.17)	(0.02)	(0.02)
QualInfoTarget	2.685	2.152	-2.533	7.403*	QuantInfo	57.626**	57.357***	-7.198***	-2.038
	(6.20)	(5.94)	(3.46)	(3.86)		(24.95)	(16.43)	(2.75)	(3.15)
α	25.347***	3.467	12.368***	-3.711	α	28.032***	5.619	9.835***	3.692
	(4.45)	(4.14)	(2.37)	(2.60)		(4.32)	(4.25)	(2.51)	(2.85)
N	1891	1898	1891	1898	N	1919	1928	1919	1928
χ^2	185.2	215.4	88.23	169.5	χ^2	163.3	111.6	275.0	182.4

(I) This table presents results from a series of mixed effects panel regressions. NoInfoTarget, QualInfo, QualInfoTarget and QuantInfo are treatment-specific dummies. α denotes the estimated constant. Robust standard errors are employed. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Figure 1: Screenshot from NoInfo and QualInfo treatments

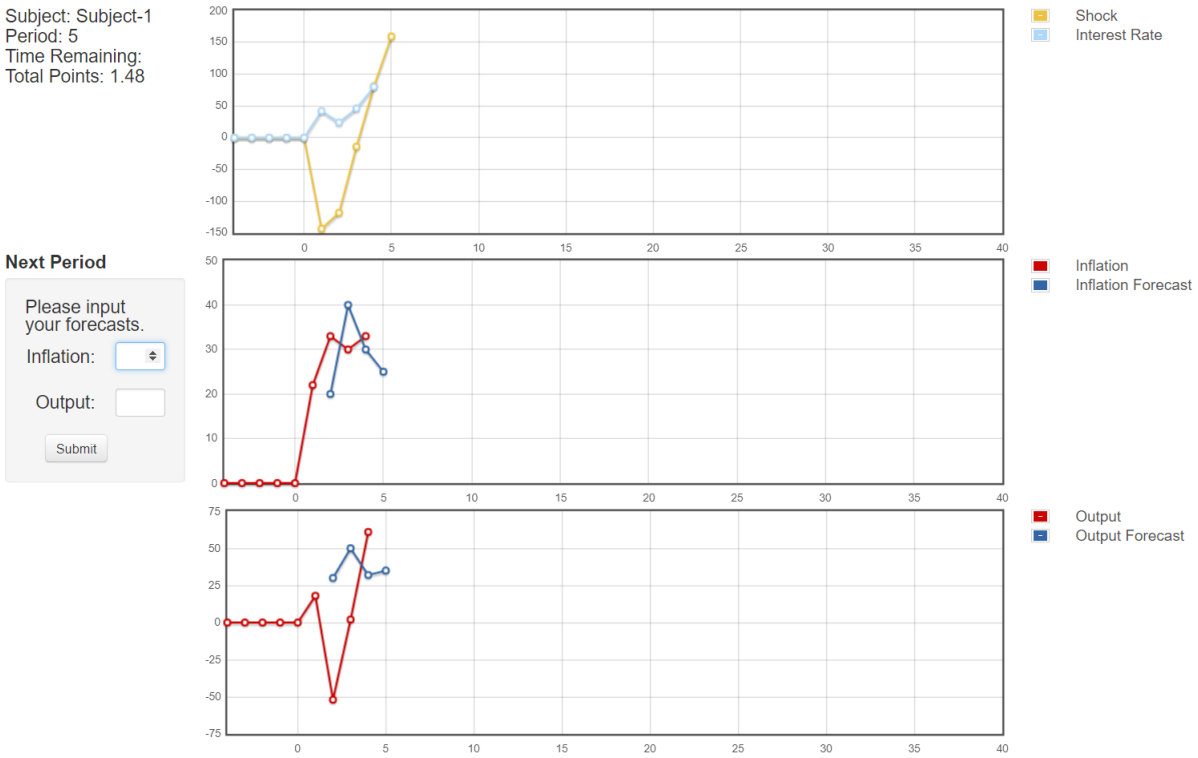

Figure 2: Empirical CDFs of output and inflation absolute forecast errors, by repetition

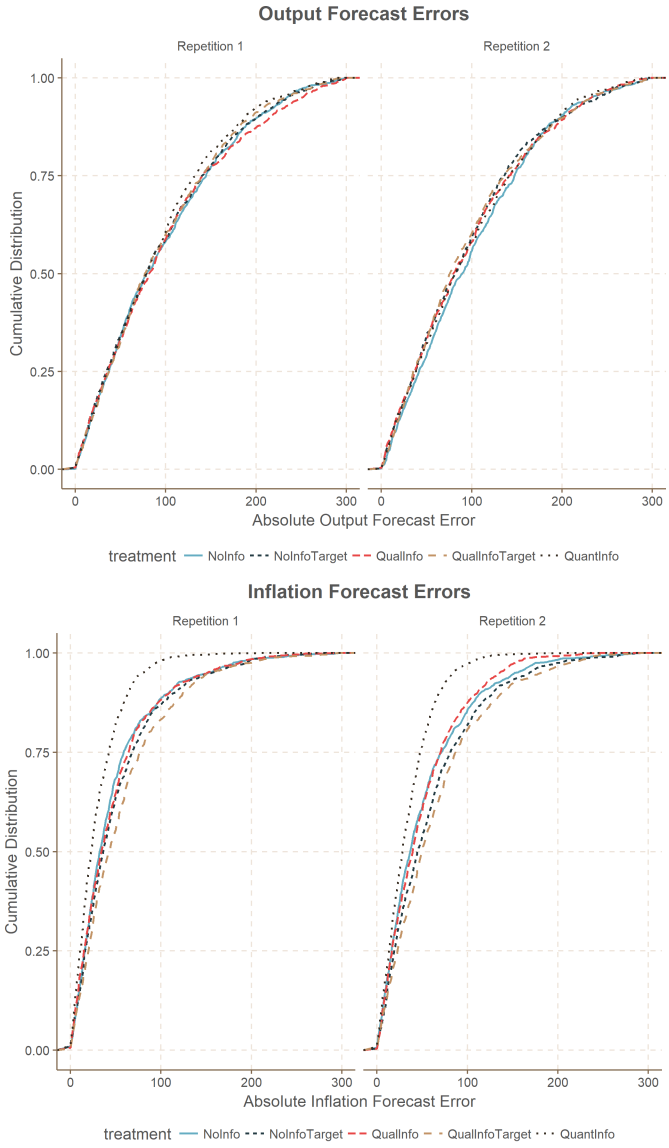

Figure 3: Empirical CDFs of output and inflation within-period disagreement, by treatment and repetition

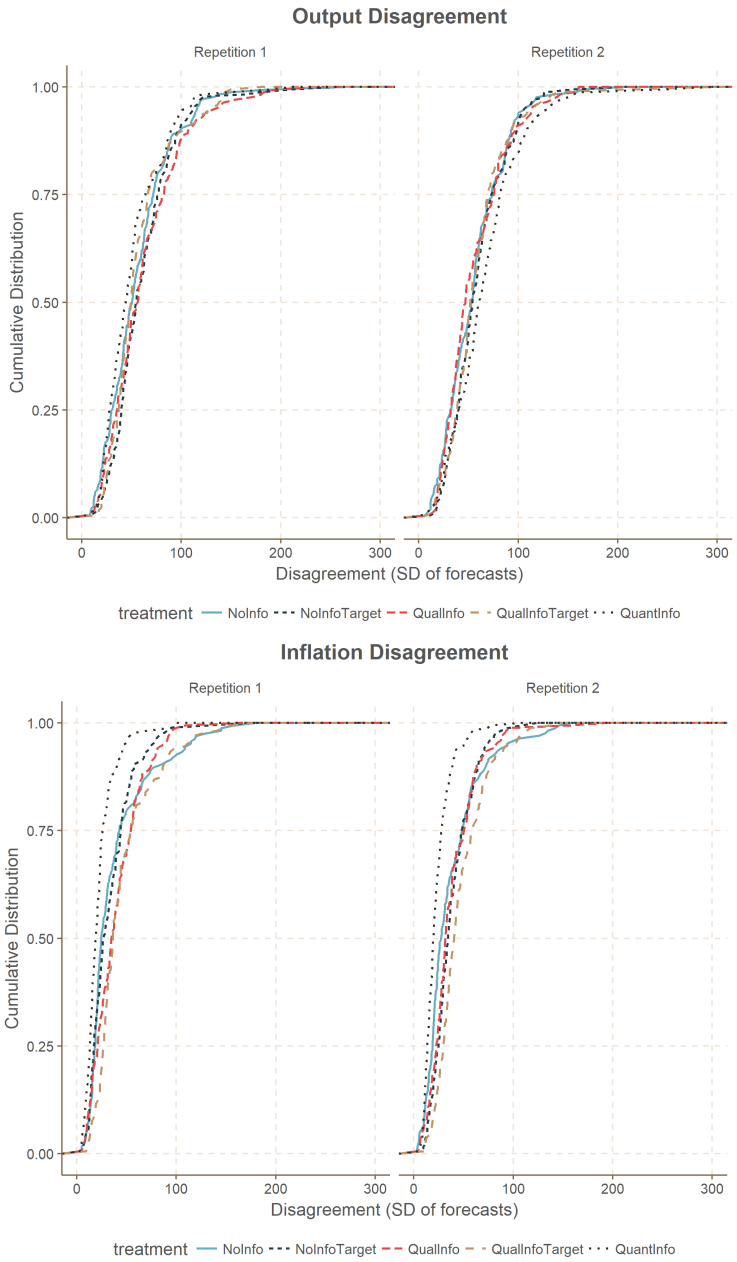

Figure 4: Box plots of output and inflation within-period disagreement, by treatment and repetition

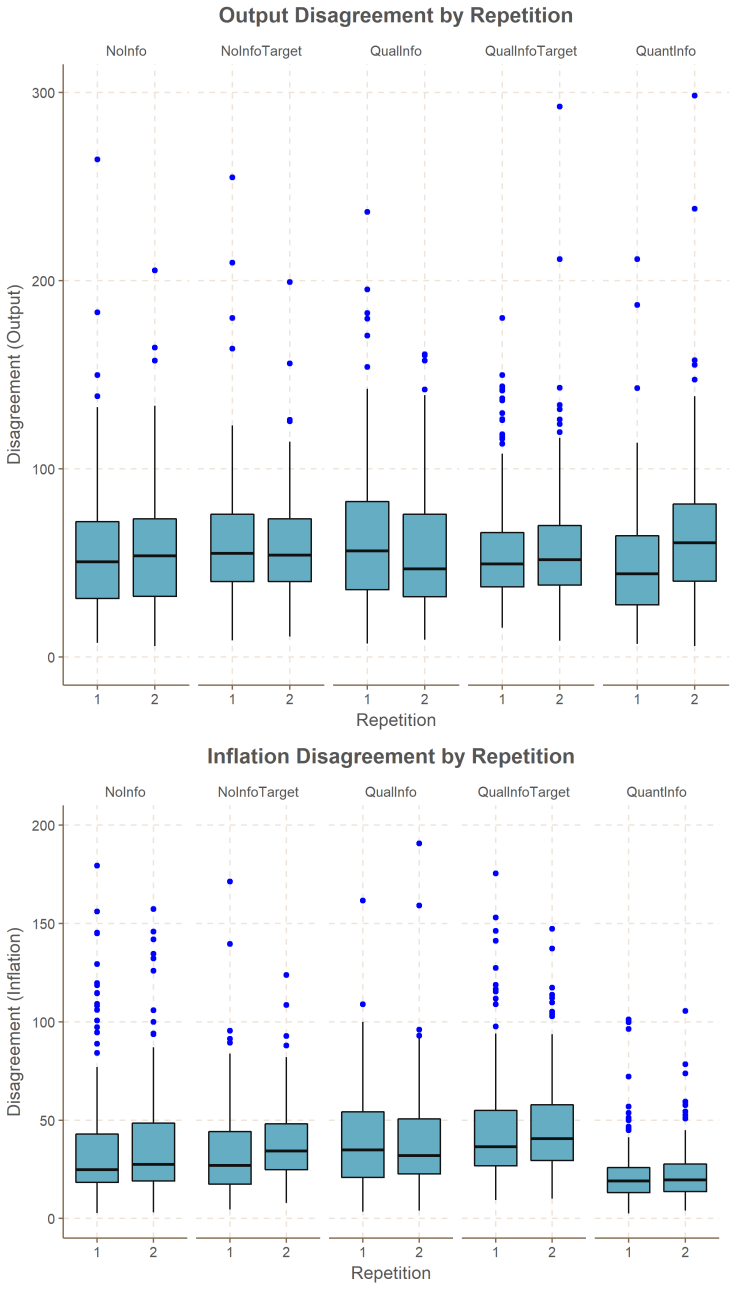

Figure 5: Distribution of heuristics, by treatment and repetition

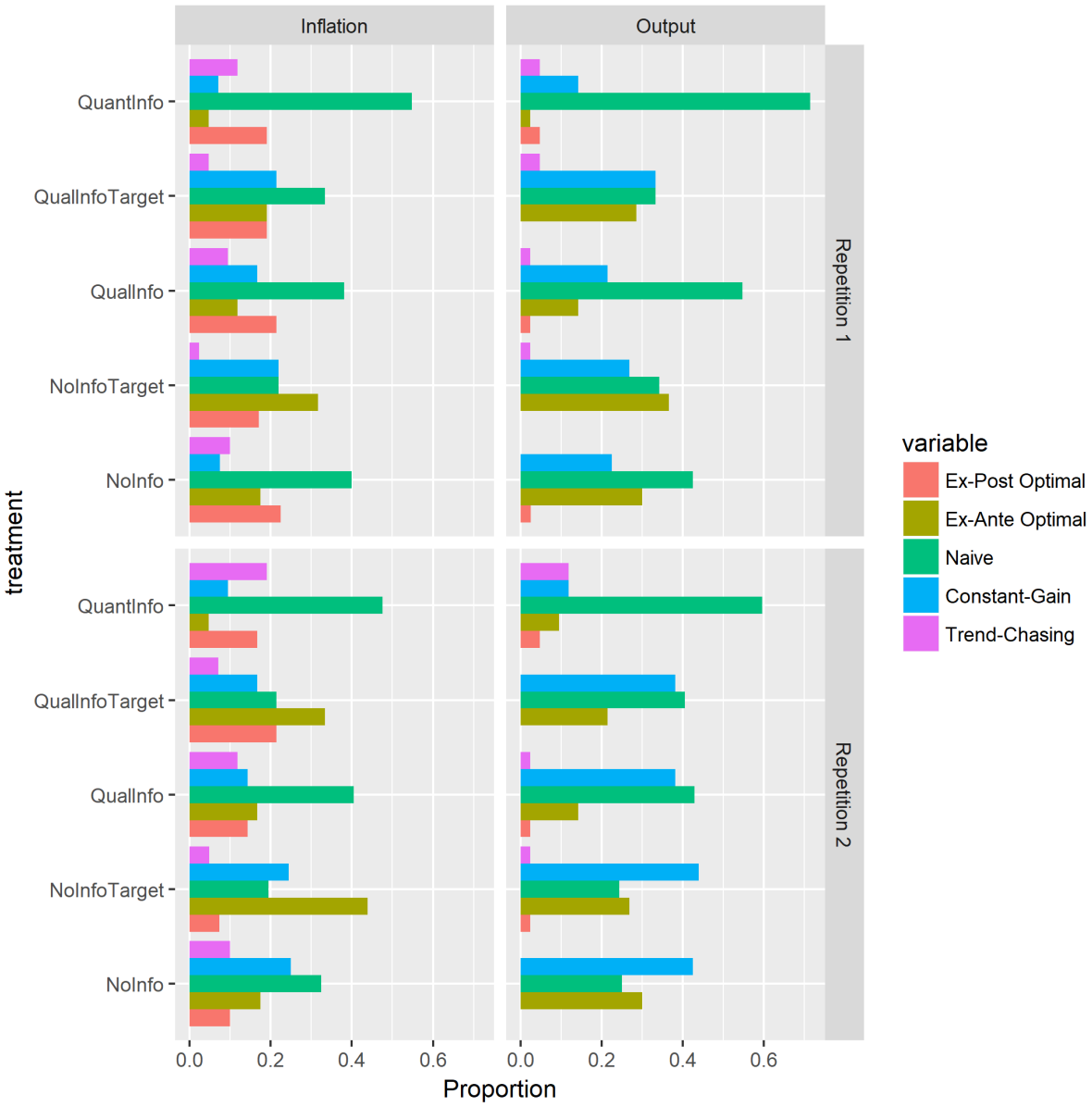

Figure 6: Time series of output and inflation, by shock sequence

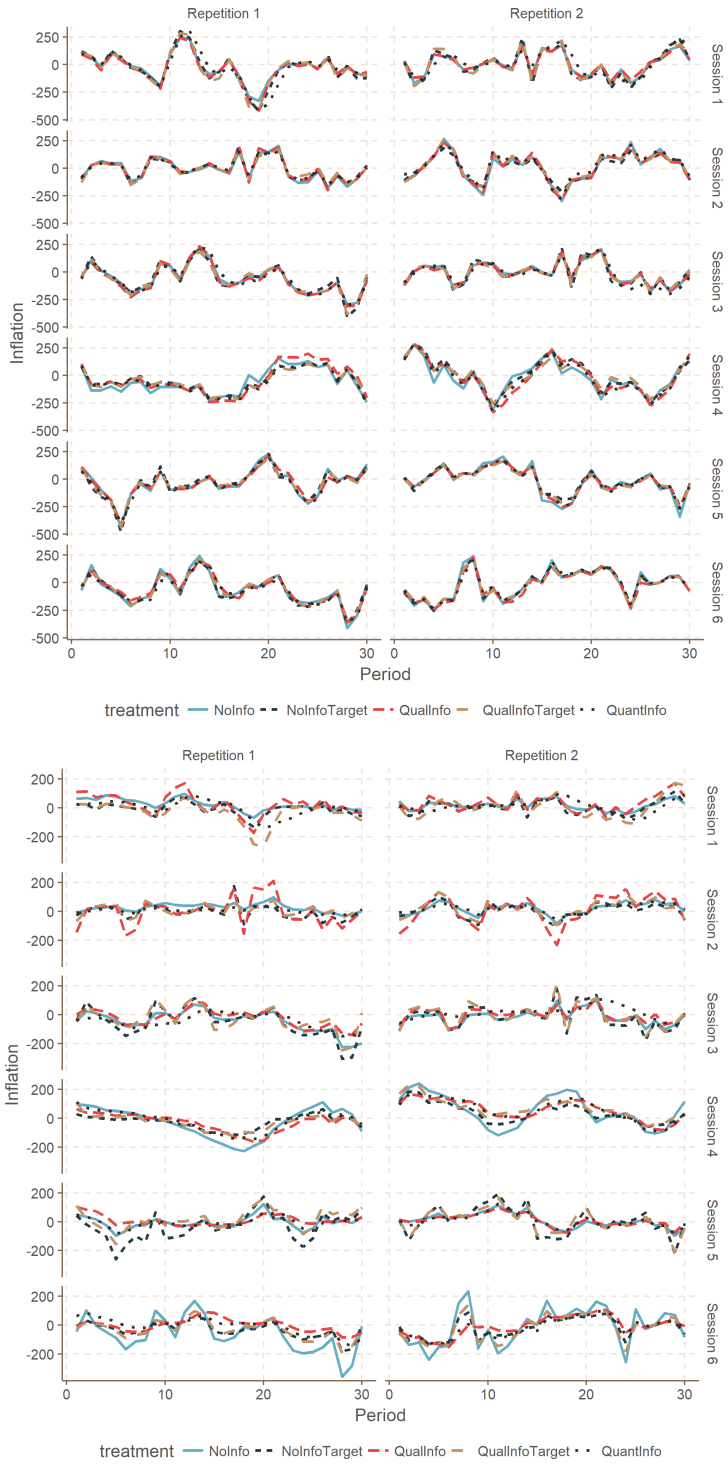

Figure 7: Box-plots of session-level standard deviations of output and inflation, by repetition

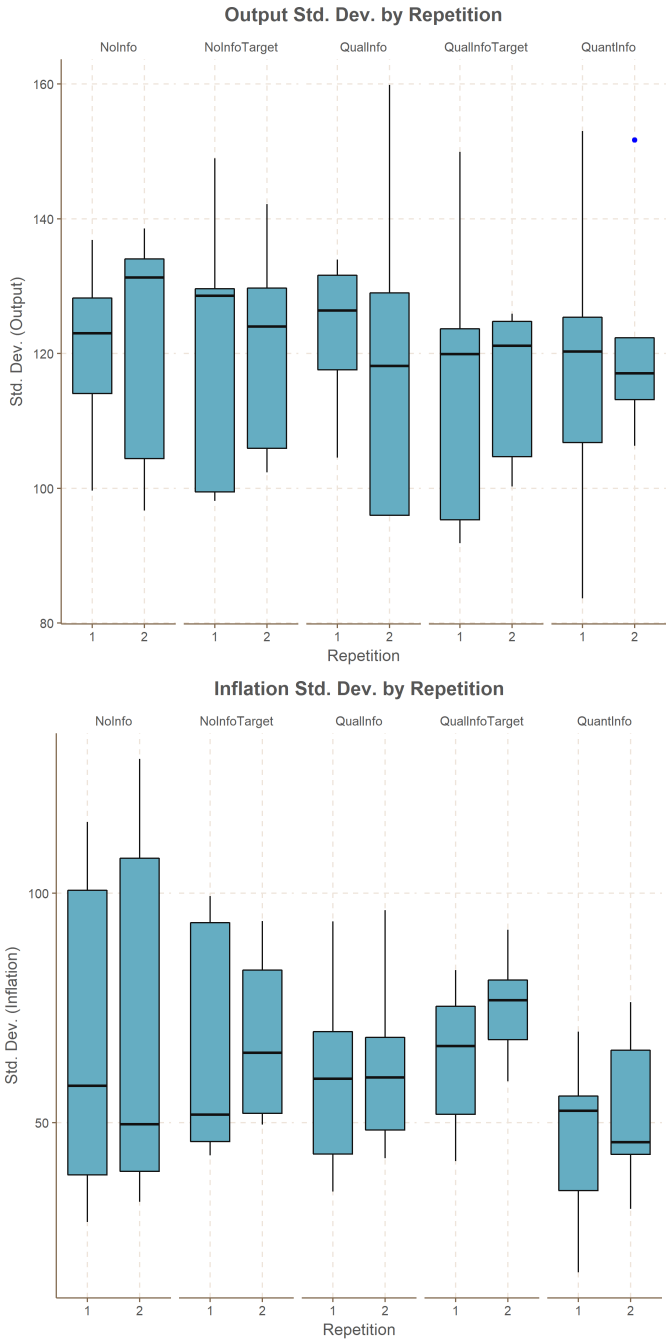

Figure 8: Box-plots of the session-level welfare loss associated with the output gap and inflation, by repetition