A Horse Race of Monetary Policy Regimes: An Experimental Investigation*

Olena Kostyshyna
Bank of Canada

Luba Petersen
Simon Fraser University
NBER

Jing Yang
Bank of Canada

December 5, 2022

Abstract

We provide a comprehensive assessment of leading monetary policy frameworks away from and at the ELB. Inflation targeting, dual mandate, average inflation targeting under 4- and 10-period horizons, price level targeting, and nominal GDP level targeting are evaluated in a laboratory setting. Contrary to theoretical prediction with full information rational expectations, participants exhibit backward-looking expectations and, consequently, rate-targeting mandates outperform level targeting. More history dependence worsens macroeconomic stability. Inflation expectations are managed better when mandates are framed in terms of inflation rates than price levels. Central bank communication significantly improves the performance of price level targeting.

JEL classifications: C9, D84, E52, E58
Keywords: expectations, monetary policy, credibility, laboratory experiment, experimental macroeconomics

* We have benefited from useful discussions with Herbert Dawid, Chris Gibbs, Seppo Honkapohja, Cars Hommes, Oleksiy Kryvtsov, Chen Lian, Tomasz Makarewicz, Rhys Mendes, Athanasios Orphanides, Isabelle Salle, Jón Steinsson, Alexander Ueberfeld, Joel Wagner, and Yang Zhang as well as seminar and conference participants at the Bank of Canada, Bielefeld University, Bundesbank, C.D. Howe Institute, Florida State University, Simon Fraser University, University College Dublin, University of California Berkeley, the Max Bell School of Public Policy Workshop on the Bank of Canada's Inflation Target Renewal. Finally, we thank Ekeme Afia, Michael Mirdamadi, Fatemeh Mokhtarzadeh, and Ryan Rholes for outstanding research assistance and Justin Baize for excellent programming support.
† Kostyshyna and Yang: Bank of Canada, 234 Wellington Street Ottawa, ON, K1A 0G9; kost@bankofcanada.ca, jyang@bankofcanada.ca; Petersen: Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada, lubap@sfu.ca.

1 Introduction

Monetary policy has evolved significantly in the aftermath of the Global Financial Crisis and the pandemic. Leading central banks, explicitly following an inflation-targeting (IT) framework for decades, drove their policy rates to their effective lower bounds to provide liquidity to markets and stimulate sluggishly low inflation expectations and economic growth. With limited traction through their policy rates, many central banks began to entertain monetary frameworks that promised to keep interest rates "lower for longer" to bolster inflation expectations. The Federal Reserve Bank, for instance, adopted an average inflation targeting (AIT) framework in August 2020 that would make up for past shortfalls in inflation by temporarily accepting above-target inflation as the US economy recovered.

Level-targeting mandates have also recently gained attention. Price-level targeting (PLT) and nominal GDP level targeting (NGDP) have been proposed as alternative policy mandates [Evans, 2012, Williams, 2017, Bernanke, 2017, Bullard, 2018]. These history-dependent regimes can be more powerful in stabilizing the economy as past misses must be made up.[1] The superior performance of PLT depends on agents' forward-looking expectations and credibility of the regime. Any advantages in the performance of such a history-dependent regime could diminish or even reverse if the expectations channel does not work well, that is, if people do not understand how the policy regime works, or if the policy regime is not credible. [Honkapohja and Mitra, 2014, 2020, Amano et al., 2020]. "People must generally understand what the central bank is doing - an admittedly high bar" [Carney, 2012]. Evidence about people's understanding of these alternative monetary policy regimes is very limited and mixed. Coibion et al. [Forthcoming] show that surveyed households in the United States had difficulty understanding the newly introduced AIT framework and distinguishing it from IT regime. Hoffmann et al. [2022] provide results that are more encouraging for AIT where German households revised their expectations in line with policy.

In this paper we provide a comprehensive assessment of five monetary policy regimes – IT, dual mandate (DM), AIT, PLT, and NGDP – in a unified experimental framework. This approach allows us to study how well people can understand and form macroeconomic expectations under competing policy frameworks both during periods of economic stability away from the ELB, during demand-driven recessions at the ELB, and during recovery. Laboratory

[1] While PLT has gained some popularity in recent monetary policy discussions, there is very limited experience with it: PLT was briefly implemented only in Sweden in the 1930s [Berg and Jonung, 1999]. The United States defined its monetary policy framework as flexible average inflation targeting only recently, and nominal GDP level targeting has not been implemented anywhere yet.

experiments are a necessary test-bed given the lack of evidence about some of these regimes in practice. Laboratory experiments have been used extensively in the literature to study various mandates in isolation or in pair-wise comparisons: inflation targeting with price-level targeting [Hommes and Makarewicz, 2021, Amano et al., 2011, Arifovic and Petersen, 2017, Salle, 2021], a dual mandate[Cornand and M'baye, 2018, Hommes et al., 2019b], and average inflation targeting [Salle, 2021]. Our paper provides a systematic evaluation of the broadest set of mandates in the literature across different underlying macroeconomic conditions.

To this end, we take the simple New Keynesian model used in the Bank of Canada's own analytical horse race to the laboratory and run an experimental horse race of the various mandates under consideration. Using a standard learning-to-forecast experimental setup [Adam, 2007, Pfajfar and Žakelj, 2014], we study how incentivized participants form expectations about future inflation and output gap. The aggregated expectations, combined with shocks to the natural rate of interest, are fed into the economies' data-generating process and drive macroeconomic dynamics. In a between-subject design, we systematically compare aggregate dynamics and expectation formation under the different monetary policy regimes over the business cycle, where different regimes are placed on even footing.

Contrary to theoretical prediction with full information rational expectations, participants exhibit backward-looking and extrapolative expectations and, consequently, rate-targeting mandates outperform level targeting. Our experiments suggest a distinct ranking of the monetary policy regimes in terms of their ability to achieve macroeconomic stability. Rate-targeting regimes such as IT, DM, and AIT significantly outperform level-targeting regimes such as PLT and NGDP in terms of their ability to minimize deviations of inflation, output gap, and nominal interest rates from the steady state. The IT regime is among the frontrunners in this monetary policy horse race despite being a predicted loser under rational expectations.

History dependence in monetary policy can be destabilizing. AIT with a shorter horizon (4-quarter) performs better than AIT with a longer horizon (10-quarter). This observation is similar to Amano et al. [2020], who find that a short horizon is optimal in the presence of backward-looking expectations. Indeed, the backward-looking expectations observed in all our treatments generate considerable inflation volatility. Mandates such as AIT-10, PLT, and NGDP require the central bank to react to past economic deviations that may no longer be relevant to the current economic situation. The targets become counter-intuitive and confusing for boundedly-rational people, and lead to a de-anchoring of expectations.

Framing also matters. Inflation expectations are better managed when mandates are framed in terms of inflation rates than price levels. Participants form more accurate and anchored forecasts about inflation when policy is formulated in terms of inflation targets rather than price level targets. This, in turn, produces more stable economic outcomes under AIT-10 than under PLT.

Our individual-level analysis further suggests that participants have difficulty understanding the basic 'directionality' associated with the various monetary policy regimes, i.e. forecasting in the rationally expected direction. This is especially evident in the level-targeting regimes. But even participants who do forecast in the correct direction do not fully internalize the stabilizing effects of monetary policy, i.e. they react too little.

Finally, credibility is more challenging to maintain in level-targeting mandates. Broadly speaking, participants 'need to see it to believe it.' While the economy regularly returns to target in the rate-targeting treatments, it takes longer in the level-targeting mandates. Participants in the PLT and NGDP treatments eventually grow skeptical about the central bank's ability to achieve its targets, especially as their economies enter into the ELB and the central bank can no longer stimulate the economy back to its intended targets.

Having observed such a poor performance of price-level targeting in the lab, we conduct a follow-up treatment (PLT Comm) to explore whether adding central bank communication of macroeconomic projections of inflation and output can improve the performance of this regime. Our results are encouraging for PLT. Compared with PLT, PLT Comm reduces economic variability by 38% in the pre-shock phase and 100% in the post-shock phase. PLT Comm even outperforms our most stable treatments in terms of losses. This improvement comes from the majority of participants forecasting in the correct direction and a reduction in extrapolative forecasting behavior. Our results suggest an important role for central bank communication in the implementation of complex mandates such as PLT.

Our paper is organized as follows. Section 2 lays out our experimental design and Section 3 presents our hypotheses. Sections 4 examines the performance of the competing monetary policy frameworks at the aggregate level, while Section 5 explore how individual-level behavior and heterogeneity in expectation formation drive our aggregate results. Results from the PLT with communication treatment are presented in Section 6 and Section 7 concludes.

2 Experimental Design

We design a laboratory experiment to collect individual-level expectations under different frameworks to inform the design of monetary policy. The data from the experiment is used to address the following questions. Do different monetary policy regimes perform in the lab as predicted by theory? Importantly, does history-dependence deliver the stability as promised by rational-expectations models? Does framing of policy objectives and the degree of history-dependence matter for the management of expectations? Are participants able to understand and incorporate monetary policy into their macroeconomic forecasts? That is, do participants update their forecasts in the correct direction and by sufficient magnitude?

2.1 Data-generating process

Our experimental environment is designed around a simple New Keynesian model that is commonly used for monetary policy analysis. We construct an economy that follows a data-generating process based on this canonical model [Woodford, 2003] that is calibrated to the Canadian economy, a model environment among many model candidates considered by the Bank of Canada for its 2021 mandate renewal [Swarbrick and Zhang, forthcoming].

The economy in which participants interact is described by the following system of equations:

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

$$x_t = \mathbb{E}_t x_{t+1} - \frac{1}{\sigma}(i_t - \mathbb{E}_t \pi_{t+1} - r_t^n) \tag{2}$$

$$r_t^n = (1-\rho)(-ln(\beta)) + \rho r_{t-1}^n + \sigma_{rn}\epsilon_t \tag{3}$$

Equation 1 describes the evolution of inflation in period t, π_t, in response to aggregate one-period-ahead inflation expectations, $\mathbb{E}_t \pi_{t+1}$, and the output gap, deviations of output from its steady state level, x_t. The output gap, given by Equation 2, is a function of aggregate expectations of one-period-ahead inflation and output gap expectations, $\mathbb{E}_t x_{t+1}$, as well as the deviations of the nominal interest rate, i_t, from the natural rate of interest, r_t^n. The natural rate of interest, described by Equation 3, is the rate of interest that keeps the economy at full employment while keeping inflation constant. The natural rate of interest is assumed to follow an AR(1) process and is subject to a sequence of demand shocks, ϵ_t.[2] Parameters

[2] In our experiments, as in related New Keynesian learning-to-forecast experiments by Arifovic and Petersen [2017], Ahrens et al. [2022], Assenza et al. [2019], Hommes et al. [2019a], Kryvtsov and Petersen [2021], we only focus on aggregate demand shocks. Pilot studies including supply shocks made the environment considerably complicated and we leave this to future research.

in our model are calibrated to quarterly data, as in Swarbrick and Zhang [forthcoming], and are consistent with Canadian data. These values are used in Kryvtsov and Petersen [2021]. β =0.994, $\sigma = 1$, $\rho = 0.8$, $\sigma_{rn} = 0.005$, $\kappa = 0.125$, π^* =0, x^* =0, r_t^{n*} =0, $\bar{r} = i^* = 60$.

To close the model, we include a policy rule that governs the evolution of the nominal interest rate, i_t. The policy rule is our key source of experimental variation. We consider six distinct ad hoc policy rules. The first three mandates we consider involve the central bank targeting various metrics of inflation and the output gap.

Under IT and DM regimes, the central bank sets the nominal interest according to the following general policy rule:

$$i_t = \bar{r} + \phi_\pi(\pi_t - \pi^*) + \phi_x(x_t - x^*) \tag{4}$$

where it seeks to minimize deviations of inflation and output gap from their targeted level of zero. Parameters ϕ_π and ϕ_x govern the reactions of the central bank to deviations of inflation and output gap from their targeted levels. The difference between IT and DM is that the weight on the output gap, ϕ_x is assumed to be considerably larger and equal to ϕ_π under a dual mandate. In the IT regime, $\phi_\pi = 5.5$ and $\phi_x = 3.0$, while in DM, $\phi_\pi = \phi_x = 4.5$.

Under the AIT regime, the central bank sets the nominal interest rate to minimize deviations of inflation from its inflation target based on the recent average inflation rate. The central bank also places some weight on the output gap when making its policy decisions. Policy coefficients are the same as in IT: $\phi_\pi = 5.5$, $\phi_x = 3.0$. We consider two horizons for average inflation —a short horizon of 4 quarters and a long horizon of 10 quarters. Our objective in studying two horizons in AIT is twofold. First, we would like to explore how AIT with different horizons perform. Such results can be useful in guiding the choice of the horizon for policymakers. Second, theory predicts that AIT approaches PLT when the horizon in computing average inflation goes to the limit of infinity. Therefore, AIT with a longer horizon may be a more feasible way to achieve results comparable to those in PLT without many of the practical challenges in implementing PLT [Amano et al., 2020]. The two AIT policy rules we implement are given by Equations 5 and 6:

$$i_t = \bar{r} + \phi_\pi(\frac{\sum_{j=0}^{3} \pi_{t-j}}{4} - \pi^*) + \phi_x(x_t - x^*) \tag{5}$$

$$i_t = \bar{r} + \phi_\pi(\frac{\sum_{j=0}^{9} \pi_{t-j}}{10} - \pi^*) + \phi_x(x_t - x^*) \tag{6}$$

Next, under the price-level targeting mandate, the central bank responds to deviations of the price level, P_t from its targeted level, P^*, as well as the output gap:

$$r_t = \bar{r} + \phi_P(P_t - P^*) + \phi_x(x_t - x^*) \tag{7}$$

where $P_t = P_{t-1} + \pi_t$. $\phi_P = 0.8$, $\phi_x = 1.3$.

Finally, a nominal GDP level targeting mandate involves the central bank instead adjusting nominal interest rates in response to deviations of the nominal GDP level, $NGDP_t$ from its targeted level, $NGDP^*$:

$$i_t = \bar{r} + \phi_{NGDP}(NGDP_t - NGDP^*) \tag{8}$$

where $NGDP_t = x_t + P_t$. ϕ_{NGDP} =1.1

Parameters in the policy rules are derived from optimizing the following loss function as implemented by Swarbrick and Zhang [forthcoming]:

$$L = \sum_{t=1}^{50} \left(\pi_t^2 + x_t^2 + 0.5(i_t - i_{t-1})^2\right) \tag{9}$$

This ad hoc loss function gives a realistic description of the goals pursued by a central bank. Central banks are concerned not only about inflation and output gap stabilization but also interest rate variation. The coefficients in the policy rules in different monetary policy regimes were chosen to minimize this loss function while putting the frameworks on comparable footing, and are in line with the theoretical horse race conducted in Swarbrick and Zhang [forthcoming].

2.2 Experimental implementation

Our experimental design follows closely the structure of previous New Keynesian learning-to-forecast experiments focused on expectation formation at the zero lower bound [Arifovic and Petersen, 2017, Hommes et al., 2019a]

General game

Our experiment consists of six independent sessions for each of the six monetary policy treatments. For each session, we invited groups of seven inexperienced participants to play the roles of professional forecasters tasked with making forecasts in 50 sequential periods.

In each period, each subject j submitted forecasts about inflation and the output gap in the subsequent period – $E_{j,t}\pi_{t+1}$ and $E_{j,t}x_{t+1}$. Actual outcomes for π_t, x_t, and i_t were determined based on the current period's realized ϵ_t and the median submitted forecasts for $t+1$ inflation and output gap according to Equations 1-3 and one of the policy rules given in Equations 4-8.

The exogenous shocks in the experimental economy were pre-drawn. This is described in the instructions to the participants. The shock sequence was chosen to implement two distinct phases in the experiment. Each session began with an initial stable phase during periods 1-19 and provided us with an opportunity to evaluate the relative performance of different policy mandates away from the effective lower bound. This phase was followed by a significant large negative demand shock in period 20 that brought the economy to the effective lower bound. The large negative demand shock dissipated rather quickly, returning to the steady-state level of zero by period 23. The remainder of the post-shock phase lasted 27 periods and enabled us to study how economies respond to and recover under the different monetary policy mandates following episodes at the ELB. Thus, we can test the stabilization properties of the monetary policy regimes during stable and unstable periods, including periods at ELB.

Treatment implementation

All treatments were parameterized to have the same steady state. At the beginning of each session, participants observed a five-period path of all variables at their steady-state levels. Once the session began, participants observed new shocks to the natural rate of interest which would drive the economy out of the steady state.

Under rational expectations, past historical data would not matter when transitioning from the steady state to the first period of the game in all regimes except AIT. Historical data would not be used in the calculation of the policy rate for IT, DM, PLT, and NGDP. However, when calculating the policy rate in AIT-4 and AIT-10 regimes, average inflation is calculated over past realized inflation. In the simulations with rational expectations presented in Section 3, average inflation is computed using the assumption that inflation is equal to steady state in periods that have not yet realized within the experimental session ("non-realized"). For example, in the simulations of AIT-4 with rational expectations, in period 1, average inflation is computed as $(0 + 0 + 0 + \pi_1)/4$.[3] If we were to implement the AIT regimes in the same way as in the RE simulations, the potency of the response to actual inflation real-

[3] In period $t = 2$, average inflation is computed as $(0+0+\pi_1+\pi_2)/4$, in period $t = 3$, average inflation is computed as $(0+\pi_1+\pi_2+\pi_3)/4$, and in period $t = 4$, average inflation is computed as $(\pi_1+\pi_2+\pi_3+\pi_4)/4$. In all periods $t > 4$, average inflation is computed based on 4 periods.

ized during the experiment would have been lower during initial periods of the experimental economy. The developments during the initial periods influence subjects' learning and the formation of their expectations. An insufficient policy response has the potential to create dynamics that were not conducive to subjects' understanding of the AIT regime and would risk the policy mandate's credibility.

Therefore, in experimentally implementing the AIT-4 and AIT-10 regimes, average inflation was computed using actual inflation realizations available during the initial three or nine periods of play, respectively.[4] This computation was explained in detail in the experimental instructions. In our view, this implementation gives the AIT policy rule the best chance to react to actual developments during the experiment to the best of its strength. Our approach to initial periods is preferable to using the assumption of steady-state inflation in "non-realized" periods.

Payoffs

Participants' earnings during the experiment are determined based on the accuracy of their inflation and output gap forecasts. The points earned by subject j in period t were based on the absolute distance between their forecasts made in period $t-1$ and realized inflation and output in period t:

$$Points_{j,t} = 0.3\left(2^{-.5|\mathbb{E}_{j,t-1}\{\pi_t\}-\pi_t|} + 2^{-.5|\mathbb{E}_{j,t-1}\{x_t\}-x_t|}\right) \tag{10}$$

Participants' total payoffs over all the forecasting periods were converted to Canadian dollars at an exchange rate of 50 cents per point.

Instructions

Participants were provided with detailed information about the economy's data-generating process, including very clear descriptions of how the central bank would set monetary policy and its impact on the economy. This information was presented both descriptively and quantitatively in the form of explicit equations. We also explained how participants' forecasts would translate into points and payoffs at the end of the experiment. The experimental instructions can be found in Online Appendix A.

[4]For example, in period $t = 1$, π_1 is used to compute deviation from inflation target, in period $t = 2$, average inflation over periods $t = 1$ and $t = 2$ is used $\frac{\pi_1+\pi_2}{2}$, etc. Starting from period $t = 4$, 4-period average inflation is used in AIT-4. And starting period $t = 10$,10-period average inflation is used in AIT-10.

Experimental interface

Participants were presented with information about their experimental economy as it evolved during the experiment. Figure B1 in Online Appendix B shows the screenshot of the computer screen seen by the subjects during the experiment. They continuously observed 4 charts presenting shocks and interest rates, inflation, inflation target and the subject's private inflation forecast, output and the subject's private output forecast, nominal GDP level and price level, and the targets of the central bank (inflation in IT, DM, AIT-4, and AIT-10, as well as the price-level target in PLT and nominal output target in NGDP). The targets were displayed continuously as a horizontal line at zero (for inflation and output gap) and 1000 for the price-level and nominal output targets.

On the left-hand side of the screen, there were two input windows where subjects submitted their forecasts of inflation and output. Subjects were given 75 seconds to submit their forecasts during the initial 10 periods and 50 seconds during the remaining periods of the experiment. If participants failed to input their forecast on time, the experiment would move on to the next round and they would simply earn zero points for their missed forecasts. The median forecast would instead be selected from the submitted forecasts. Subjects could submit any number they wished, positive, negative, or zero, with no upper or lower bounds on their forecasts.

Participant pool and lab implementation

Experiments were conducted online over Zoom with 252 undergraduate students from Simon Fraser University and Texas A&M University from May to July 2020 and from May to June 2021. Online sessions were necessary given health restrictions due to the pandemic and the closure of physical labs.[5] Participants were recruited using SONA and ORSEE [Greiner, 2015] recruiting systems. The sessions for each treatment were equally split between the two institutions. Each session lasted approximately two hours, during which instructional time was about 40 minutes and 10 minutes of four rounds of practice with the experimental interface.[6] Participants were able to ask questions to the experimenter throughout the session. Subjects were paid a show-up fee of $7 in addition to pay linked to their performance, with an average total pay of $25. Payments were made by e-Transfer in Canada and via Venmo in the United States.

[5] Petersen and Rholes [2022] show that, during the pandemic, forecasting behavior was not significantly different in basic LtFEs across online and lab settings.

[6] A link to a web-hosted PDF of the instructions was sent to each participant through zoom at the beginning of the session, allowing them to reference it at any point during the experiment. The PDF could not be downloaded and the URL to the instructions was changed after every session.

3 Experimental Hypotheses

Owing to the nature of the various policy mandates, the policy regimes are predicted to generate noticeably distinct aggregate dynamics. Figure 1 presents the rational expectation equilibrium solutions for inflation, output gap, and the nominal interest rate for considered monetary policy mandates associated with our pre-selected shock sequence. Note that while inflation deviates significantly more from the steady state under IT and DM than under NGDP and PLT, it is relatively more stable.

We formulate theoretical predictions about the stabilization performance of different monetary policy regimes based on the loss function in Equation 9. Our model is simulated with each monetary policy regime under rational expectations using the sequence of demand shocks implemented in the experiment. Then we compute the average total loss in each regime as a square root of total loss (Equation 9) divided by 50 periods. The results are presented in Table 1. We break down the total loss into the losses associated with deviations of inflation, output, and interest rates from the steady state (which in our case is the target) by policy mandate.[7]

[7] Details about the breakdown of losses are in Table C1 in Online Appendix C.

Figure 1: Simulations with rational expectations

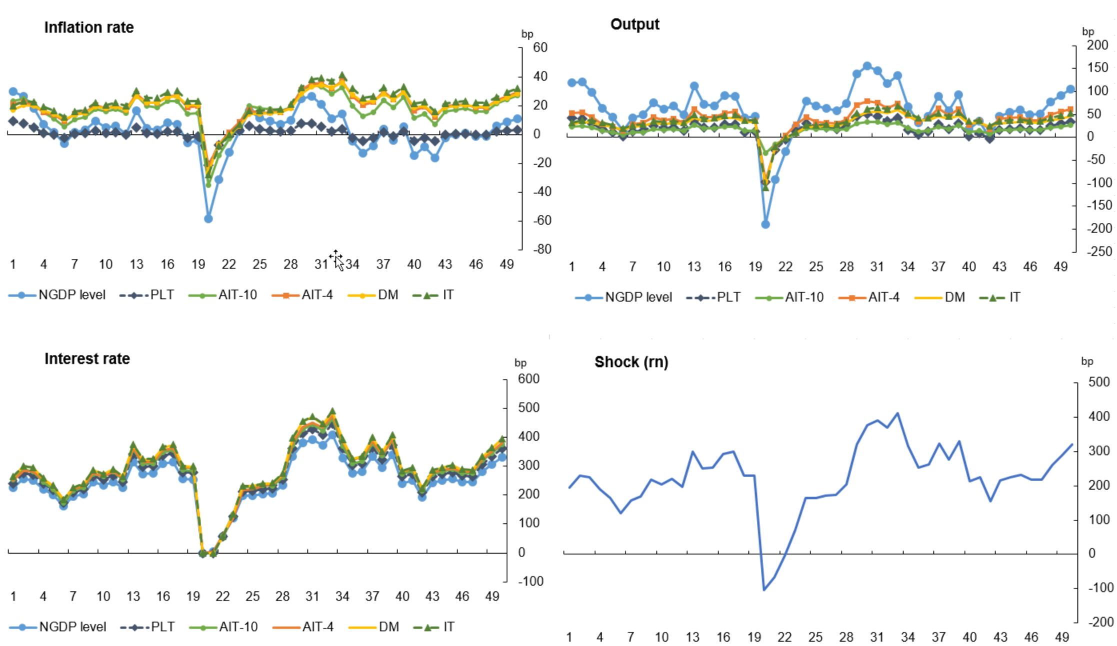


Table 1: Losses in REE

regime	periods 1-50	periods 1-19	periods 20-50
NGDP	168.2	153.8	176.4
PLT/PLT Comm	169.8	155.9	177.7
AIT-10	179.7	164.0	188.7
AIT-4	180.8	165.3	189.7
DM	184.4	168.5	193.5
IT	186.9	170.8	196.2

Given our simulated sequence of shocks, the overall total loss (as well as the loss associated with inflation) is predicted to be lowest under NGDP, followed closely by PLT. Thereafter, AIT with a 10-period horizon performs better than AIT with a 4-period horizon. DM and IT are predicted to produce relatively larger losses than the other regimes. It should be noted that losses across these regimes are quite close in REE. Using these simulations and calculated losses, we form our key testable hypothesis:
Hypothesis 1: The realized losses under the five mandates are ordered as follows $L_{NGDP} < L_{PLT} < L_{AIT-10} < L_{AIT-4} < L_{DM} < L_{IT}$.

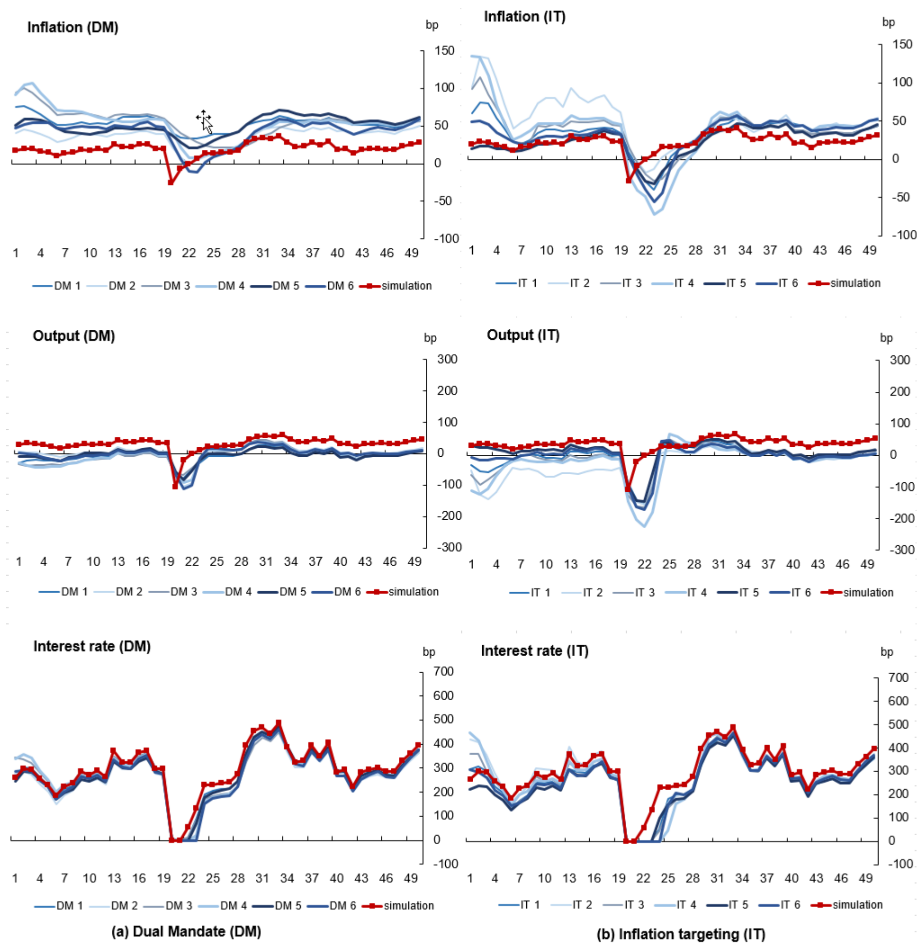
In Online Appendix C, we extend our computational analysis with varying shares of naive agents. Our simulations show that IT and DM are more robust to the presence of non-rational expectations in their ability to stabilize the economy and perform better than history-dependent regimes. Naive expectations do not have a forward-looking aspect and, as a result, weaken the expectations channel on which history-dependent regimes rely for their superior performance in models with rational expectations.

4 Results

4.1 Overview of aggregate dynamics

Figure 2 illustrates the dynamics of inflation, output, and interest rate in each of six sessions in treatments with DM and IT. The dynamics of inflation and output exhibit impressive consistency across six sessions for each of these regimes and are very similar to the simulations of our model in REE. Both DM and IT experience stable inflation and output and a brief episode at the ELB at the time of large demand shock. These economies recover relatively quickly from this shock, although somewhat more slowly than in the simulation

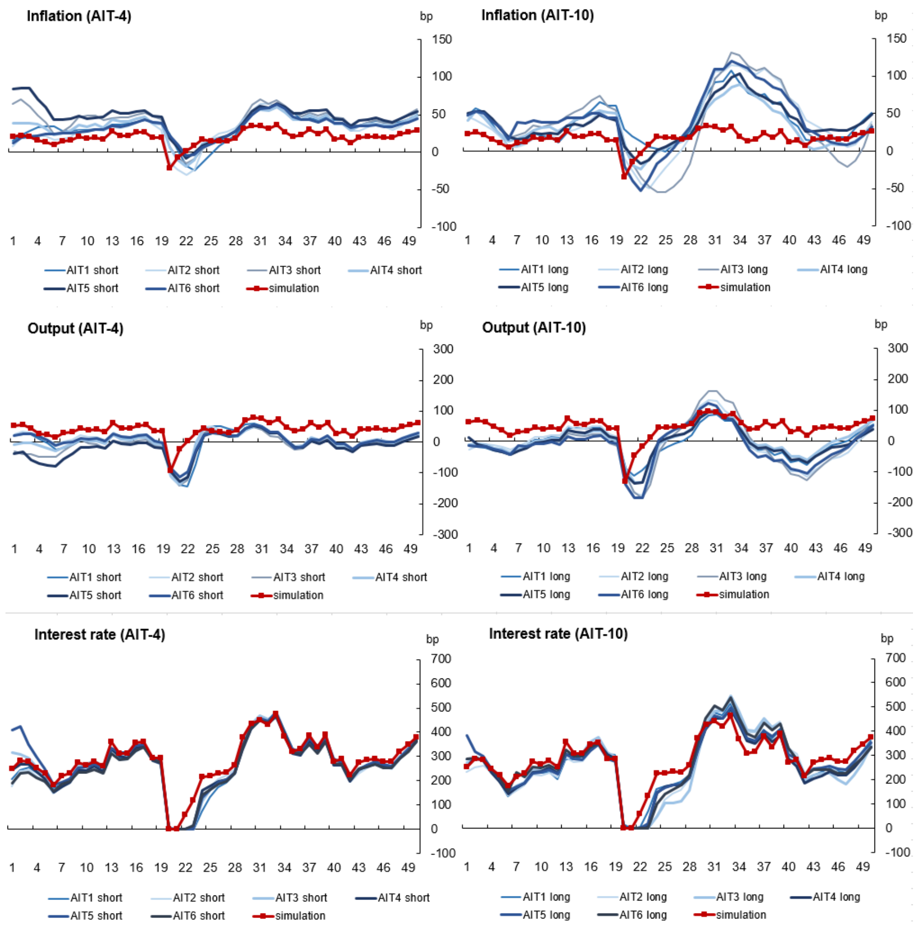
Figure 2: Aggregate dynamics of inflation, output and interest rate in dual mandate (DM) and inflation targeting (IT) treatments



with RE. The experimental economies take 3-4 (4-5) periods in DM (IT) to lift off from the ELB compared to 2 periods under RE.

The stability of IT and DM in our experiments may be due to the relatively high responsiveness of policy to both output and inflation. The coefficients in our DM are the strongest considered in the literature. For example, Cornand and M'baye [2018] and Hommes et al. [2019b] study flexible IT mandates (interest rate responding to both inflation and the output gap) with a relatively small coefficient on output gap ($\phi_x = 0.5$), while Kryvtsov and Petersen [2013] consider coefficients of as high as $\phi_x = 1$. Our simulations with naive agents show that IT and DM are more robust than other regimes. Hommes et al. [2019b] show that in the presence of backward-looking expectations, a strong response to output is important for stabilizing output and inflation.

Figure 3: Aggregate dynamics of inflation, output, and interest rate in average inflation targeting (AIT) treatment



Aggregate dynamics from AIT-4 and AIT-10 are presented in Figure 3. The dynamics of inflation and output indicate that AIT-4 is capable of stabilizing the economy similarly to IT and DM, whereas AIT-10 delivers less stability and reports less consistency across sessions than in IT, DM, and AIT-4. The liftoff from the ELB takes four periods in both AIT treatments, compared to two periods under RE. This result is consistent with Amano et al. [2020], who find that in a two-agent New Keynesian model with a fraction of backward-looking price setters, a shorter horizon is optimal in AIT, and in experimental results by Salle [2021].

Recent evidence from a household survey by Coibion et al. [Forthcoming] indicates that people have difficulty understanding the AIT framework and do not react to the treatment with information about average inflation target differently than to information about inflation target. Our evidence shows that AIT performs relatively well, suggesting that experimental participants understood AIT reasonably well, especially AIT of a shorter horizon. Indeed, the outcomes in AIT-4 treatment are comparable to those in IT, consistent with Coibion et al. [Forthcoming], showing that people react similarly to AIT as they do to IT.

Finally, we present results from the level-targeting frameworks. Figure 4 presents results from the NGDP level-targeting treatment, and Figure 5 shows results from the PLT treatment. These figures present time series separately for the stable periods of experiment and the remaining periods following ELB shock.

During the stable periods, both NGDP and PLT show dynamics with larger amplitude of fluctuations and more persistence than in the RE simulation. Before the ELB shock, treatments with NGDP and PLT show more volatility and less consistency in dynamics than all other treatments. Following the ELB shock, all sessions in each of these regimes unravel into spiraling deflation and declining output. Only one session in each of NGDP and

Figure 4: Aggregate dynamics of inflation, output, and interest rate in NGDP level targeting (NGDP) treatment

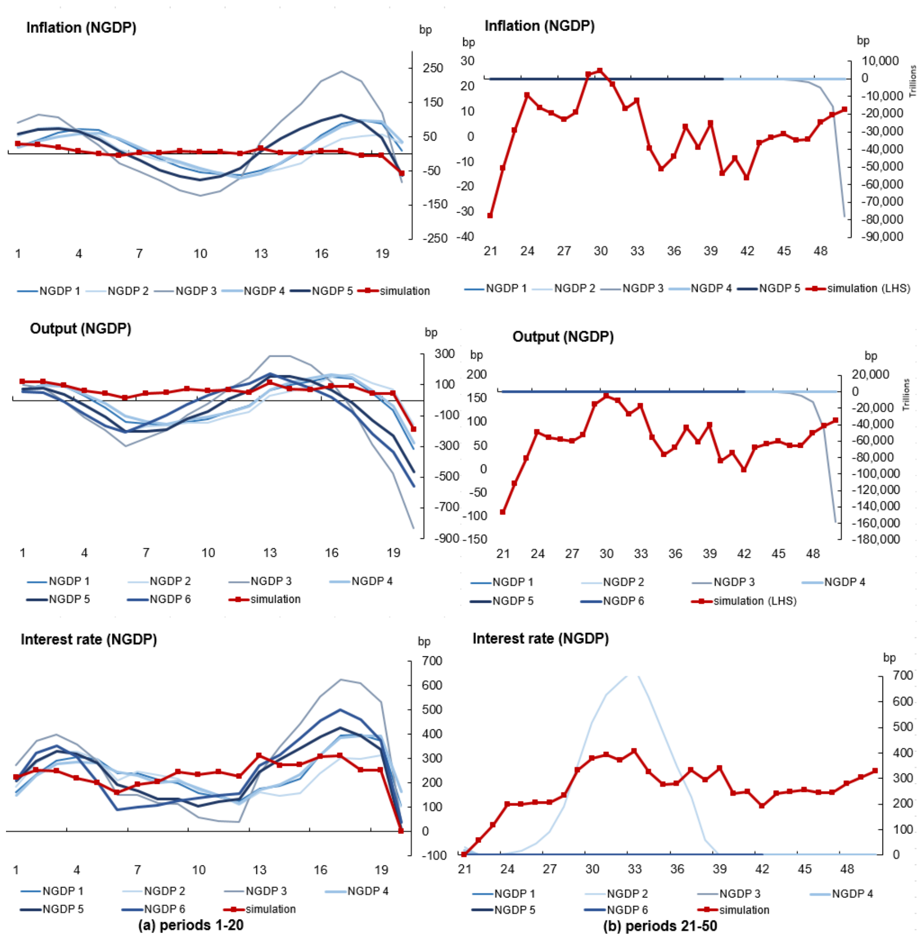
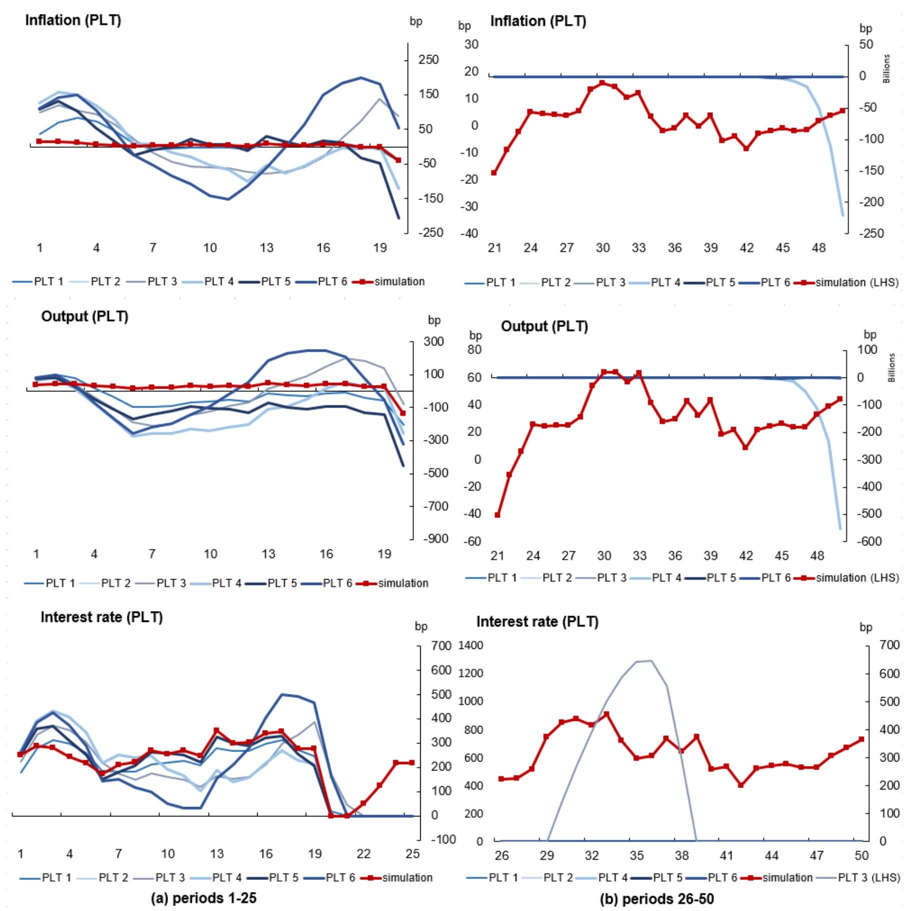


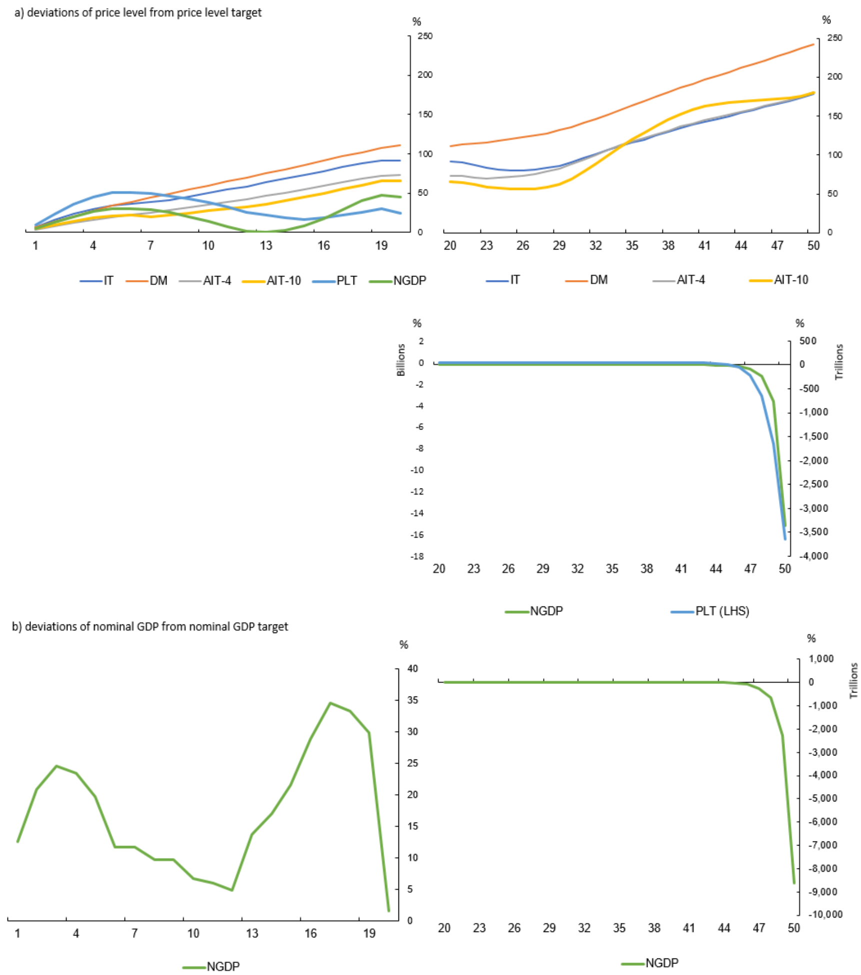
Figure 5: Aggregate dynamics of inflation, output, and interest rate in price-level targeting (PLT) treatment



PLT experience some liftoff from the ELB, but eventually return back in a subsequent, expectations-driven, recession. Such unraveling deflationary dynamics were not observed on other policy regimes in our experiments. Other experimental studies such as Arifovic and Petersen [2017], Hommes et al. [2019a], and Assenza et al. [2019] have reported evidence of deflationary spirals in the experimental economies that face ELB. Unraveling in these treatments may be a consequence of participants' panic as they observe their economies increasingly deviating from the central bank's announced targets.

When it comes to stabilizing the price level, NGDP and PLT perform the best during the the pre-shock phase of the experiment (Panel (a) Figure 6). The level-targeting regimes are followed by AIT-10 and AIT-4, as these regimes incorporate some history-dependence and are closer in their design to PLT than IT and DM. Interestingly, during the initial 10 periods of the experiment, AIT-10 and AIT-4 are better at keeping the price level close to the target than PLT. This happens as it likely takes a bit longer for learning to happen in PLT and

Figure 6: Deviations of price level from price-level target, percent



for this regime to come into its full force. And finally, IT and DM result in relatively higher deviations of price level than other regimes by the end of the pre-shock phase. However, following the ELB shock, PLT loses its performance properties and becomes incapable of stabilizing the price level, as this regime unravels into deflationary spirals. NGDP performs even worse than PLT. IT, AIT-10, and AIT-4 are comparable in their ability to stabilize the price level. DM finishes after IT, AIT-10, and AIT-4, as this regime places relatively less weight on stabilizing inflation.

Panel (b) of Figure 6 illustrates deviations of nominal output from its target level in NGDP level targeting. This figure shows that NGDP can be effective in achieving nominal output target during the stable period, however, this performance is lost following the ELB shock: deviations of nominal output from target unravel as experimental economies with NGDP regime unravel into very strong deflationary spirals.

4.2 The horse race

4.2.1 Macroeconomic stability of different regimes

Next, we summarize the performance of the six monetary policy regimes in terms of their ability to stabilize inflation, output, and interest rate. We use the loss function (Equation 9) as a summary measure describing the stabilization performance of each regime. Table 2 presents mean losses by treatment and phase, ordered from lowest to highest.

Table 2: Losses in laboratory experiments

Regime	Periods 1-50	Periods 1-19	Periods 20-50
AIT-4	172	154	182
DM	176	168	181
IT	177	170	181
AIT-10	186	152	203
PLT	31×10^9	213	39×10^9
NGDP	4×10^{15}	221	5×10^{15}
PLT Comm	2723 (155*)	131	3435 (169*)

This table presents losses averaged across all sessions of each treatment. * Values in brackets excludes single outlier session.

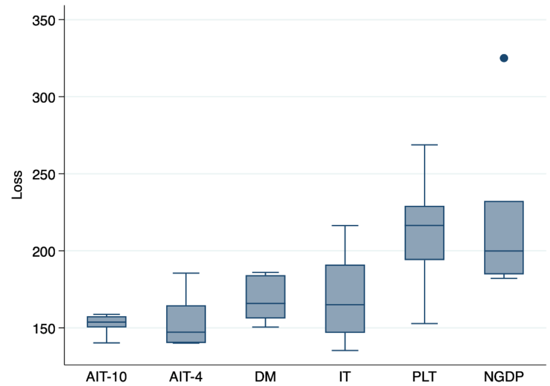
The session-level distribution of average losses for each monetary policy framework are presented by phase in Figure 7.[8] We observe a distinct ranking between rate and level targeting regimes. This ranking is somewhat different before the ELB shock and after it. During the stable periods 1-19, AIT-4 and AIT-10 perform better than DM and IT, which are followed by NGDP and PLT. After the ELB shock, the performance of AIT-4 and AIT-10 deteriorates below that of DM and IT, which outperform NGDP and PLT. Overall, after the ELB shock, the rankings of the regimes decline as the degree of their history dependence increases.

We perform a Wilcoxon rank order test (details are in Table E1) to assess whether the distribution of losses in each treatment has statistically identical medians. This table shows that losses in DM are statistically significantly different from losses in AIT-10, AIT-4, NGDP, and PLT at 1% to 5% levels, and losses in AIT-4 and AIT-10 are statistically significantly different from losses in NGDP and PLT at 1% to 5% levels. The difference between losses in the rate targeting treatments, and between the level targeting treatments are not statis-

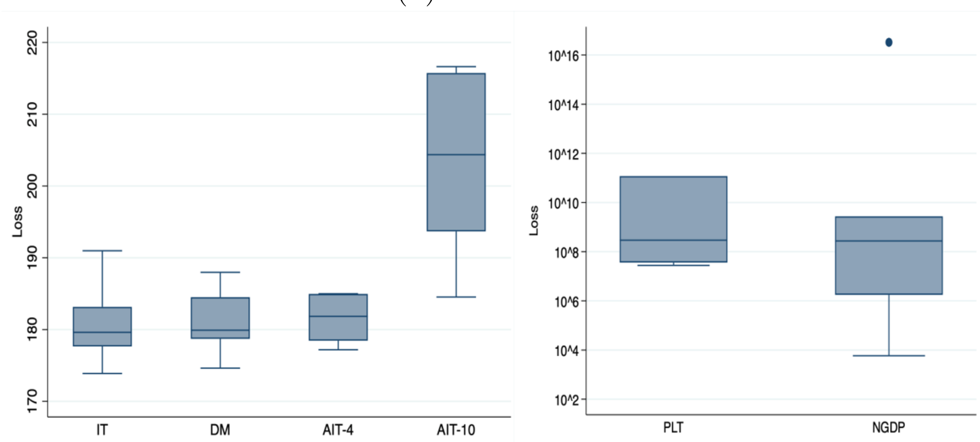
[8] The y-axis in Panel (b) of Figure 7 for PLT and NGDP is a logarithmic scale. Dots in the figures represent values outside upper adjacent value (upper quartile +3/2 Interquartile range).

Figure 7: Distribution of session losses, by phase

(a) Pre-shock



(b) Post-shock



tically significant.

The ranking of the regimes in the experiments is different from the ranking of the regimes in the model with REE (Table 1).[9] Based on the evidence from experiments, we reject our hypothesis about the relative stabilization performance of the six monetary policy frameworks outlined in Section 3. In the experiments, monetary regimes responding to concurrent inflation and output such as IT, DM, AIT-4, and AIT-10 outperform the most history-dependent regimes, PLT and NGDP. AIT-4 outperforms AIT-10, i.e. less history dependence results in better stabilization.

Overall, the performance of the regimes declines with an increase in the extent of history dependence. Better performance of rate-targeting regimes IT, DM, AIT-4, and AIT-10 rel-

[9] A summary of the rankings of the regimes in the simulations with RE, naive expectations and in the experimental data is provided in Table C2 in Online Appendix C.

ative to level-targeting rules PLT and NGDP may be due to the "divine coincidence" in the model: in the presence of only demand shocks, a strong response of policy rule to both output and inflation achieves stabilization of both. In our parameterization, both IT and DM regimes benefit from responding relatively strongly to deviations to both inflation and output.

While our findings of the performance of monetary policy regimes contrast with those in the model with RE, they are consistent with results from the horse race in ToTEM with an estimated positive share of non-rational rule-of-thumb price and wage setters in the Canadian economy [Swarbrick and Zhang, forthcoming].[10] Wagner et al. [forthcoming] confirm the ToTEM results regarding underperformance of highly history-dependent frameworks in a model with bounded rationality, or "cognitive discounting" [Gabaix, 2020]. The presence of boundedly rational rule-of-thumb price-setters weakens the performance of the history-dependent rule because it weakens the expectations channel. For example, in the PLT regime, missing the price-level target in the current period means that, in the future, inflation must be higher to catch up with the target price level. Rational agents understand this mechanism and increase their inflation expectations, leading to higher inflation and output. However, the expectations channel does not work well with boundedly rational expectations. If inflation expectations are backward-looking, agents do not raise their inflation expectations in response to missed price-level targets, and consequently the price level continues to fall short of its target without some further inflationary pressures. This is especially important at the ELB, where monetary policy cannot rely on lowering the interest rate to stimulate demand, and the expectations channel is key to managing real interest-rate expectations. As we will show in Section 5.1.1, the experimental participants use forecasting rules that are mostly backward looking/trend chasing, i.e. not rational.

4.2.2 Is AIT good enough? The role of framing in policy regimes

Theoretically, under RE, the performance of AIT approaches that of PLT as the horizon of AIT increases to infinity. But our experimental results show that both AIT-4 and AIT-10 perform better than PLT. Overall, performance declines with an increase in the extent of history dependence with the most history-dependent rule, PLT, performing the worst. This is likely because of the framing of these regimes: AIT is framed in terms of the inflation rate, while PLT is framed in terms of the price level. AIT is less cognitively demanding when participants are tasked with forecasting inflation.

[10] Rule-of-thumb price setters set prices either based on past inflation or inflation target not in the optimal way; for more details, see Dorich et al. [2013].

5 Why do history-dependent regimes not work better?

History-dependent regimes can be more difficult than rate-targeting regimes for people to understand. The forecasting task in the AIT, PLT, and NGDP regimes is more computationally and cognitively demanding than in IT and DM. In AIT, participants need to keep track of average inflation over a certain number of periods and to figure out what inflation rate the policy rule aims for in the next period to make up for past misses. In PLT, participants need to track deviations from the target price level and figure out what deviations of the price level from the target imply for the inflation rate, and how the policy rule would achieve this target in future periods. Similarly, NGDP participants need to interpret deviations of nominal output from its target to understand goals of policy rule for the next period's inflation and output.

In all history-dependent regimes, the goal for the next period's inflation (and output) evolves depending on the realized history. This is the key strength of such regimes in theoretical models with rational forward-looking expectations. However, history-dependent objectives can be a weakness working against these regimes in an experimental environment for several reasons. First, it is more challenging for participants to understand a time-varying policy goal that does not possess the clarity and stability of unchanging inflation targets as in IT and DM. The clearly communicated, constant target in IT and DM can serve as a salient focal point for anchoring expectations. Second, time-varying goals can erode the credibility of the regime by creating the impression of unfocused or wavering intentions. Lastly, these issues can be further exacerbated when the central bank fails to achieve these goals, with forecasters simply unable to "see it to believe it". Arifovic and Petersen [2017] find that the implementation of PLT by communicating time-varying targets for the inflation rate does not deliver the expected stability of this regime. Even when the subjects do not need to figure out the time-varying inflation target and are provided with this information by the experimenters, the PLT regime can be unstable.

In this section, we show that the weak performance of history-dependent regimes in the experiments is due to a combination of participants having difficulty understanding these regimes ("don't get it") and the central bank having difficulty establishing their credibility ("don't buy it"). Limited comprehension of the regimes manifests itself in two ways: not enough participants forecast in the correct direction and, of those that do, forecasts fall short of what is rational and necessary to pull economies out of their deflationary spiral ("too little"). Even those who do try to forecast in the correction do so "too late."

5.1 Do people understand the monetary policy regimes?

The expectations channel is key to effective monetary policy and especially to the performance of history-dependent regimes. We begin by presenting evidence about participants' forecasting heuristics and how this may differ across competing regimes.

5.1.1 Forecasting heuristics

Our simulations discussed in Section 3 suggest that the nature of expectation formation – rational versus naive – can crucially affect the relative performance of different monetary policy mandates. It is typically assumed that how aggregate expectations are formed is policy invariant. However, Assenza et al. [2019] show that expectations of experimental participants tend to self-organize on different forecasting rules depending on the monetary policy regime in place. We next consider how different monetary policy mandates influence participants' forecasting heuristics.

We analyze the distribution of forecasting heuristics observed in each of the policy treatments. We consider several types of expectation formation observed in surveys and used in the estimation of DSGE models [Milani, 2012]: ex-ante rational or model consistent expectations [Muth, 1961, Sargent and Wallace, 1975], cognitive discounting [Gabaix, 2020], constant gain learning [Branch and Evans, 2006], anchoring on targets [Coibion et al., 2018], and extrapolative trend-chasing [Frankel and Froot, 1990]. We assign a type to each participant that best fits their forecasting behaviour. Interested readers can find more details of our approach and results in Online Appendix D.

The most striking result is the rarity of the rational expectations type in the experimental data. Fewer than 5 percent of participants in any treatment can be classified as rational or *model consistent*, and in some cases, this share is close to zero. The consistent lack of rational expectations suggests that participants do not broadly appreciate how economic fundamentals and monetary policy influence aggregate dynamics. Participants' expectations show substantial deviations from model-consistent expectations (Figure E1 in Online Appendix E). These deviations, however, are smaller in the rate-targeting regimes IT, DM, AIT-4 and AIT-10 (under 50 bps) than in level-targeting regimes PLT and NGDP (between 60 and 100 bps pre-shock, and exploding post-shock). The fact that AIT-10 delivers macroeconomic stabilization better than PLT because it is easier for subjects to understand, as a result,enables them to make more accurate forecasts.

Backward-looking expectations – trend-chasing and constant gain learning – are the dominant forecasting heuristics in most of our treatments. Together, these backward-looking heuristics describe more than 90 percent of participants in each treatment. In inflation forecasts, the largest share of participants uses trend-chasing during both pre-shock and post-shock phases, with trend-chasing becoming more prevalent post-shock with a decline of constant gain learning. We observe similar composition and evolution of heuristics in output forecasts across all treatments.

How strong is trend chasing?

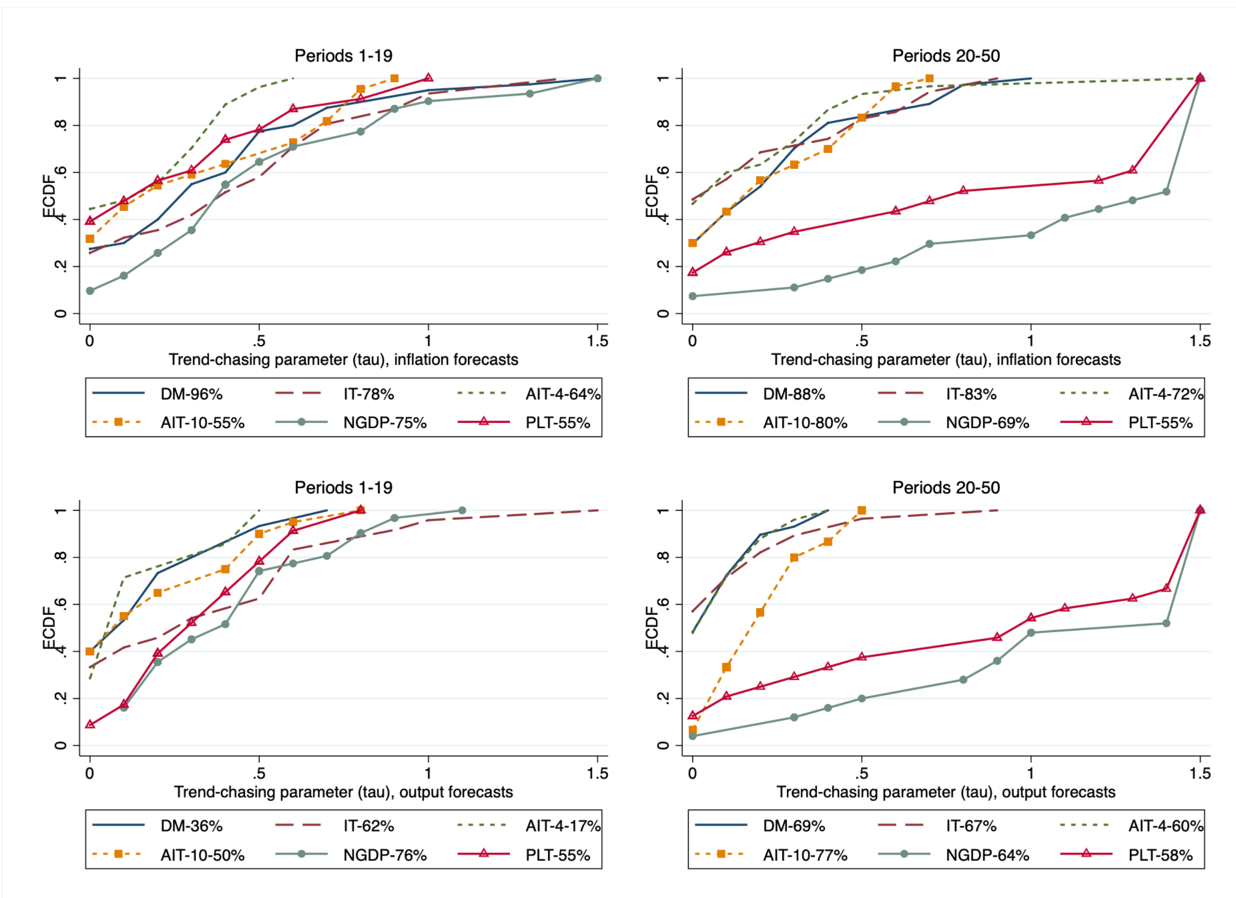
Most of the participants, irrespective of the regime, form backward-looking trend-chasing expectations. However, some of the regimes remain more stable than others, despite the marked absence of rational expectations in all of them. To better understand this observation, we explore how the degree of backward-looking expectations differs across the regimes. We compare the empirical cumulative distribution function (CDF) of the trend-chasing parameter τ assigned to trend-chasing forecasters across treatments. Figure 8 plots these distributions for inflation and output forecasts pre-shock and post-shock.

Pre-shock, there is relatively little difference across treatments in the distribution of the strength of the response to past trends. The median subject has an assigned τ parameter between 0.1 to 0.4, depending on the treatment. NGDP and PLT exhibit higher trend-chasing reactions compared with other treatments, but the differences aren't quantitatively large. The strength of trend-chasing is relatively low in AIT-4, in addition to a low incidence of trend-chasers in this regime.

By contrast, in the post-shock phase, we observe notable differences in how participants extrapolate trends across treatments. Subjects in rate-targeting treatments are significantly less responsive to past trends in inflation and output than those in level-targeting regimes. In the rate-targeting treatments, trend-chasing is characterized by a median value of a trend-chasing parameter of roughly zero, indicative of simple naive forecasting. In PLT and NGDP, trend-chasing is very strong, with parameter τ close to or greater than 1. In the monetary policy regimes with more history dependence, participants react more to history following a large shock. It may also be difficult for participants to understand what these monetary policy regimes should do in such cases. Instead of correcting past misses by expecting a reversal in inflation and output, participants react more strongly to recent trends to catch up with economic dynamics. Strong trend-chasing forecasts further destabilize the economy and lead to explosive deflationary dynamics in PLT and NGDP regimes following the ELB shock.

What are the implications of these strongly trend-chasing expectations for the stability of PLT? Honkapohja and Mitra [2014] suggest that PLT parameters of $\phi_P = 0.25$ and $\phi_x = 1$ would be sufficient to stabilize the economy if agents form their expectations through recursive learning. By contrast, Hommes and Makarewicz [2021] show that a non-linear New Keynesian system is stable when PLT responds very strongly to price deviations when agents are naive ($\phi_P = 3$ and $\phi_x = 2$). Our results suggest that forecasters are mostly trend-chasing, with a relatively small share of naive expectations. The explosive dynamics in our experiments with PLT suggest that PLT would need to be even stronger than in the case of naive or constant-gain expectations to effectively manage expectations.

Figure 8: Distribution of trend-chasing parameter τ in inflation and output forecasts by phase



This figure presents CDFs of parameter τ for participants whose forecasts were classified as trend-chasing. Percentages in the legend indicate the proportion of forecasts classified as trend-chasing for the corresponding monetary policy regime during the phase depicted in the chart.

5.1.2 Forecast dispersion

We next evaluate how each mandate reduces the disagreement across forecasters. We use the session-level interquartile range of inflation and output gap forecasts as our measure of disagreement. Average dispersion of inflation and output forecasts for each treatment is presented in Figure 9.

Figure 9: Dispersion of inflation and output forecasts

This figure presents the dispersion of inflation and output forecasts as measured by interquartile range, averaged for each period across all sessions of each treatment.

Disagreement is notably high at the beginning of each session in all regimes. This is to be expected as participants are just beginning to understand how to forecast and have no historical data to learn from. There is no consistent pattern in relative dispersion in the first few periods of play. By period 4, however, we observe a distinct separation between treatments. The interquartile ranges of forecasts in the level-targeting regimes are much higher than those in the rate-targeting regimes. Difficulty in understanding level-targeting regimes gives rise to greater heterogeneity in forecasts in PLT and NGDP.

On impact of the ELB shock, dispersion increased in all regimes (Figure 9). This finding is consistent with observations in surveyed expectations following the COVID-19 outbreak: consumer surveys have reported a sharp increase in divergence in consumers' inflation forecasts [CSC, 2020]. Interestingly, during the post-shock period, the dispersion of forecasts declines in all rate-targeting regimes (except output forecast in AIT-10). Dispersion grows larger in PLT and NGDP post-shock as participants become uncertain about what values to forecast. A lack of understanding of level-targeting regimes gives rise to greater heterogeneity among participants.

Bias in inflation and output forecasts

Our analysis shows a positive bias in inflation expectations and negative bias in output gap expectations persists in most of our treatments (Table E3 and Figure E1 in Online Appendix E). Inflation forecasts are above rational in all rate-targeting regimes IT, DM, AIT-4, and AIT-10, whereas output forecasts are below rational except in AIT-10. This suggests that participants tend to expect higher inflation and weaker output. This may be due to forecasters perceiving inflation as "bad for the economy," i.e. higher inflation is associated with weaker economic activity. Evidence of such views has been documented in survey data by Candia et al. [2020] among consumers across several countries, by Kamdar [2019] among consumers in the U.S., and by Coibion et al. [2020] among firms in Italy.

The bias in the experiments may be a consequence of participants not fully appreciating the stabilizing effects of monetary policy on inflation. While participants are aware of how monetary policy directly affects output and how output contemporaneously affects inflation, the transmission of monetary policy to inflation may be less obvious. This limited ability of forecasters to connect the causal effects of monetary policy on inflation has also been documented in Mokhtarzadeh and Petersen [2020] and Kryvtsov and Petersen [2021].

5.1.3 Do people forecast in the direction of rational expectations?

While participants' expectations are substantially different from rational, to what extent do they form their forecasts in the direction of rational model-consistent expectations? We begin by considering a generous definition of *basic rationality* whereby participants forecast in the direction of the REE solution. We base this evaluation on the approach in Amano et al. [2011] who call it "directionality." Specifically, a forecast is considered in the direction of the REE solution if a participant forecasts higher (lower) values relative to the previous period's outcomes if the REE forecast is higher (lower) than the previous period's outcome.[11] The shares of inflation and output forecasts exhibiting basic rationality in our experiments are presented for each treatment in Table 3.[12] The columns labelled *Both* report the share of both inflation and output forecasts made in the direction of the REE solution.

Table 3: Share of forecasts exhibiting basic rationality

	Preshock (Periods 1-19)			Shock (Periods 20-21)			Postshock (Periods 22-50)		
	Inflation	Output	Both	Inflation	Output	Both	Inflation	Output	Both
NGDP	0.59	0.49	0.26	0.85	0.22	0.12	0.29	0.29	0.18
PLT	0.49	0.64	0.29	0.71	0.58	0.49	0.36	0.38	0.26
DM	0.47	0.49	0.25	0.63	0.56	0.27	0.36	0.57	0.16
IT	0.54	0.54	0.33	0.70	0.49	0.26	0.46	0.54	0.21
AIT-4	0.48	0.62	0.32	0.58	0.52	0.20	0.43	0.56	0.25
AIT-10	0.48	0.57	0.33	0.63	0.80	0.51	0.51	0.65	0.36
PLT Comm	0.71	0.74	0.62	0.92	0.85	0.82	0.66	0.62	0.50

This table presents the share of forecasts in the direction of the rational expectations equilibrium solution. A forecast exhibits *basic rationality* if it is higher (lower) than the previously realized outcome when the predicted REE is above (below) the previous outcome. *Both* indicates the share of the participants simultaneously submitting inflation and output forecasts both exhibiting basic rationality.

During the pre-shock phase, about 50–60% of inexperienced subjects exhibit basic rationality in forecasts of inflation or output, and about 30% of subjects demonstrate it for both forecasts. There is little difference in the prevalence of basic rationality across policy regimes. On impact of the shock, participants' basic rationality sharply increases in periods 20-21 when it comes to forecasting inflation, as participants observe an unexpected large negative demand shock. Interestingly, at the time of the shock, PLT and AIT-10 report the highest shares of basic rationality in inflation and output forecasts, NGDP has the highest basic rationality (85%) in inflation forecasts followed by PLT and DM, and AIT-10 has the highest presence of basic rationality in output forecasts (80%). That is, large shocks increase

[11] In Amano et al. [2011], subjects are defined as forecasting in the correct direction if, in the regression analysis, subjects' forecasts are negatively related to past log deviation of the price level from its target.

[12] Time series are presented on Figure E2 in Online Appendix E.

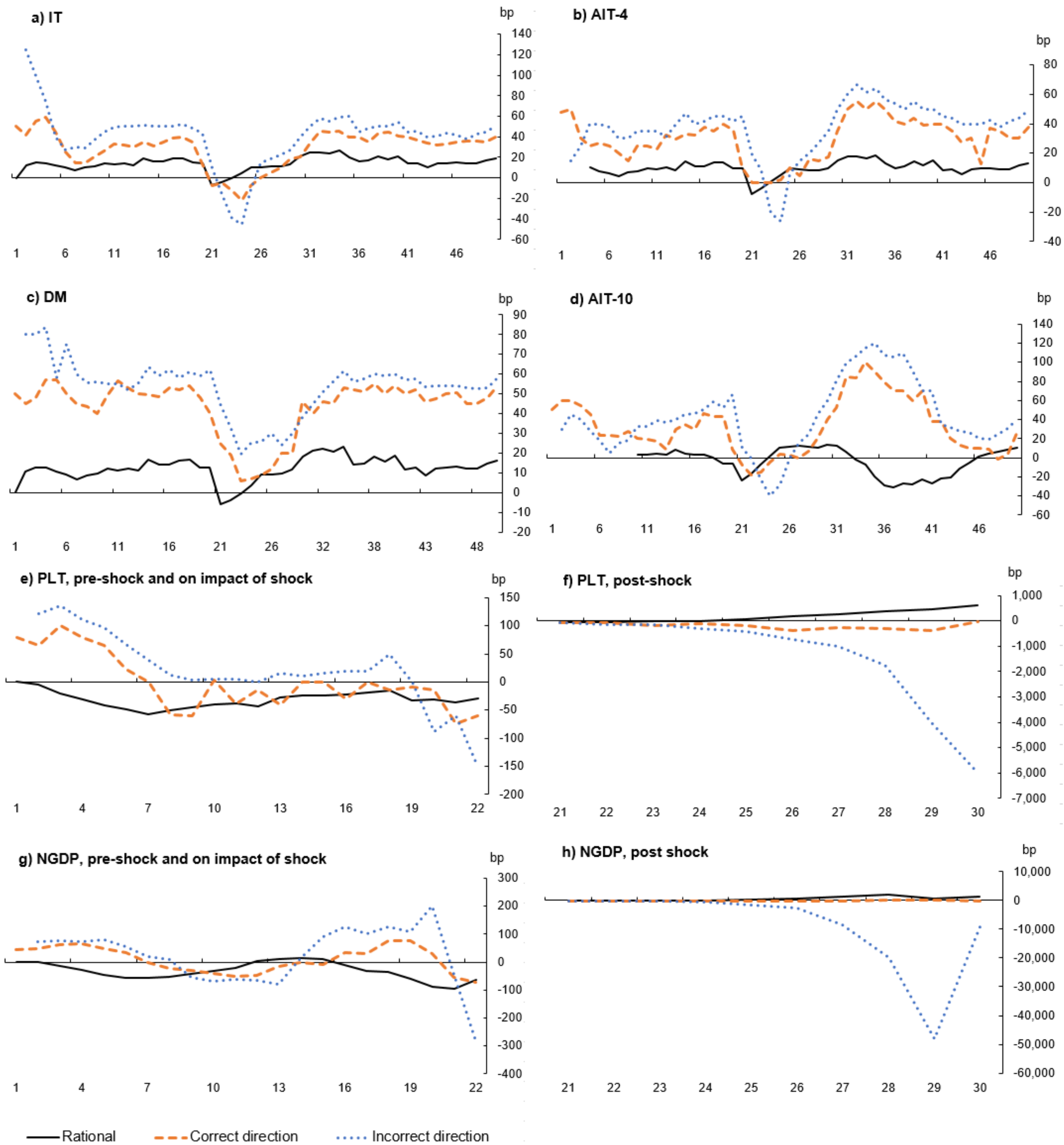
appear to reduce participants' inattention [Sims, 2010]. The timing in this spike of attention during a significant and unfamiliar shift to the ELB is consistent with theoretical work by Mackowiak and Wiederholt [2009] who show that increased uncertainty increases rationally-inattentive agents' attention to shocks.

Despite the sharp increase in the share of people forecasting in the correct direction in PLT and NGDP on the impact of the shock, these regimes start to destabilize following the shock. As the shock dissipates and fundamentals revert to the steady state, attention to the shock declines and the share of participants exhibiting basic rationality falls significantly in most treatments. The decline is most pronounced under the NGDP and PLT mandates for both inflation and output forecasts. Only 18% of NGDP participants and 26% of PLT participants forecast inflation and output in the correct direction in the post-shock phase. We attribute the relatively larger decline in basic rationality in these treatments to the higher cognitive load associated with processing these level-targeting regimes.

When participants forecast in the correct direction, do they go as far as REE? Or do their forecasts fall short? Figures 10 and 11 illustrate that the forecasts of those exhibiting basic rationality fall short of what is required by REE. Forecasts of those without basic rationality deviate from rational expectations even more because they do not even move in the necessary direction. For example, in PLT, when the price level is above target, rational expectations implies expecting lower inflation than in the previous period. Those with basic rationality forecast lower inflation than in the previous period, but their median forecast is not low enough to bring the price level back to target. And during the post-shock period, their forecasts do not incorporate fully the reversal prescribed by the PLT regime and therefore remain below rational value. Those without basic rationality do not even understand the reversal in inflation embedded in the PLT regime, and therefore, they do not lower their inflation forecasts below past-period inflation. As a result, forecasts of those with basic rationality are closer to the rational solution than forecasts of those without basic rationality, but they fall short of what would be expected under rational forecasts.

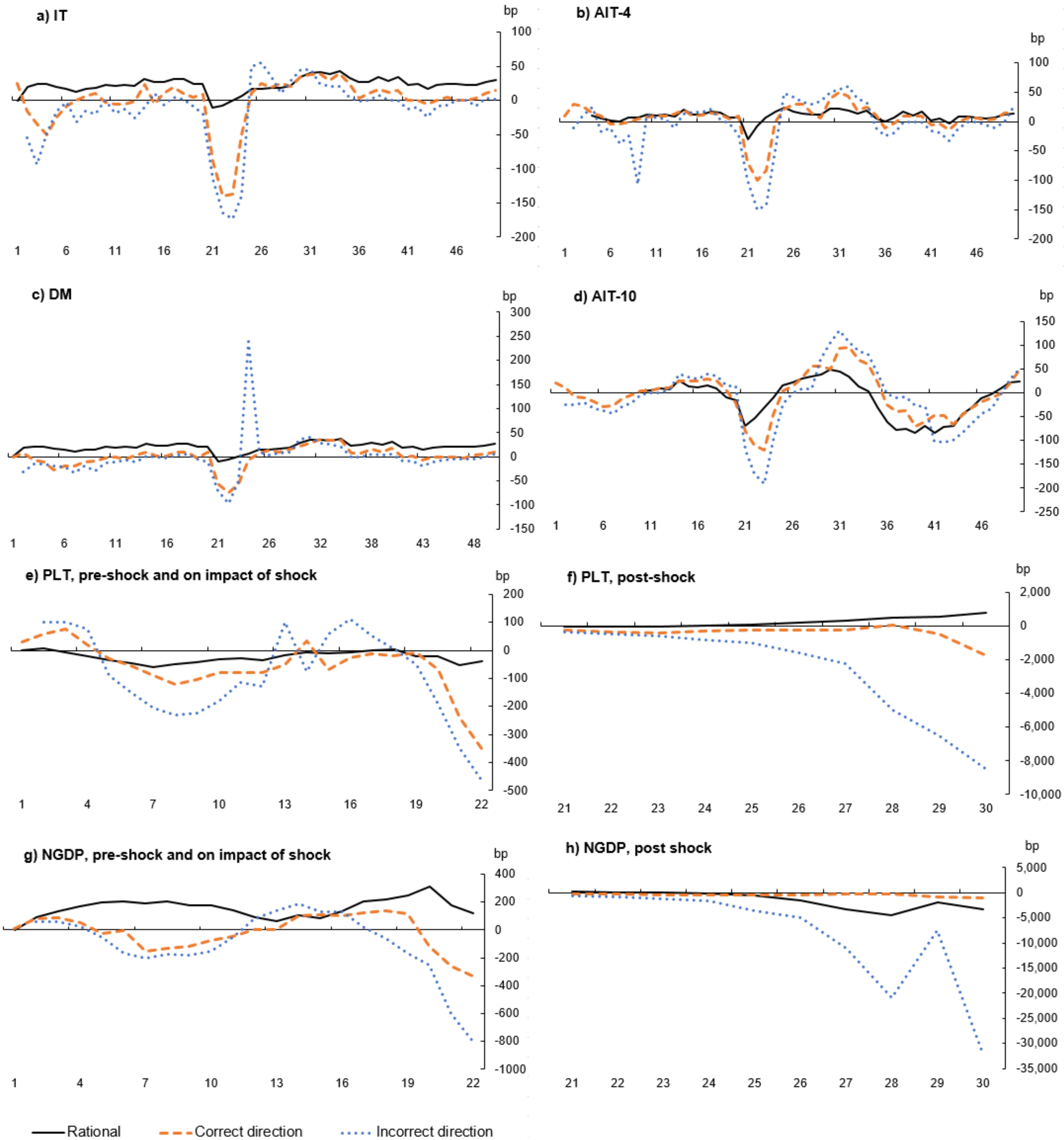
Our analysis reveals two behavioural challenges for history-dependent mandates. First, in all treatments, the fraction of participants forecasting in the correct direction is relatively low. Second, even those who forecast in the correct direction do not forecast sufficiently close to the rational solution. These behavioural challenges result in macroeconomic dynamics that deviate from the rational solution and are less stable. The consequences of these two issues result in explosive dynamics in the level-targeting regimes PLT and NGDP. While

Figure 10: Inflation forecasts by participants forecasting in the correct direction of rational forecast and incorrect direction



This figure presents the median forecast of participants of each type (correct and incorrect direction), averaged across all sessions of each treatment.

Figure 11: Output forecasts by participants forecasting in the correct direction of rational forecast and incorrect direction



This figure presents the median forecast of participants of each type (correct and incorrect direction), averaged across all sessions of each treatment.

rate-targeting regimes also suffer from a lack of basic rationality, the frameworks are more robust to heterogeneity and deviations from RE.

Our results about basic rationality in PLT are similar to those in Amano et al. [2011], who find that, on average, participants forecast in the correct direction in PLT. We would like to highlight the important difference between our experimental designs. In Amano et al., the experimental economy is exogenously generated and is simulated using a model with rational expectations, i.e. the expectations of their subjects' expectations do not feed into actual outcomes. In contrast, our experimental economy is self-referential: expectations feed into actual outcomes, and subjects observe actual outcomes and learn from them. In our experiments, as discussed, subjects' expectations are overwhelmingly non-rational, with stronger trend-chasing expectations in PLT (and NGDP) than in rate-targeting regimes. Non-rational expectations generate dynamics that are significantly different from dynamics in the model with rational expectations. For example, economies in PLT and NGDP treatments unravel into deflationary spirals. As a result, the forecasting task of subjects in our experiments is considerably more difficult than in Amano et al., and it may not be surprising that a relatively small shares of participants forecast in the correct direction. As dynamics in the experimental economies deviate further from REE solutions, as in PLT, they lead participants' forecasts further from those consistent with rational expectations.

History-dependent policy mandates such as NGDP and PLT demand a high level of rationality to be effective at stabilizing the economy. Our results suggest there is an insufficient level of basic rationality, let alone full rationality, at both the individual and aggregate levels, to advocate enthusiastically for such level mandates.

5.1.4 Too little, too late: It does not work

Subjects react too little, too late in history-dependent regimes. On impact of the ELB shock in period $t = 20$, the share of basic rationality increases in inflation forecasts in PLT and NGDP and output forecasts in NGDP. Inflation forecasts become closer to REE (panel (a) of Figure E1 in Online Appendix E) , while output forecasts fall below REE forecast (panel (b) of Figure E1 in Online Appendix E) in all regimes. And while inflation and output forecasts recover towards REE in IT, DM, AIT-4, and AIT-10, they continue to unravel below REE in PLT and NGDP. Not enough people adjust their expectations in the direction of basic rationality. Quantitatively, expectations are not sufficiently adjusted upward towards REE to pull the economies out of the crisis. In other words, the expectations channel on which PLT and NGDP rely is *too little* or too weak. As a result, the ELB episode becomes much longer in PLT and NGDP than in other regimes. With time, as participants see their economy unravel into a deflationary spiral, basic rationality declines and deviations of forecasts from REE increase in PLT and NGDP.

5.1.5 Need to see it to believe it

Another challenge for history-dependent regimes is that participants often *"need to see it to believe it."* In the pre-shock phase, the deviations of inflation from the price level and NGDP level targets are large. By the time the large negative demand shock occurs, there is little credibility in the central bank's ability to achieve its target. As destabilization continues, people do not see that policy is working and, therefore, they do not believe it as evidenced by the decline of the share of participants exhibiting basic rationality when the deviations from target grow.

We next evaluate how a central bank's performance in achieving its targets influences its credibility. A candidate proxy for a subject's credibility in the central bank is if the subject forms her inflation expectations in line with the rational expectations solution. We use the indicator variable $\mathbb{1}_{i,t}^{BasicRationality}$ that takes the value of one if participant i exhibits basic rationality in period t when forming their $t+1$ forecast, as described in the previous subsections, and zero otherwise. Recent central bank performance, $AbsDevFromTarget_{t-1}$, is measured as the absolute deviation of the key macroeconomic variable from the central bank's stated target. For IT and DM, we calculate the absolute deviation of inflation from the inflation target of zero. For AIT-4 and AIT-10, we compare the average inflation over the past and current four and 10 periods to zero. For PLT and NGDP, we compare the price level and nominal GDP level to their respective targets of 1000. Finally, we control for persistence in credibility by including a one-period lag of the indicator variable in our specifications. We estimate the following panel logit regressions by treatment over our pre- and post-shock data:

$$\mathbb{1}_{i,t}^{BasicRationality} = \alpha + \beta_1 \mathbb{1}_{i,t-1}^{BasicRationality} + \beta_2 AbsDevFromTarget_{t-1} + \beta_4 \mu_i + \epsilon_{i,t} \qquad (11)$$

where μ_i is the subject fixed effect and $\epsilon_{i,t}$ are robust standard errors. Results are presented in Table 4.

Table 4: Central bank credibility

Dep. Var. Preshock: Periods 1-19 $\mathbb{1}_{i,t}^{BasicRationality}$	IT	DM	AIT-4	AIT-10	PLT	NGDP	PLT Comm	
$\mathbb{1}_{i,t-1}^{BasicRationality}$	0.622	0.364	0.044	-0.512	0.491	0.492	0.757	0.816
	(0.16)	(0.16)	(0.22)	(0.26)	(0.16)	(0.16)	(0.19)	(0.20)
$\lvert\pi_{t-1} - \pi^*\rvert$	0.024	0.027						
	(0.00)	(0.01)						
$\lvert\frac{1}{4}\sum_{t-4}^{t-1}\pi_j - \pi^*\rvert$			0.088					
			(0.02)					
$\lvert\frac{1}{10}\sum_{t-4}^{t-1}\pi_j - \pi^*\rvert$				0.051				
				(0.01)				
$\lvert P_{t-1} - P^*\rvert$					0.000		0.003*	
					(0.00)		(0.00)	
$\lvert NGDP_{t-1} - NGDP^*\rvert$						-0.001		
						(0.00)		
$\lvert\pi_{t-1} - \pi_{t-2,t-1}^{Proj}\rvert$								0.038
								(0.01)
N	747	699	529	339	711	715	702	663
χ^2	51.13	15.62	34.78	29.53	9.567	13.99	16.94	37.82
p	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000
Dep. Var. Post-shock: Periods 20–50 $\mathbb{1}_{i,t}^{BasicRationality}$	IT	DM	AIT-4	AIT-10	PLT	NGDP	PLT Comm	
$\mathbb{1}_{i,t-1}^{BasicRationality}$	0.604	0.788	0.463	0.828	1.545	2.176	1.172	1.185
	(0.12)	(0.12)	(0.13)	(0.13)	(0.13)	(0.16)	(0.14)	(0.14)
$\lvert\pi_{t-1} - \pi^*\rvert$	0.007	0.021						
	(0.00)	(0.00)						
$\lvert\frac{1}{4}\sum_{t-4}^{t-1}\pi_j - \pi^*\rvert$			0.006					
			(0.00)					
$\lvert\frac{1}{10}\sum_{t-4}^{t-1}\pi_j - \pi^*\rvert$				0.001				
				(0.00)				
$\lvert P_{t-1} - P^*\rvert$					0.000		-0.000	
					(0.00)		(0.00)	
$\lvert NGDP_{t-1} - NGDP^*\rvert$						0.000		
						(0.00)		
$\lvert\pi_{t-1} - \pi_{t-2,t-1}^{Proj}\rvert$								-0.000
								(0.00)
N	1286	1287	1167	1185	1224	1044	1291	1287
χ^2	25.42	67.82	15.65	46.18	140.9	208.6	77.03	79.00
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

This table presents results from a series of fixed-effects logit panel regressions. The dependent variable is an indicator variable that takes the value of 1 if participant i in period t inflation exhibits basic rationality. α denotes the estimated constant. NGDP post-shock results are estimated and reported with a random-effects specification due to convergence issues. Standard errors are in the brackets.

Credibility in IT and DM strengthens as inflation deviate further from the central bank's target. A one-basis-point increase in the deviation of inflation from target increases the log odds of forecasting rationally by 0.04 when participants are inexperienced. The effect is smaller but remains significant in the post-shock phase. Likewise, we observe a similar re-coordination on RE when the economy deviates further from the target in AIT-4 and AIT-10 in the pre-shock phase. These results suggest that larger, more salient deviations from the target are quickly corrected by participants adjusting their expectations in line with the central bank's target.

This is not the case in PLT and NGDP. Participants do not consistently respond to deviations from the target by adjusting their direction of forecasting. In fact, for inexperienced NGDP participants, credibility declines as the NDGP level deviates further from its target. Given the lack of credibility in the pre-shock phase, it is not surprising that credibility is weak after entering the ELB and that monetary policy has no further capacity to provide stability.

We find that there is very strong persistence in subjects' basic rationality. If participants were previously forming expectations in line with rational expectations, and thus the central bank's mandate, they are more likely to continue to do so in most of our specifications. The persistence is strongest in IT, PLT, and NGDP, indicative of stable learning. AIT-4 and AIT-10 participants are somewhat different in the pre-shock phase, in that there is limited persistence in their basic rationality and, in AIT-10, are more likely to switch away from rational forecasting if they had used it in the past. However, in all treatments in the post-shock phase, participants' heuristics become highly entrenched and participants are much less likely to switch in and out of their basic rationality.

Both "don't get it" and "don't buy it" contribute to unstable dynamics in the PLT and NGDP treatments. Although participants form backward-looking, trend-chasing expectations in all regimes, PLT and NGDP give rise to the strongest extrapolation of past trends in the future. The trend-chasing expectations are more destabilizing in the level-targeting regimes than in other regimes.

6 Improving learning of level-targeting mandates with central bank communication

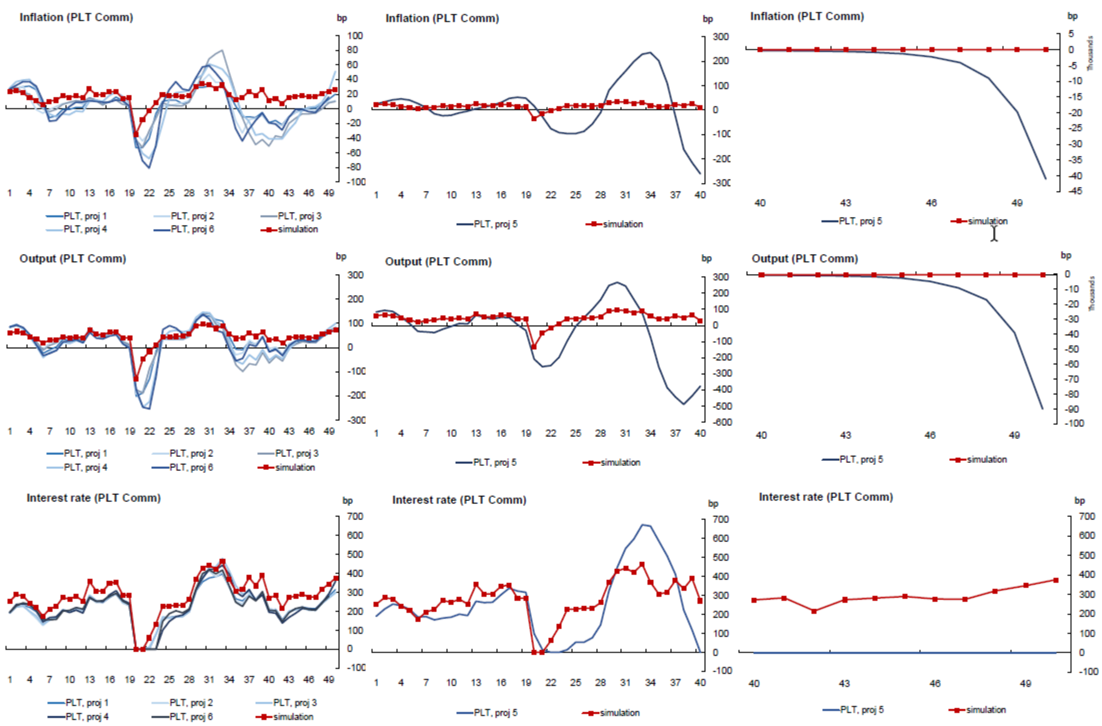
Our experiments demonstrate how challenging it is for people to forecast under level-targeting mandates. Many people fail to understand that the central bank must bring the price level back to target, and even more do not understand how much of a make-up strategy is necessary. One solution is to provide more central bank guidance about the implied path of inflation. Recent research has shown that precisely communicated point projections can effectively guide expectations in similar laboratory experiments [Mokhtarzadeh and Petersen, 2020, Rholes and Petersen, 2021, Petersen and Rholes, 2022].

In a final treatment, PLT Comm, we extend our PLT treatment by introducing central bank communication of precise projected paths of inflation and the output gap. Participants were informed whether the price level was above (below) target, that the central bank would respond by increasing (decreasing) the interest rate, and the impact higher (lower) rates would be expected to have on inflation. Precise point projections were presented to participant as green point forecasts extending beyond the inflation and output gap time series for the next five periods. Subjects were told in the instructions that the projections were constructed using the data-generating process, and in particular, the exogenous shocks and recent price level. Six sessions of PLT Comm were conducted in July 2022 using the same protocols described in Section 2.

Aggregate dynamics from PLT Comm are presented in Figure 12. Figures showing heuristics and comparison of losses to other regimes are available in Online Appendix F. In both phases of the experiment, the communicated projections results in highly stable inflation and output gap dynamics, significantly outperforming all treatments in terms of minimizing aggregate losses. Only one of the six sessions experience a significant deflationary episode at the ELB that does not recover. The session-level differences in aggregate losses between PLT Comm and all other treatments are highly significant pre-shock when we consider all sessions, and post-shock when we exclude the outlier session 5 (Wilcoxon rank sum pairwise test in each phase, $p < 0.05$).

The increased economic stability is due to a notable change in the distribution of forecasting heuristics. The share of PLT participants exhibiting trend-chasing heuristics in their inflation forecasts declines from 55% without communication (both phases) to 36% in the pre-shock phase and 43% in the post-shock phase when provided with projections. We also

Figure 12: Aggregate dynamics of inflation, output, and interest rate in PLT Comm treatment



observe an increase in ex-ante rational forecasting in the pre-shock phase from 2.2% without the communication to 7.1% with. Basic rationality is much higher as well. Table 3 shows that the majority of participants in all phases of the game are able to forecast both variables simultaneously in the correct direction. More than 60 percent of PLT Comm participants are able to form basically-rational expectations during the pre-shock phase, while it is less than one-third in all other treatments.

In Table 4 we present results from the PLT Comm treatment on central bank credibility. We evaluate the effects of recent deviations of price level from target, and the most recently observed inflation from the central bank's projected value, $|\pi_{t-1} - \pi^{Proj}_{t-2,t-1}|$. Results are presented in the last two columns. Credibility is considerably more persistent than in all other treatments. Like the rate-targeting treatments, we find that, pre-shock, PLT Comm participants maintain their confidence in the central bank when either the price level deviates more from target or inflation deviates more from the projected value. This confidence, however, is not sustained in the post-shock phase, where participants lose a small but significant amount of credibility in the CB when it fails to achieve its targets or projections.

7 Conclusion

In December 2021, Canada renewed its flexible inflation targeting framework for the next five years. The results of our experiments, together with model simulations and public consultation, informed this policy decision [ITR, 2021]. Contrary to the theoretical predictions with full information rational expectations, our results indicate that rate-targeting regimes such as IT, DM, and AIT outperform level-targeting regimes such as PLT and NGDP. More history dependence worsens macroeconomic stability, especially once economies experience some time at the ELB. While the Bank discussed the many potential benefits to adopting AIT and DM, neither was judged to provide an improvement over the current flexible inflation targeting strategy. In particular, inflation targeting was viewed as highly robust to more realistic, boundedly-rational expectations.

Policymakers cannot take for granted that people form rational expectations. Individual-level analysis of our experimental data suggests that people struggle with even the basic implications of monetary policy and do not fully internalize the goals of the policy regimes [Carvalho and Nechio, 2014, Kryvtsov and Petersen, 2021]. Only a relatively small share of participants form both expectations in the direction of the rational expectations solution ("basic rationality"). Forecasts of those with basic rationality quantitatively fall short of rational, i.e. they do not expect strong enough reversal in the dynamics. Furthermore, the share of the participants with basic rationality declines during the unstable post-shock period. In sum, these relatively rational forecasts are too few and not strong enough to reverse post-shock deflationary spirals.

Rather, the vast majority of our participants hold some form of backward-looking expectations. Specifically, most participants use a trend-chasing heuristic, with varying degrees of trend extrapolation. Moreover, the degree of their trend-extrapolation can increase significantly during lengthy episodes at the ELB in the level-targeting regimes, PLT and NGDP. This observed shifting of heuristics within and across policy regimes has been well-documented in the learning-to-forecast experimental literature [Assenza et al., 2019, Pfajfar and Žakelj, 2014] and reinforces the fact that expectations are *not policy invariant*. People shift their forecasting heuristics when their environment becomes more unstable and unpredictable, particularly when monetary policy is inactive and when rates reach their effective lower bounds [Arifovic and Petersen, 2017, Ahrens et al., 2022, Assenza et al., 2019, Hommes et al., 2019a, Kryvtsov and Petersen, 2021]. We find that rate-targeting regimes such as IT are more effective than level-targeting mandates at producing economic stability, and that in turn, leads to better anchored expectations and forward-looking behavior. This result is consistent with empirical evidence that people's personal experiences shape their inflation expectations [Malmendier and Nagel, 2016].

Our experimental results may be useful for the practical implementation of AIT. Recent evidence from household surveys regarding AIT are mixed. Hoffmann et al. [2022] show that German households are able to revise their expectations consistently with a hypothetical introduction of AIT. On the other hand, Coibion et al. [Forthcoming] show that U.S. consumers had difficulty understanding the newly introduced AIT framework of the Federal Reserve in August 2020 [Powell, Clarida, 2021]. The details of implementation of AIT were not clearly presented to both the public and surveyed participants and may contribute to the mixed evidence on the understanding of AIT. Orphanides and Williams [2005] illustrates that it is even more important to have credible inflation targets under bounded rationality than under rational expectations. However, a vague definition of policy goal may have some advantages in maintaining credibility when goals are not achieved [Stein, 1989, Jia and Wu, 2022].

We find that, even with full information about the central bank's horizon and the explicit specification of the AIT policy rule, people still have difficulty internalizing the make-up strategy of AIT-10. The solution to this issue is to design policy in which the make-up strategy is clear to the public and by responding to deviations over a shorter horizon such as IT, DM, AIT-4. Our finding that AIT with a shorter horizon performs better than AIT with a longer horizon is similar to that in Amano et al. [2020]. They find that AIT with a relatively short horizon is optimal in a two-agent New Keynesian model with a fraction of firms forming backward-looking expectations, and that in this case, the properties of the economy are quantitatively similar to those under a price-level target. The latter conclusion is, however, in contrast with our experimental results, where PLT is much more unstable than AIT of both short and long horizon. Our experimental evidence suggests that AIT of a shorter horizon is more likely to stabilize the economy than PLT.

History-dependent rules like PLT may be better implemented if supplemented with well-designed central bank communication. The promising results from our follow-up PLT Comm treatment indicate that communicating projections together with an explanation of the goals of monetary policy can guide expectations and achieve stabilization. Communication plays a very important role in central banking [Blinder et al., 2022, Haldane and McMahon, 2018]. Levin [2014] emphasizes the importance of clarity about the monetary transmission mechanism in the joint design of monetary policy strategy and communications. Ehrmann et al. [2022] show that communication about new inflation targets can have longer lasting effects on expectations if supplemented with background context, explaining why and how the target helps to stabilize the economy. Otherwise, on-going communication with the public is vital for long-term management of expectations. Our experimental design incorporates these principles and illustrates their success.

Our paper provides empirical support for recent theoretical work on macroeconomic dynamics under level-targeting mandates when agents form non-rational expectations. Honkapohja and Mitra [2014] and Honkapohja and Mitra [2020] show that properties of NGDP-level targeting and PLT do not hold when agents form expectations adaptively. If agents learn about forecasting inflation and output from historical data (as we find experimental participants do), then these history-dependent regimes do not perform better than IT. Both of these regimes require additional guidance about the path of targeted variables to improve their performance. Our results suggest that communication about the expected path of inflation and output gaps is sufficient to guide expectations under PLT. However, more research is necessary to determine which types of communication are most effective in level-targeting regimes.

References

Canadian Survey of Consumer Expectations, Q2 2020. URL https://www.bankofcanada.ca/2020/07/canadian-survey-of-consumer-expectations-second-quarter-of-2020.

Monetary Policy Framework Renewal — December 2021. *Bank of Canada*, 2021.

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

Steffen Ahrens, Joep Lustenhouwer, and Michele Tettamanzi. The stabilizing effects of publishing strategic central bank projections. *Macroeconomic Dynamics*, pages 1–43, 2022.

Robert Amano, Jim Engle-Warnick, and Malik Shukayev. Price-level targeting and inflation expectations: Experimental evidence. Technical report, Bank of Canada, 2011.

Robert Amano, Stefano Gnocchi, Silvain Leduc, and Joel Wagner. Average is good enough: Average-inflation targeting and the ELB. *Bank of Canada Staff Working Paper*, 2020-31, 2020.

Jasmina Arifovic and Luba Petersen. Stabilizing expectations at the zero lower bound: Experimental evidence. *Journal of Economic Dynamics and Control*, 82:21–43, 2017.

Tiziana Assenza, Peter Heemeijer, Cars H Hommes, and Domenico Massaro. Managing self-organization of expectations through monetary policy: A macro experiment. *Journal of Monetary Economics*, 2019.

Claes Berg and Lars Jonung. Pioneering price level targeting: The Swedish experience. *Journal of Monetary Economics*, 43:525–551, 1999.

Ben S. Bernanke. Monetary policy in a new era. *Paper prepared for conference on Rethinking Macroeconomic Policy, Peterson Institute, Washington, DC, October 12-13*, 2017.

Alan S Blinder, Michael Ehrmann, Jakob De Haan, and David-Jan Jansen. Central bank communication with the general public: Promise or false hope? 2022.

William A Branch and George W Evans. A simple recursive forecasting model. *Economics Letters*, 91(2):158–166, 2006.

James Bullard. A primer on price level targeting in the U.S. *Presentation, CFA Society of St. Louis, St. Louis, Mo.*, 2018.

Bernardo Candia, Olivier Coibion, and Yuriy Gorodnichenko. Communication and the beliefs of economic agents. *NBER Working Paper*, 2020.

Mark Carney. Guidance. *Speech, CFA Society Toronto, December 12*, 2012.

Carlos Carvalho and Fernanda Nechio. Do people understand monetary policy? *Journal of Monetary Economics*, 66:108–123, 2014.

Richard H. Clarida. The Federal Reserve's new framework: Context and consequences. *At "The Road Ahead for Central Banks," a seminar sponsored by the Hoover Economic Policy Working Group, Hoover Institution, Stanford, California (via webcast)*, 2021.

Olivier Coibion, Yuriy Gorodnichenko, and Saten Kumar. How do firms form their expectations? *American Economic Review*, 108:2671–2713, 2018.

Olivier Coibion, Yuriy Gorodnichenko, and Tiziano Ropele. Inflation expectations and firms' decisions: Causal evidence. *Quarterly Journal of Economics*, 135:165–219, 2020.

Olivier Coibion, Yuriy Gorodnichenko, Edward S. II Knotek, and Raphael Schoenle. Average inflation targeting and household expectations. *Journal of Political Economy – Macroeconomics*, Forthcoming.

Camille Cornand and Cheick Kader M'baye. Does inflation targeting matter? An experimental investigation. *Macroeconomic Dynamics*, 22(2):362–401, 2018.

Jose Dorich, Michael Johnston, Rhys Mendes, Stephen Murchison, and Yang Zhang. ToTEM II: An Updated Version of the Bank of Canada's Quarterly Projection Model. *Bank of Canada Technical Report No. 100*, October 2013.

Michael Ehrmann, Dimitris Georgarakos, and Geoff Kenny. Credibility gains from communicating with the public: Evidence from the ecb's new monetary policy strategy. *ECB Working Paper*, 2022.

Charles M. Evans. Monetary policy in low-inflation environment: Developing a state-contingent price-level target. *Journal of Money, Credit and Banking*, 44(1):147–155, 2012.

Jeffrey A Frankel and Kenneth A Froot. Chartists, fundamentalists, and trading in the foreign exchange market. *The American Economic Review*, 80(2):181–185, 1990.

Xavier Gabaix. A behavioral New Keynesian model. *American Economic Review*, 110(8): 2271–2327, 2020.

Ben Greiner. Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125, 2015.

Andrew Haldane and Michael McMahon. Central bank communications and the general public. In *AEA papers and proceedings*, volume 108, pages 578–83, 2018.

Mathias Hoffmann, Emanuel Moench, Lora Pavlova, and Guido Schultefrankenfeld. Would households understand average inflation targeting? *Journal of Monetary Economics*, 2022.

Cars Hommes and Tomasz Makarewicz. Price level versus inflation targeting under heterogeneous expectations: A laboratory experiment. *Journal of Economic Behavior and Organization*, 182:39–82, 2021. ISSN 0167-2681.

Cars Hommes, Domenico Massaro, and Isabelle Salle. Monetary and fiscal policy design at the zero lower bound: Evidence from the lab. *Economic Inquiry*, 57(2):1120–1140, 2019a.

Cars Hommes, Domenico Massaro, and Matthias Weber. Monetary policy under behavioral expectations: Theory and experiment. *European Economic Review*, 118:193–212, 2019b.

Seppo Honkapohja and Kaushik Mitra. Targeting nominal GDP or prices: Guidance and expectation dynamics. *Bank of Finland Research Discussion Papers*, 2014.

Seppo Honkapohja and Kaushik Mitra. Price level targeting with evolving credibility. *Journal of Monetary Economics*, 116:88–103, 2020.

Chengcheng Jia and Jing Cynthia Wu. Average inflation targeting: Time inconsistency and intentional ambiguity. Technical report, National Bureau of Economic Research, 2022.

Rupal Kamdar. The inattentive consumer: Sentiment and expectations. *Manuscript*, 2019.

Oleksiy Kryvtsov and Luba Petersen. Expectations and monetary policy: Experimental evidence. Technical report, Bank of Canada, 2013.

Oleksiy Kryvtsov and Luba Petersen. Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics*, 117:760–780, 2021.

Andrew T Levin. The design and communication of systematic monetary policy strategies. *Journal of Economic Dynamics and Control*, 49:52–69, 2014.

Bartosz Mackowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803, 2009.

Ulrike Malmendier and Stefan Nagel. Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87, 2016.

Fabio Milani. The modeling of expectations in empirical dsge models: a survey. *DSGE models in macroeconomics: Estimation, evaluation, and new developments*, 2012.

Fatemeh Mokhtarzadeh and Luba Petersen. Coordinating expectations through central bank projections. *Experimental Economics*, 24(3):1–36, 2020.

John F Muth. Rational expectations and the theory of price movements. *Econometrica*, pages 315–335, 1961.

Athanasios Orphanides and John Williams. Imperfect knowledge, inflation expectations and monetary policy. In Ben S. Bernanke and Michael Woodford, editors, *Inflation Targeting*. University of Chicago Press, 2005.

Luba Petersen and Ryan Rholes. Macroeconomic expectations, central bank communication, and background uncertainty: a COVID-19 laboratory experiment. *Journal of Economic Dynamics and Control*, page 104460, 2022.

Damjan Pfajfar and Blaž Žakelj. Experimental evidence on inflation expectation formation. *Journal of Economic Dynamics and Control*, 44:147–168, 2014.

Jerome H. Powell. New economic challenges and the Fed's monetary policy review. *Speech at "Navigating the Decade Ahead: Implications for Monetary Policy," an economic policy symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 27, 2020*.

Ryan Rholes and Luba Petersen. Should central banks communicate uncertainty in their projections? *Journal of Economic Behavior & Organization*, 183:320–341, 2021.

Isabelle Salle. What to target? Insights from theory and lab experiments. *Bank of Canada Staff Working Paper*, 2021.

Thomas J Sargent and Neil Wallace. "Rational" expectations, the optimal monetary instrument, and the optimal money supply rule. *Journal of political economy*, 83(2):241–254, 1975.

Christopher A Sims. Rational inattention and monetary economics. In *Handbook of Monetary Economics*, volume 3, pages 155–181. Elsevier, 2010.

Jeremy Stein. Cheap talk and the Fed. *American Economic Review*, 79(1):32–42, 1989.

Jonathan Swarbrick and Yang Zhang. A horse race of monetary policy strategies for Canada. *Bank of Canada Staff Working Paper*, forthcoming.

Joel Wagner, Yang Zhang, and Thomas Schlanger. A horse race of alternative monetary policy regimes under bounded rationality. *Bank of Canada Staff Working Paper*, forthcoming.

John C. Williams. Preparing for the next storm: Reassessing frameworks and strategies in a low r-star world. *FRB Economic Letter*, 2017-13:1–8, 2017.

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

Online Appendix for
A horse race of monetary policy regimes: An experimental investigation*

Olena Kostyshyna
Bank of Canada

Luba Petersen
Simon Fraser University
NBER

Jing Yang
Bank of Canada

December 1, 2022

This appendix provides the following supplementary materials.

A. Instructions
B. Experimental interface screen shot
C. Simulations under rational and naive expectations
D. Classification of forecasting heuristics
E. Additional experimental results
F. Additional PLT Comm results

* † Kostyshyna and Yang: Bank of Canada, 234 Wellington Street Ottawa, ON, K1A 0G9; kost@bankofcanada.ca, jyang@bankofcanada.ca; Petersen: Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada, lubap@sfu.ca.

A. Instructions

We provide all the common instructions and indicate where the instructions differed across treatments.

1. Inflation Targeting
2. Dual Mandate
3. Average Inflation Targeting - 4 period ahead horizon
4. Average Inflation Targeting - 10 period ahead horizon
5. Price Level Targeting
6. Nominal GDP Level Targeting

Experimental Instructions

Welcome! You are participating in an economics experiment at SFU Experimental Economics Lab. In this experiment you will participate in the experimental simulation of the economy. If you read these instructions carefully and make appropriate decisions, you may earn a considerable amount of money that will be immediately paid out to you in cash at the end of the experiment.

Each participant is paid CDN$7 for attending. Throughout this experiment you will also earn points based on the decisions you make. Every point you earn is worth $0.50. We reserve the right to improve this in your favour if average payoffs are lower than expected.

During the experiment you are not allowed to communicate with other participants. If you have any questions, the experimenter will be glad to answer them privately. If you do not comply with these instructions, you will be excluded from the experiment and deprived of all payments aside from the minimum payment of CDN $7 for attending.

The experiment is based on a simple simulation that approximates fluctuations in the real economy. Your task is to serve as private forecasters and provide real-time forecasts about future output and inflation in this simulated economy. The instructions will explain what output, inflation, and the interest rate are and how they move around in this economy, as well as how they depend on forecasts. We will allow you to practice making forecasts for several unpaid periods before we begin paid periods in this experiment. You will then participate in 50 paid periods.

In this simulation, households and firms (whose decisions are automated by the computer) will form forecasts identically to yours. So to some degree, outcomes that you will see in the game will depend on the way in which you form your forecasts. Your earnings in this experiment will depend on the accuracy of your individual forecasts.

On the next page we will discuss what inflation and output are, and how to predict them. All values will be given in basis points, a measurement often used in descriptions of the economy. All values can be positive, negative, or zero at any point in time.

Your task

Your task in this experiment is to forecast future output and inflation as correctly as possible.
You will submit forecasts for the next period's inflation and output, measured in basis points:

- 1% = 100 basis points
- 3.25% = 325 basis points
- -0.5% = -50 basis points
- -4.8% = -480 basis points

These are just a handful of examples of how basis points work. You can submit any forecast you wish, positive or negative or zero, but please only submit integers.

How the economy evolves

We will now explain the factors that influence output and inflation and the relationships between the different variables in the economy.

The economy consists of six main variables: shocks, inflation, output, interest rate, price level and nominal output. Each period, you will receive the following information that will help you make forecasts.

Current Shock

A shock is a random "event" that directly affects how much people want to spend, and consequently, how much will be produced.

The shocks change every period and are influenced by a random component and past shocks.

More precisely, the shocks that you observe will follow the process specified in your instructions.

At any time period *t*, the shock is calculated as follows:

$$Shock_t = 60 + 0.8(Shock_{t-1}) + Random\ Component_t$$

- The random component is 0 on average, and has a standard deviation of 93 basis points.
- Roughly 2/3rds of the time, the shock is between -155 and +155 basis points
- 95% of the time, the shock will be between -310 and +310 basis points

Intuitively, you can think of the shocks as weather shocks. Over the long run, the weather has no effect on how much consumers want to buy. However, from day to day, there may be random changes to the weather. You can think of a positive shock as unexpectedly nice weather. When the weather is especially nice, consumers are spending more time out of their homes and increasing their expenditures (for example, buying ice cream, going out for a nice dinner, going to the beach). A negative shock can be thought of unexpectedly terrible weather, where no one wants to leave their homes, causing expenditures to be relatively low. Gradually, the shocks, like weather, will revert back to their long-run levels. As the shocks dissipate, new random events occur that will make consumers want to increase or decrease their expenditures.

Consider the following examples:

$$Shock_1 = 30$$

$$Shock_2 = 60 + 30 * 0.8 + Random\ Component_2$$

$$= 60+24 + Random\ Component_2$$

$$Shock_2 = 84 + (-150)$$

$$= -66$$

$$Shock_3 = \ldots$$

Each period, you and the other forecasters will be submitting your beliefs about the following period's output and inflation. The median of each of the forecasts will be employed as the aggregate forecast in the given period and play an important role in determining the current level of output and inflation. The median, rather than the average forecast, is used so that a small number of subjects cannot have a significant effect on the economy.

Output

Output refers to a measure of the quantity of goods the economy is over- or under producing in a given period.

At any time period *t*, output is calculated as follows:

$$Output_t = Median\ Forecast\ of\ output_{t+1} + Median\ Forecast\ of\ inflation_{t+1} - Interest\ Rate_t + Shock_t$$

The value of today's output is determined by the median expectations (forecasts) of tomorrow's output and inflation, as well as today's shock and interest rate. If you, the forecasters, predict that the future economy will be producing more output and there will be more inflation, consumers will want to spend more in the current period. Firms will then produce more to meet consumer demand.

Likewise, positive shocks to consumer demand will have a positive effect on how much will be produced.

Increases in the nominal interest rate will make it more expensive for consumers to borrow and will create more incentive for them to save. With higher interest rates, consumers will decrease their demand for goods, leading to lower production, which will indirectly reduce inflation.

Inflation

Inflation is the rate at which overall prices change between two periods.

At any time period *t*, inflation is calculated as follows:

$$Inflation_t = 0.998(Median\ Forecast\ of\ Inflation_{t+1}) + 0.125(Output_t)$$

Inflation is determined largely by your forecast about future inflation. The idea behind this is simple: If you, the professional forecasters, communicate to the public that inflation is likely to rise in the future, consumers will spend more immediately to avoid paying relatively higher prices (positive inflation) in the future. This increase in demand will cause prices to start rising, i.e. current inflation will increase.

Current output will also have a small positive effect on current inflation. Importantly, variables that affect output will also have a small positive effect on inflation.

You will also have information about other macroeconomic variables that evolve over time.

Price level

The price level is an index measuring the price of output in the economy. The price level evolves with the rate of inflation:

$$Price\ Level_t = Price\ Level_{t-1} + Inflation_t$$

When inflation is positive, the price level increases. When inflation is negative, the price level declines. The price level is shown as an index with a starting value of 1000.

Example 1. Suppose the price level in the previous period is 1000. The inflation rate in the current period is 200 basis points. The price level in the current period is:
Price level = 1000 + 200 = 1200.

Example 2. Instead, suppose the inflation rate in the current period is -200 basis points. The price level in the current period is
Price level = 1000 – 200 = 800.

Nominal output level

The nominal output level is the nominal value of output in the experimental economy. The nominal output level evolves with both output and inflation over time:

$$Nominal\ Output\ Level_t = Output_t + Price\ Level_t$$

The nominal output level is higher when output (production) and the price level are higher, and vice versa. Nominal output is shown as an index with a starting value of 1000.

INFLATION TARGETING TREATMENT

Central Bank Policy

The main objective of the central bank in this experiment is to keep the nominal output level at its target level. The target for nominal output is 1000. The central bank sets the interest rate to bring nominal output to its target.

Interest Rate

The interest rate is the rate at which consumers and firms borrow and save in this experimental economy.

The interest rate responds to the distance between the current inflation rate and its target zero. The interest rate also responds to deviations of output from 0 as they are linked to deviations of inflation from its target. The response to output is much weaker than the response to inflation as output is not the principal target of the Central Bank's policy

At any time period *t*, the interest rate is calculated as follows:

$$\begin{aligned} \textit{Interest Rate}_t &= 60 + 1.5(\textit{Inflation}_t - 0) + 0.125(\textit{Output}_t - 0) \quad && \textit{if Interest Rate}_t > 0 \\ &= 0 && \textit{otherwise} \end{aligned}$$

When inflation is high and above its target of 0 basis points, the central bank will increase interest rates more than one-for-one with inflation. The central bank will also increase interest rate, though less aggressively, in response to positive output. When inflation is further above its target, the increase in the interest rate is larger.

The increase in the interest rate has a direct negative effect on consumer demand and output, and an indirect negative effect on inflation. *When inflation is above target, a higher interest rate leads to lower inflation and thus helps bring it back towards its target.*

When inflation is below the target of 0 basis points, the central bank will decrease interest rates more than one-for-one with negative inflation. The central bank will also decrease the interest rate in response to negative output, though less aggressively. When inflation is further below its target, the decrease in the interest rate is larger.

Lower interest rates have a direct positive effect on consumer demand and output, and an indirect positive effect on inflation. *When inflation is below target, a lower interest rate leads to higher inflation and thus helps bring it back towards its target.*

It is also important for you to realize that, even though the central bank is aiming for inflation at its target of zero, it may not be able to accomplish this every period because of the persistent random shocks that are occurring each period and the public's (your) expectations. However, the economy will be kept relatively more stable as a consequence of the central bank's reaction to inflation and output.

Note that the central bank cannot lower interest rates below zero. For large negative values of inflation and output, the central bank will simply set the interest rate at zero.

You will not observe the current interest rate when you are forming your forecast about the following period's inflation and output. After you submit your forecasts, the computer will simultaneously solve for the current period's inflation, output and interest rate taking into consideration the forecasts and the realized shock.

DUAL MANDATE TREATMENT

Central Bank Policy

The main objective of the central bank in this experiment is to keep the inflation rate and output at their targets. The inflation target is equal to 0 basis points. The target for output is 0 basis points as well. The central bank sets the interest rate to bring the inflation rate and output to their targets.

Interest Rate

The interest rate is the rate at which consumers and firms borrow and save in this experimental economy.

The interest rate responds to the distance between the current inflation rate and its target zero. The interest rate also responds to the distance between the current output and its target zero.

At any time period *t*, the interest rate is calculated as follows:

$$\textit{Interest Rate}_t = 60 + 4.5(\textit{Inflation}_t - 0) + 4.5(\textit{Output}_t - 0) \quad \textit{if Interest Rate}_t > 0$$
$$= 0 \quad \textit{otherwise}$$

When inflation and output are high and above their targets of 0 basis points, the central bank will increase interest rates more than one-for-one with inflation and output. When inflation and output are further above their targets, the increase in the interest rate is larger.

The increase in the interest rate has a direct negative effect on consumer demand and output, and an indirect negative effect on inflation. *When inflation and output are above target, a higher interest rate leads to lower inflation and output and thus helps bring both back towards their targets.*

When inflation and output are low and below their targets of 0 basis points, the central bank will decrease interest rates more than one-for-one with negative inflation and output. When inflation and output are further below their target, the decrease in the interest rate is larger.

Lower interest rates have a direct positive effect on consumer demand and output, and an indirect positive effect on inflation. *When inflation and output are below target, a lower interest rate leads to higher inflation and output and thus helps bring both back towards their targets.*

It is also important for you to realize that, even though the central bank is aiming for inflation at its target of zero, it may not be able to accomplish this every period because of the persistent random shocks that are occurring each period and the public's (your) expectations. However, the economy will be kept relatively more stable as a consequence of the central bank's reaction to inflation and output.

Note that the central bank cannot lower interest rates below zero. For large negative values of inflation and output, the central bank will simply set the interest rate at zero.

You will not observe the current interest rate when you are forming your forecast about the following period's inflation and output. After you submit your forecasts, the computer will simultaneously solve for the current period's inflation, output and interest rate taking into consideration the forecasts and the realized shock.

AVERAGE INFLATION TARGETING – 4 PERIOD AND 10 PERIOD HORIZON TREATMENTs

Central Bank Policy

The main objective of the central bank in this experiment is to keep the average inflation rate over 4 (10) periods at its target. The average inflation target is equal to 0 basis points. The central bank sets the interest rate to bring the average inflation rate to its target.

Interest Rate

The interest rate is the rate at which consumers and firms borrow and save in this experimental economy.

The interest rate responds to the distance between the average inflation rate over the current and past 3 (10) periods and its target zero. The interest rate also responds to deviations of output from 0 as they are linked to deviations of inflation from its target. The response to output is much weaker than the response to inflation as output is not the principal target of the Central Bank's policy.

At any time period *t*, the interest rate is calculated as follows:

$$\begin{aligned} \textit{Interest Rate}_t &= 60 + 5.5(\textit{Average Inflation}_t - 0) + 3(\textit{Output}_t - 0) && \textit{if Interest Rate}_t > 0 \\ &= 0 && \textit{otherwise} \end{aligned}$$

where

$\textit{Average Inflation}_1 = \textit{Inflation}_1$	in Period 1
$\textit{Average Inflation}_2 = (\textit{Inflation}_1 + \textit{Inflation}_2)/2$	in Period 2
$\textit{Average Inflation}_t = (\textit{Inflation}_1 + \textit{Inflation}_2 + \textit{Inflation}_3)/3$	in Period 3
$\textit{Average Inflation}_t = (\textit{Inflation}_t + \textit{Inflation}_{t-1} + \textit{Inflation}_{t-2} + \textit{Inflation}_{t-3}) / 4$	in Periods 4+
$(\textit{Average Inflation}_t = (\textit{Inflation}_t + \textit{Inflation}_{t-1} + \textit{Inflation}_{t-2} + \ldots + \textit{Inflation}_{t-9}) / 4$	in Periods 10+)

When average inflation is high and above its target of 0 basis points, the central bank will increase interest rates more than one-for-one with average inflation. The central bank will also increase interest rate, though less aggressively, in response to positive output. When average inflation is further above its target, the increase in the interest rate is larger.

The increase in the interest rate has a direct negative effect on consumer demand and output, and an indirect negative effect on inflation. *When inflation is above target, a higher interest rate leads to lower inflation and thus helps bring average inflation back towards its target.*

When average inflation is below the target of 0 basis points, the central bank will decrease interest rates more than one-for-one with negative average inflation. The central bank will also decrease the interest rate in response to negative output, though less aggressively. When average inflation is further below its target, the decrease in the interest rate is larger.

Lower interest rates have a direct positive effect on consumer demand and output, and an indirect positive effect on inflation. *When average inflation is below target, a lower interest rate leads to higher inflation and thus helps bring average inflation back towards its target.*

It is also important for you to realize that, even though the central bank is aiming for average inflation at its target of zero, it may not be able to accomplish this every period because of the persistent random shocks that are occurring each period and the public's (your) expectations. However, the economy will be kept relatively more stable as a consequence of the central bank's reaction to inflation and output.

Note that the central bank cannot lower interest rates below zero. For large negative values of average inflation and output, the central bank will simply set the interest rate at zero.

You will not observe the current interest rate when you are forming your forecast about the following period's inflation and output. After you submit your forecasts, the computer will simultaneously solve for the current period's inflation, output and interest rate taking into consideration the forecasts and the realized shock.

PRICE LEVEL TARGETING TREATMENT

Central Bank Policy

The main objective of the central bank in this experiment is to keep the price level at its target level. The target for the price level is 1000. The central bank sets the interest rate to bring nominal output to its target.

Interest Rate

The interest rate is the rate at which consumers and firms borrow and save in this experimental economy.

The interest rate responds to the distance between the price level and its target level of 1000. The interest rate also responds to deviations of output from 0 as they are linked to deviations of the price level from its target.

At any time period *t*, the interest rate is calculated as follows:

$$\begin{aligned} \textit{Interest Rate}_t &= 60 + 0.8(\textit{Price Level}_t - 1000) + 1.3(\textit{Output}_t - 0) && \textit{if Interest Rate}_t > 0 \\ &= 0 && \textit{otherwise} \end{aligned}$$

When the price level is high and above its target of 1000 basis points, the central bank will increase interest rates. The central bank will also increase interest rate in response to positive output. When the price level is further above its target, the increase in the interest rate is larger.

The increase in the interest rate has a direct negative effect on consumer demand and output, and an indirect negative effect on inflation, and thus the price level. *When the price level is above target, a higher interest rate leads to lower inflation and thus helps bring the price level back towards its target.*

When the price level is below the target of 1000 basis points, the central bank will decrease interest rates. The central bank will also decrease the interest rate in response to negative output. When the price level is further below its target, the decrease in the interest rate is larger.

Lower interest rates have a direct positive effect on consumer demand and output, and an indirect positive effect on inflation, and thus the price level. *When the price level is below target, a lower interest rate leads to higher inflation and thus helps bring the price level back towards its target.*

It is also important for you to realize that, even though the central bank is aiming for the price level to be at its target of 1000, it may not be able to accomplish this every period because of the persistent random shocks that are occurring each period and the public's (your) expectations. However, the economy will be kept relatively more stable as a consequence of the central bank's reaction to the price level and output.

Note that the central bank cannot lower interest rates below zero. For low price levels and large negative values of output, the central bank will simply set the interest rate at zero.

You will not observe the current interest rate when you are forming your forecast about the following period's inflation and output. After you submit your forecasts, the computer will simultaneously solve for the current period's inflation, output and interest rate taking into consideration the forecasts and the realized shock.

NOMINAL GDP LEVEL TARGETING TREATMENT

Central Bank Policy

The main objective of the central bank in this experiment is to keep the nominal output level at its target level. The target for nominal output is 1000. The central bank sets the interest rate to bring nominal output to its target.

Interest Rate

The interest rate is the rate at which consumers and firms borrow and save in this experimental economy.

The interest rate responds to the distance between nominal output level and its target level of 1000.

At any time period *t*, the interest rate is calculated as follows:

$$\textit{Interest Rate}_t = \begin{cases} 60 + 1.1(\textit{Nominal Output}_t - 1000) & \textit{if Interest Rate}_t > 0 \\ = 0 & \textit{otherwise} \end{cases}$$

When the level of nominal output is above its target level of 1000, the central bank will increase interest rates more than one-for-one in response to this discrepancy. When nominal output is further above its target, the increase in the interest rate is larger.

The increase in the interest rate has a direct negative effect on consumer demand and output, and an indirect negative effect on inflation. When inflation decreases, the price level decreases. As output and the price level decrease, nominal output decreases. Thus, *when nominal output is above its target, higher interest rate leads to lower nominal output and thus helps bring it back towards its target.*

When the level of nominal output is below its target level of 1000, the central bank will decrease interest rates more than one-for-one in response to this discrepancy. When nominal output is further below its target, the decrease in the interest rate is larger.

The decrease in the interest rate has a direct positive effect on consumer demand and output, and an indirect positive effect on inflation. When inflation increases, the price level increases. As output and price level increase, nominal output increases. *Thus, when nominal output is below its target, lower interest rate leads to higher nominal output and thus helps bring it back towards its target.*

Lower interest rates have a direct positive effect on consumer demand and output, and an indirect positive effect on inflation, and thus the price level. *When the price level is below target, a lower interest rate leads to higher inflation and thus helps bring the price level back towards its target.*

It is also important for you to realize that, even though the central bank is aiming for a stable level of nominal output at its target of 1000, it may not be able to accomplish this every period because of the persistent random shocks that are occurring each period and the public's (your) expectations. However, the economy will be kept relatively more stable as a consequence of the central bank's reaction to the nominal output from its target.

Note that the central bank cannot lower interest rates below zero. For low nominal outputs, the central bank will simply set the interest rate at zero.

You will not observe the current interest rate when you are forming your forecast about the following period's inflation and output. After you submit your forecasts, the computer will simultaneously solve for the current period's inflation, output and interest rate taking into consideration the forecasts and the realized shock.

Score

Your score will depend on the accuracy of your inflation and output forecasts. The absolute difference between your forecasts and the actual values for output and inflation are your absolute forecast errors.

Absolute Forecast Error = absolute (Your Forecast – Actual Value)
Total Score = 0.30(2^-0.01(Absolute Forecast Error for Output)) + 0.30(2^-0.01 (Absolute Forecast Error for Inflation))

The maximum score you can earn each period is 0.6 points.
Your score will decrease as your forecast error increases. Suppose your forecast errors for each of output and inflation are:

0	-Your score will be 0.6	300	-Your score will be 0.075
50	-Your score will be 0.42	500	-Your score will be 0.02
100	-Your score will be 0.3	1000	-Your score will be 0
200	-Your score will be 0.15	2000	-Your score will be 0

Information about the Interface, Actions, and Payoffs

During the experiment, your main screen will display information that will help you make forecasts and earn more points.

At the top left of the screen, you will see your subject number, the current period, time remaining, and the total number of points earned. You will also see four history plots.

The top history plot displays past interest rates and past and current shocks.

The second plot displays your past forecasts of inflation and realized inflation levels. (*IT/DM/AIT Treatments:* You will also be shown the central bank's inflation target of 0 in orange.)

The third plot displays your past forecast of output and realized output levels.

Your forecasts will always be shown in blue while the realized value will be shown in red. You can see the exact value for each point on a graph by placing your mouse at that point. The difference between your forecasts and the actual realized levels constitutes your forecast errors.

The fourth plot will show price level and nominal output. The price level will be presented on the left axis in purple while the nominal output will be presented on the right axis in green. (*PLT Treatment:* You will also be shown the central bank's nominal output level target of 1000 in orange.) (*NGDP Treatment:* You will also be shown the central bank's nominal output level target of 1000 in orange.)

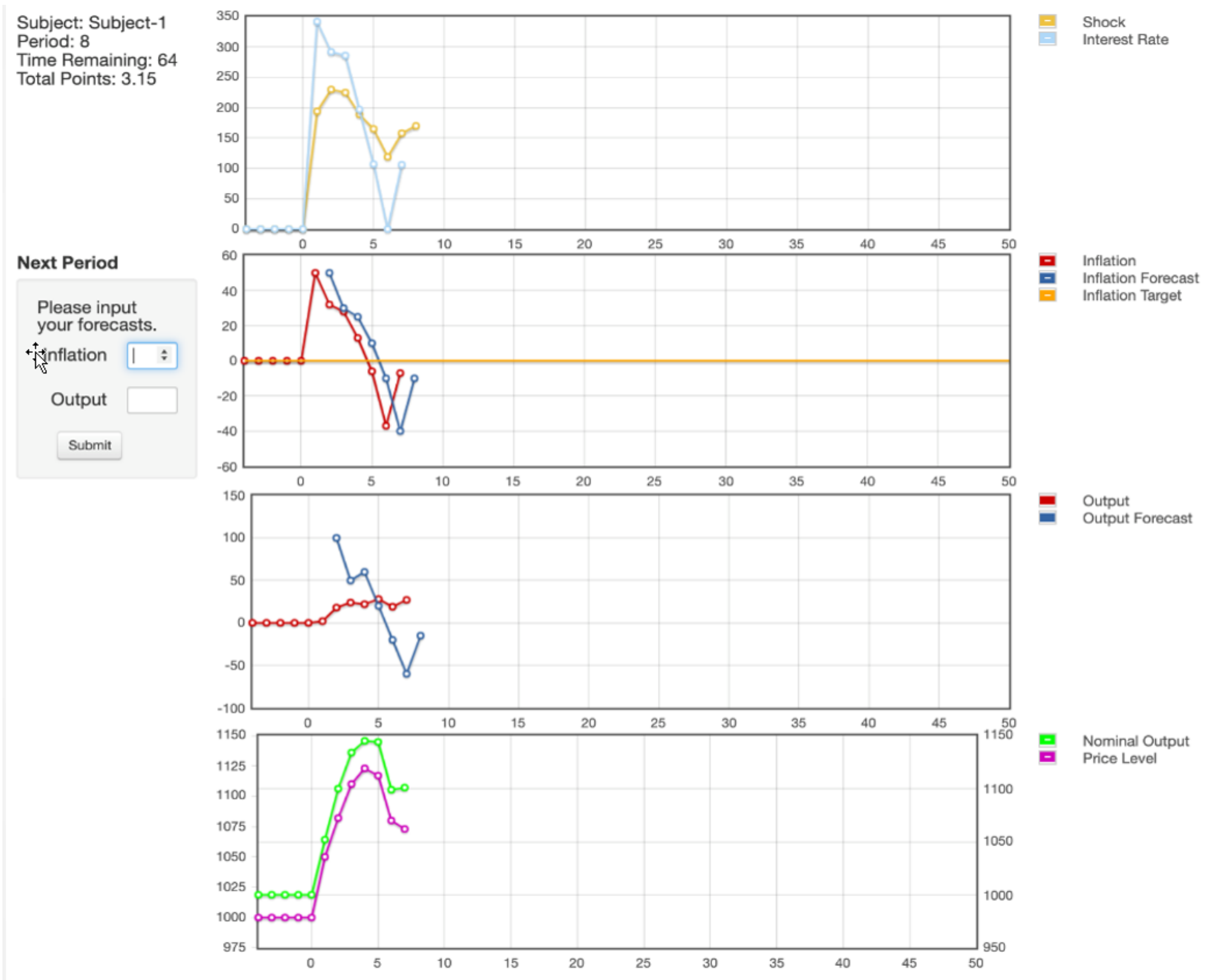
You may submit positive, negative or zero forecasts. Please use whole numbers. Please review your forecasts before pressing the SUBMIT button. Once the SUBMIT button has been clicked, you will not be able to revise your forecasts until the next period. You will earn zero points if you do not submit both forecasts.

You will have 75 seconds to submit forecasts for output and inflation for the first 10 rounds, and 60 seconds for the remaining 40 periods. Your score converted into Canadian dollars ($0.50 per point) plus the show up fee will be paid to you in cash at the end of the experiment.

B. Experimental interface

Participants interacted in an online interface where they repeatedly made inflation and output forecasts. Figure B1 presents a sample screenshot from the inflation targeting treatment.

Figure B1: Screenshot of participants' screens during the experiment



C. Simulations with rational and naive expectations

Table C1 presents the breakdown of losses associated with deviations of inflation, output, and nominal interest rates from target, for each phase and treatment. Simulations are conducted under the assumption that expectations are rational/model-consistent.

Table C1: Losses associated with inflation, output, and interest rate in REE

	$\sqrt{\text{total loss}/T}$	$\sqrt{\sum^T(\pi-\pi^*)^2/T}$	$\sqrt{\sum^T(x-x^*)^2/T}$	$\sqrt{\sum^T(i-i^*)^2/T}$
periods 1-50				
NGDP	168.2	14.8	82.8	206
PLT	169.8	9.3	40.4	232.8
AIT-10	179.7	20.2	56.5	239.6
AIT-4	180.8	22	47.4	244.8
DM	184.4	21.7	38.1	253.4
IT	186.9	24.8	43	254.8
periods 1-19				
NGDP	153.8	11.8	75.5	188.8
PLT	155.9	7.5	34.3	214.8
AIT-10	164	18.2	49.8	219.5
AIT-4	165.3	20	42.1	224.2
DM	168.5	19.6	32.1	232.4
IT	170.8	22.4	36.8	233.7
periods 20-50				
NGDP	176.4	16.4	87	215.8
PLT	177.7	10.2	43.7	243.2
AIT-10	188.7	21.3	60.2	251.2
AIT-4	189.7	23.1	50.3	256.6
DM	193.5	22.8	41.3	265.4
IT	196.2	26.1	46.5	267

Loss and standard deviations of inflation, output, and interest rate were computed from simulations with rational expectations and are expressed in basis points.

Simulations with naive agents

Other experimental studies of monetary policy regimes illustrate that participants' expectations are mostly non-rational [Anufriev et al., 2013, Assenza et al., 2019]. Given this evidence, we introduce a very simple form of bounded rationality – naive expectations – into our model to understand the implications for stabilization properties of different monetary policy regimes. Naive expectations are set as $E_t\pi_{t+1} = \pi_{t-1}$ and $E_t x_{t+1} = x_{t-1}$. We find that the presence of naive agents can be disruptive to economies with certain monetary policy regimes. Level-targeting regimes such as PLT and NGDP can break down at certain shares of naive agents. The threshold share of naive agents is 33% in the PLT regime and 45% in NGDP; economies become unstable with shares of naive agents above the threshold level. IT, DM, and AIT tolerate 100% of

Table C2: Ranking of regimes in the simulations with RE, naive expectations and in the data from laboratory experiments

ranking	REE Table C1	naive =33% Table C3	lab data Table 2
Periods 1-19			
1	NGDP	NGDP	AIT-10
2	PLT	PLT	AIT-4
3	AIT-10	AIT-10	DM
4	AIT-4	IT	IT
5	DM	DM	PLT
6	IT	AIT-4	NGDP
Periods 20-50			
1	NGDP	NGDP	IT
2	PLT	IT	DM
3	AIT-10	DM	AIT-4
4	AIT-4	AIT-4	AIT-10
5	DM	PLT	PLT
6	IT	AIT-10	NGDP
All Periods			
1	NGDP	NGDP	AIT-4
2	PLT	IT	DM
3	AIT-10	DM	IT
4	AIT-4	AIT-4	AIT-10
5	DM	PLT	PLT
6	IT	AIT-10	NGDP

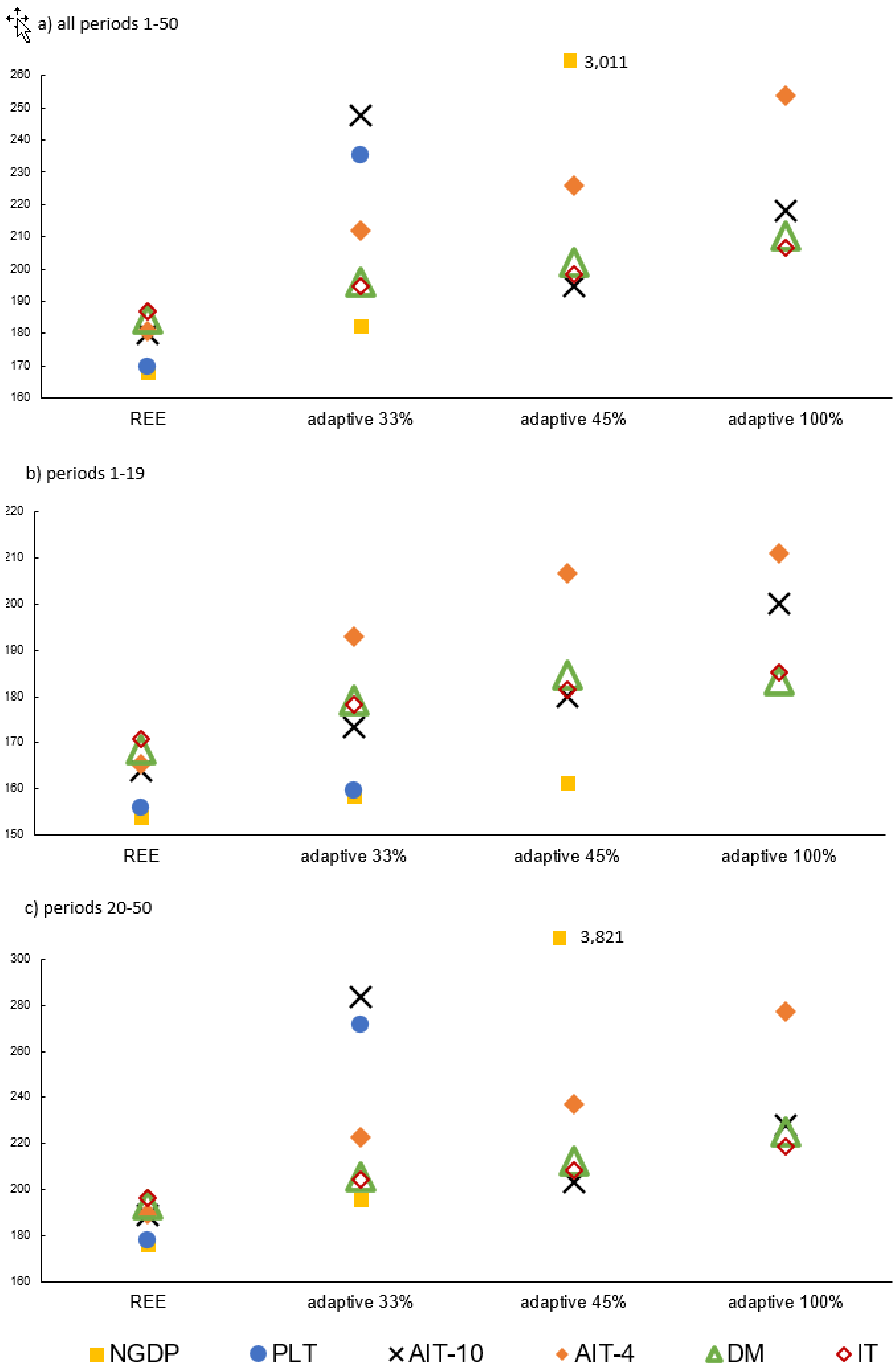
Note: Ranking in column "naive=33%" is based on simulations with 33% of naive agents in all regimes. For details on losses for these simulations, see Table C3.

naive agents, remaining stable. In other words, PLT and NGDP are the least robust to the presence of naive expectations.

We have simulated our model with different shares of naive expectations using a sequence of demand shocks implemented in the experiment. Figure C1 presents losses from the simulations with rational expectations and simulations with different shares of naive agents – 33%, 45%, and 100%. Table C3 reports results from simulation with 33% of naive agents in all regimes, Table C4 reports results from simulation with 33% in PLT regime and 45% in the rest of the regimes, and Table C5 reports simulations with 33% in PLT, 45% in NGDP, and 100% in IT, DM and AIT.

The results presented in Figure C1 lead to the following interesting observations. First, the presence of naive agents leads to higher losses across all regimes. Second, the increase of the share of naive agents leads to the increase of the losses for all regimes (except for AIT-10 following the shock). Third, the ranking of monetary policy regimes changes with an increase of the share of naive agents: performance of history-dependent regimes (AIT, PLT, NGDP) deteriorates relative to regimes responding to current inflation and output gap (IT and DM). We would like to note that AIT performs better than level-targeting regimes PLT and NGDP. These simple simulations with naive expectations illustrate the important role expectations play in the ability of different policy regimes to stabilize the economy.

Figure C1: Summary of losses from simulations



This figure shows results from simulations of New Keynesian model with rational expectations and simulations with naive agents. We vary the shares of naive agents from 0% (REE) to 33% (threshold in PLT), 45% (threshold in NGDP) and 100%.

Table C3: Losses associated with inflation, output, and interest rate from simulations with both rational and naive agents

	$\sqrt{\text{total loss}/T}$	$\sqrt{\sum^T (\pi-\pi^*)^2/T}$	$\sqrt{\sum^T (x-x^*)^2/T}$	$\sqrt{\sum^T (i-i^*)^2/T}$
periods 1-50				
NGDP	182.5	29.5	87.9	222.3
IT	194.8	34.8	27.5	268.3
DM	195.8	35.3	25.6	270
AIT-4	212	54.9	53.8	279.5
PLT	235.3	63.3	144.3	247.1
AIT-10	247.8	68.5	154.9	255.8
periods 1-19				
NGDP	158.5	17.4	63	204.2
PLT	159.6	11.6	29.2	221.2
AIT-10	173.4	27.8	44	233.9
IT	178.3	31.3	25	245.7
DM	179.2	31.8	23.4	247.2
AIT-4	193.1	48.9	45.9	256
periods 20-50				
NGDP	195.8	34.9	100.2	232.7
IT	204.3	36.8	28.9	281.2
DM	205.3	37.3	26.9	283
AIT-4	222.9	58.2	58	293
PLT	271.4	79.9	181.8	261.6
AIT-10	283.9	84.3	193.7	268.4

Losses are computed from simulations with combination of rational expectations (67%) and naive expectations (33%) and are expressed in basis points.

Next, we discuss the mechanism through which naive expectations weaken the performance of history-dependent regimes. As described above, naive agents form their expectations for the next period based on the realization in the previous period. These expectations are purely backward-looking and do not incorporate an understanding of what a monetary policy regime aims to achieve (stabilize inflation and output) and how it works to achieve it. In other words, naive expectations do not have a forward-looking aspect and, as a result, they weaken the expectations channel on which history-dependent regimes rely for their superior performance in models with rational expectations. In addition, in the economy with naive agents, rational agents account for the non-rationality of naive agents and adjust their expectations relative to expectations in the model with only rational agents. Thus, the presence of naive agents diminishes the effectiveness of the expectations channel. Given that history-dependent regimes such as AIT, PLT, and NGDP rely heavily on the expectations channel, the performance of these regimes deteriorates as the expectations channel weakens.

The simulations with naive agents suggest that IT and DM may be more robust to the presence of non-rational expectations in their ability to stabilize the economy. And these simulations may be indicative of how these regimes may perform in the laboratory, where expectations are likely to deviate from rationality.

Table C4: Losses associated with inflation, output, and interest rate from simulations with both rational and naive agents

	$\sqrt{\text{total loss}/T}$	$\sqrt{\sum^T (\pi - \pi^*)^2/T}$	$\sqrt{\sum^T (x - x^*)^2/T}$	$\sqrt{\sum^T (i - i^*)^2/T}$
periods 1-50				
AIT-10	194.7	37	43.6	263.2
IT	198.4	38.5	23.8	273.2
DM	202.2	42.6	21.1	278
AIT-4	226	69	44.8	297.7
PLT	235.3	63.3	144.3	247.1
NGDP	3010.6	1285.2	2691.2	581.7
periods 1-19				
PLT	159.6	11.6	29.2	221.2
NGDP	161.4	21.6	58	210.9
AIT-10	180.1	35.3	41	242.9
IT	181.4	34.7	21.1	250
DM	184.5	37.7	19.6	253.9
AIT-4	206.8	63	41	272.4
periods 20-50				
AIT-10	203.1	38	45.1	274.9
IT	208.2	40.7	25.3	286.5
DM	212.3	45.3	22	291.7
AIT-4	237	72.5	47	312.2
PLT	271.4	79.9	181.8	261.6
NGDP	3821.4	1632.2	3417.5	720.1

Losses are computed from simulations with a combination of rational expectations and naive expectations and are expressed in basis points. Shares of naive expectations: 33% in PLT and 45% in all other regimes.

Post-shock dynamics in the presence of naive agents

We highlight two additional observations from Figure C1, both about the performance of monetary policy regimes following the ELB shock. First, in the presence of naive agents (33%), NGDP performs better than other regimes, and notably better than PLT, another level-targeting regime. Second, the performance of AIT-10 does not deteriorate monotonically with the increase in the share of naive agents in the post-shock period.

Post-shock dynamics in NGDP and PLT with naive agents

During a stable period, with the share of naive agents at 33%, NGDP performs somewhat better than PLT, and these two regimes outperform other regimes. However, following the ELB shock, the performance of NGDP remains better than that of other regimes, while PLT performs much worse. In other words, during stable times, PLT can handle the presence of naive agents as well as NGDP does, but after ELB shock, PLT deteriorates substantially relative to NGDP and other regimes. Why can NGDP handle the period after ELB shock better than PLT? And why does PLT deteriorate so much?

Table C5: Losses associated with inflation, output, and interest rate from simulations with both rational and naive agents

	$\sqrt{\text{total loss}/T}$	$\sqrt{\sum^T(\pi-\pi^*)^2/T}$	$\sqrt{\sum^T(x-x^*)^2/T}$	$\sqrt{\sum^T(i-i^*)^2/T}$
periods 1-50				
IT	206.7	44.5	28.5	282.6
DM	210.1	49.9	23.5	286.7
AIT-10	217.9	54.1	52	289.3
PLT	235.3	63.3	144.3	247.1
AIT-4	254	95.6	42	327.4
NGDP	3010.6	1285.2	2691.2	581.7
periods 1-19				
PLT	159.6	11.6	29.2	221.2
NGDP	161.4	21.6	58	210.9
DM	183.4	34.2	25	252.3
IT	185.1	35.2	26.4	254.3
AIT-10	200	50	35.8	269.2
AIT-4	211	63.2	46.1	277.1
periods 20-50				
IT	218.9	49.4	29.7	298.6
DM	224.9	57.4	22.6	305.9
AIT-10	228.2	56.5	59.8	301
PLT	271.4	79.9	181.8	261.6
AIT-4	277.1	110.8	39.3	354.8
NGDP	3821.4	1632.2	3417.5	720.1

Losses are computed from simulations with combination of rational expectations and naive expectations and are expressed in basis points. Shares of naive expectations: 33% in PLT, 45% in NGDP, and 100% in IT, DM, AIT-4, and AIT-10.

The deterioration of PLT's performance after ELB shock is mostly due to higher volatility of output (Table C3). The focus of PLT is on the stabilization of price level and making up for all past misses from the price-level target; therefore, price-level stability comes at a price of higher output volatility. In contrast, NGDP aims at the stability of nominal output where deviations of price level can be compensated with deviations in output level, keeping nominal output stable. Indeed, NGDP has lower volatility of output than PLT during post-shock periods 20-50. And so, after ELB shock, PLT overreacts to deviations of the price level from the target, leading to higher output volatility, thus reducing its stabilization performance. Such a focus on price level also reduces the performance of PLT relative to IT and DM. This result is related to the finding in Hommes et al. [2019] who shows that in the presence of backward-looking expectations responding to output is important for stabilizing inflation and output.

It is worth noting that with an increase in the share of naive agents to the 45% threshold level (the threshold level for NGDP), the performance of NGDP declines below that of all other regimes. A larger presence of naive agents further weakens the expectations channel and consequently, NGDP can no longer outperform other regimes.

Post-shock dynamics in AIT-10 with naive agents
The relationship between the presence of naive agents and the performance of AIT-10 is non-monotonic during the post-shock period: losses increase when the share rises from 0 to 33%, then decline when the share is 45% and increase when the share increases further to 100% (Figure C1, panel c), while during stable periods 1-19, an increase in the share of naive agents leads to the monotonic decline in the performance of AIT-10 (Figure C1, panel b).

The presence of naive agents brings two effects. First, naive expectations carry over strength in inflation and output from past periods, which leads to smaller declines in these variables after ELB shock than in the case with rational expectations. Second, expectations channels weaken directly because of naive agents and indirectly because rational agents adjust their expectations to account for the presence of naive agents.

AIT-10 performs worse at 33% than at 0% because the presence of naive agents weakens the expectations channel (the second effect dominates). When the share of naive agents increases from 33% to 45%, the first effect becomes more pronounced. As a higher share of backward-looking expectations carries over some of the past strength in economic variables, the ELB episode is less severe and shorter with AIT-10 than in other regimes, resulting in its better performance. However, when the share of naive agents reaches 100%, the complete absence of rational agents destroys forward-looking expectations and the expectations channel necessary for AIT-10 to work. As a result, at 100% AIT-10 performed worse than at 45%.

D. Classification of forecasting heuristics

We use experimental data on participants' forecasts to determine how participants form their forecasts. We consider several types of expectation formation and assign a type to each participant that best fits their forecasting behaviour. Table D1 summarizes all mechanisms we have considered. As we discussed above, the experimental economies significantly deviate from rational expectations equilibrium paths. Therefore, we need to consider other types of forecasting mechanisms in addition to rational expectations. We study several heuristics that have been previously used in the literature on the formation of expectations in macroeconomic models. The simplest deviation from rational expectations is cognitive discounting à la Gabaix [2020], with expectations that are somewhat short (a share α) of a rational expectations solution. Cognitive discounting weakens the expectations channel and makes history-dependent rules less successful than rational expectations. We also consider a heuristic in which participants forecasts' are based on steady state/target.

We also consider backward-looking mechanisms in which the formation of expectations is history driven. Constant-gain learning has been widely used in the literature on the role of learning in macroeconomics [Evans and Honkapohja, 2001] and implications for monetary policy [Bullard and Mitra, 2002] and is supported by the evidence of such expectations in the survey data [Branch, 2004]. We also consider trend-chasing expectations that have been shown to be used in the experimental data [Hommes et al., 2019, Anufriev et al., 2013, Assenza et al., 2019]. Trend-chasing expectations nest naive expectations $E_{i,t}x_{t+1} = x_{t-1}$ under a trend-chasing parameter $\tau = 0$, and survey data provides empirical evidence of the use of naive expectations [Branch, 2004].

Table D1: Forecasting heuristics: models of expectations as functions of exogenous or historical data.

Model	Heuristic Name	Model
M1	Ex–Ante Rational	$E_{i,t}x_{t+1} = f(r^n_{t-1}, \epsilon_t)$
		$E_{i,t}\pi_{t+1} = f(r^n_{t-1}, \epsilon_t)$
M2	Cognitive Discounting	$E_{i,t}x_{t+1} = \alpha f(r^n_{t-1}, \epsilon_t)$
		$E_{i,t}\pi_{t+1} = \alpha f(r^n_{t-1}, \epsilon_t)$
M3	Constant Gain	$E_{i,t}x_{t+1} = E_{i,t-1}x_t - \gamma(E_{i,t-2}x_{t-1} - x_{t-1})$
		$E_{i,t}\pi_{t+1} = E_{i,t-1}\pi_t - \gamma(E_{i,t-2}\pi_{t-1} - \pi_{t-1})$
M4	Steady State/Target	$E_{i,t}x_{t+1} = 0$
		$E_{i,t}\pi_{t+1} = 0$
M5	Trend Chasing	$E_{i,t}x_{t+1} = x_{t-1} + \tau(x_{t-1} - x_{t-2})$
		$E_{i,t}\pi_{t+1} = \pi_{t-1} + \tau(\pi_{t-1} - \pi_{t-2})$

$\alpha \in [0.1, 0.9]$, γ and $\tau \in [0, 1.5]$ in increments of 0.1.

We determine the forecasting heuristic for each participant that best fits their forecasting behavior during each of the phases of the experiment. To do so, we compute the mean absolute error of each participant's expectations for each of the heuristics presented in Table D1. For some heuristics such as M2 Cognitive Discounting, M4 Constant Gain, and M5 Trend-Chasing, we consider a wide range of parameterizations. The cognitive discounting parameter α can take values in the range of $[0.1, 0.9]$. Constant gain parameter γ and trend-chasing parameter τ are in the range of $[0, 1.5]$. We consider values of these parameters from these ranges with an increment of 0.1. We assign each participant the heuristic and its parameter value (if applicable) that produces the lowest mean absolute error. Participants' forecasts can be classified under different heuristics for two different phases of the experiment. The prevalence of the assigned heuristics, by treatment and phase, are presented in Figure D1 for inflation forecasts and Figure D2 for output forecasts.

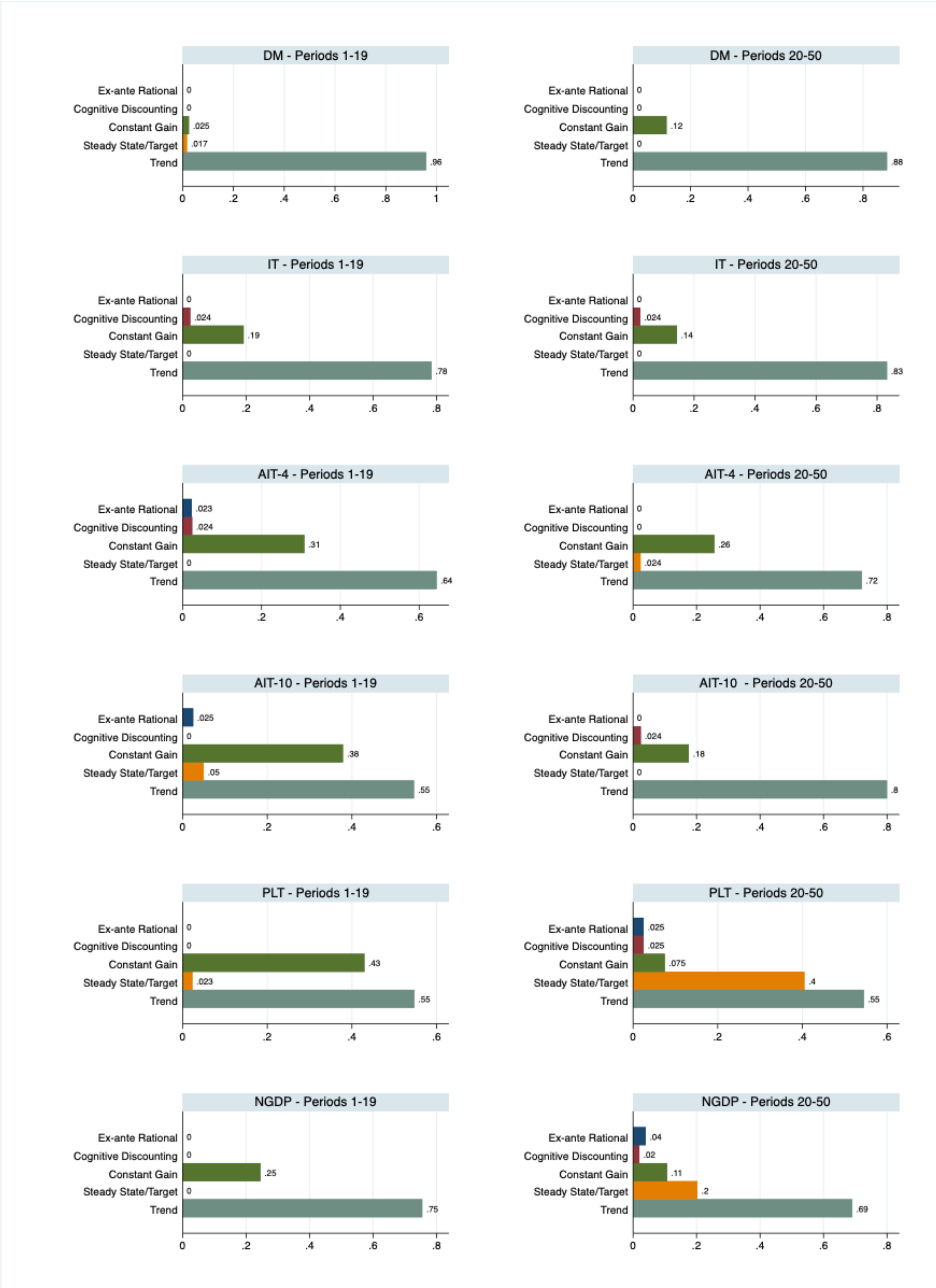
The most striking result is the rarity of the rational expectations in the experimental data. Fewer than 5% of participants in any treatment can be classified as rational or *model consistent*, and in some cases, this share is close to zero. The consistent lack of rational expectations suggests that participants do not broadly appreciate how economic fundamentals influence aggregate dynamics. Participants' elicited expectations show little evidence that they made their forecasts in a forward-looking way responding to the dynamics of shocks and internalizing the stabilizing role of monetary policy. There remains a possibility that participants may understand these elements but not fully, i.e. they may use cognitive discounting, which, like rational expectations, assumes agents are forward-looking and respond solely to aggregate demand shocks, albeit in a more muted manner. However, we observe very little use of the cognitive discounting model. In most treatments, we observe under 5% of participants using cognitive discounting, and frequently this share is close to zero. We observe the highest incidence of cognitive discounting in the post-shock phase of the NGDP treatment, with 11% of participants. While this small minority of participants tried to use their understanding of this policy regime, their forecasts were insufficiently rational and their share was too small to pull NGDP economies out of deflationary spiral in the post-shock phase.

We emphasize that although participants in the rate-targeting treatments are not especially model consistent in their forecasting, their beliefs are not wildly different from rational expectations. They are relatively well anchored on the inflation target. Moreover, the consistency in aggregate dynamics across sessions in rate-targeting treatments suggests a consistent aggregate forecasting heuristic and not just random or confused expectations.

Backward-looking expectations – trend-chasing and constant gain learning – are the dominant forecasting heuristics in most of our treatments. Together, these backward-looking heuristics make up the majority of participants' forecasts. In inflation forecasts, the largest share of participants use trend-chasing during both pre-shock and post-shock phases, with trend-chasing becoming more prevalent post-shock with a decline of constant gain learning. We observe similar composition and evolution of heuristics in output forecasts across all treatments, except for AIT-4 and DM. AIT-4 and DM stand out among the treatments: the share of constant-gain learning is considerably larger than the share of trend-chasing in the pre-shock phase, although it declines below that of trend-chasing post-shock.

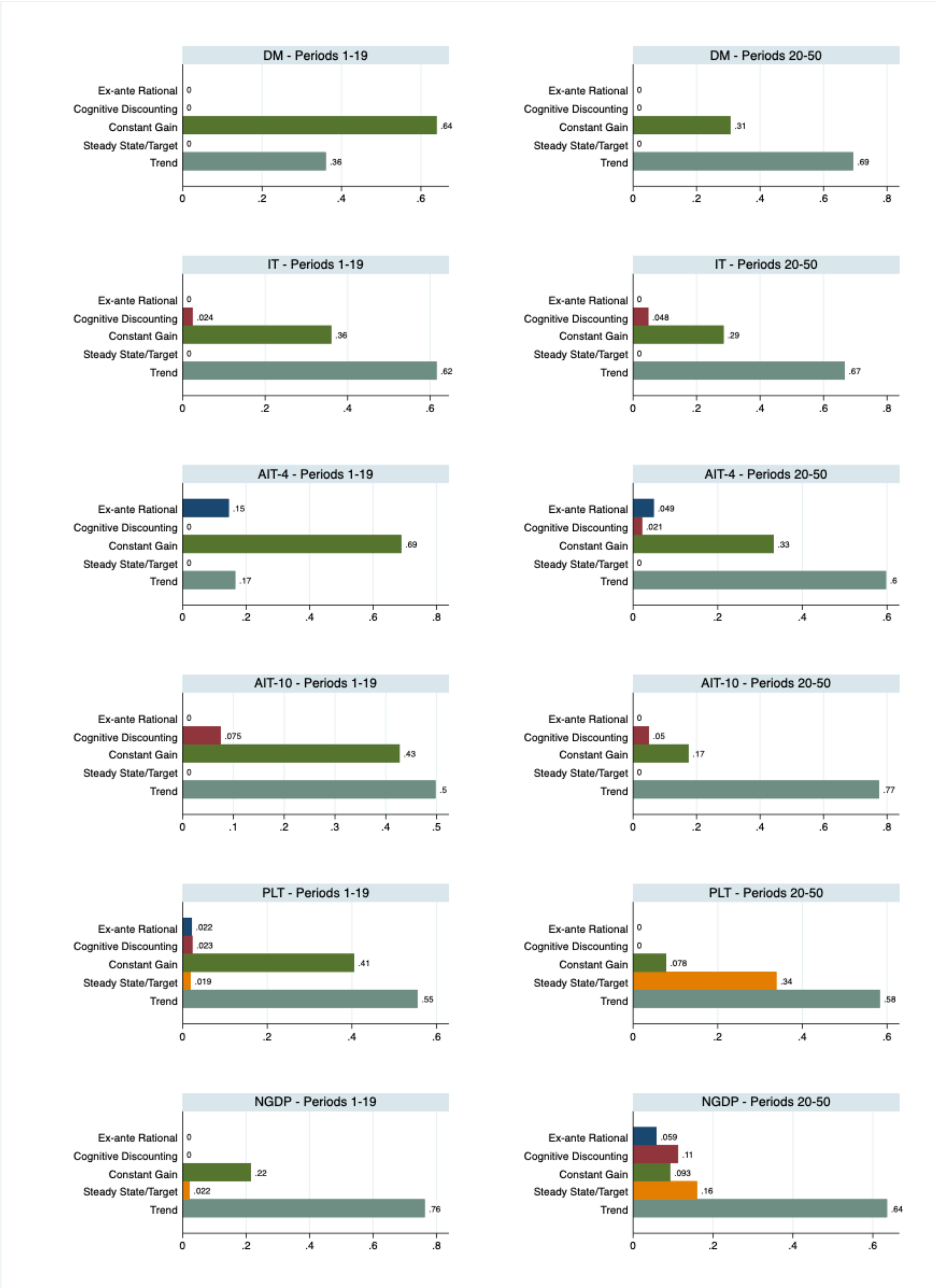
Also noteworthy is the large minority of participants in the NGDP and PLT treatments classified as using steady-state/target forecasting during the post-shock phase. A detailed analysis of individual forecasts shows that only a small minority of these participants actually forecast the steady-state values of zero. Rather, they submit forecasts closer to the steady-state value of zero than to values implied by other heuristics, but their values are negative. This behavior is certainly not rational, as an ex-ante rational agent would anticipate very high levels of inflation and output gap given the observed deflation and negative output gaps.

Figure D1: Distribution of forecasting heuristics for inflation forecasts, by treatment and phase



This figure presents the share of participants in each treatment and phase classified into a given heuristic.

Figure D2: Distribution of forecasting heuristics for output forecasts, by treatment and phase



This figure presents the share of participants in each treatment and phase classified into a given heuristic.

E. Additional experimental results

Table E1: Wilcoxon rank order test, statistical significance

Periods 1-19						
	NGDP	PLT	DM	IT	AIT-4	AIT-10
NGDP						
PLT	0.87					
DM	0.0163	0.037				
IT	0.0782	0.0542	0.8728			
AIT-4	0.0104	0.0103	0.1093	0.3367		
AIT-10	0.0039	0.0161	0.0547	0.4233	0.5218	
Periods 20-50						
	NGDP	PLT	DM	IT	AIT-4	AIT-10
NGDP						
PLT	0.4225					
DM	0.0039	0.0039				
IT	0.0039	0.0039	0.631			
AIT, short	0.0039	0.0039	0.7488	0.631		
AIT-4	0.0039	0.0039	0.0065	0.0065	0.0104	
AIT-10						
	NGDP	PLT	DM	IT	AIT-4	AIT-10
NGDP						
PLT	0.4225					
DM	0.0039	0.0039				
IT	0.0039	0.0039	0.8728			
AIT-4	0.0039	0.0039	0.3367	0.3367		
AIT-10	0.0039	0.0039	0.2002	0.2002	0.025	

Results from Wilcoxon rank order test based on the average losses from each of 6 sessions for all treatments. These results are for the hypothesis that losses in the treatments listed in the rows are equal to the losses in the treatments listed in the columns.

Table E2: Wilcoxon rank order test, probability that regimes in rows have lower losses than regimes in columns

Periods 1-19						
	NGDP	PLT	DM	IT	AIT-4	AIT-10
NGDP						
PLT	0.528					
DM	0.917	0.861				
IT	0.806	0.833	0.472			
AIT-4	0.944	0.944	0.778	0.667		
AIT-10	1	0.917	0.833	0.639	0.389	
Periods 20-50						
	NGDP	PLT	DM	IT	AIT-4	AIT-10
NGDP						
PLT	0.639					
DM	1	1				
IT	1	1	0.417			
AIT-4	1	1	0.444	0.417		
AIT-10	1	1	0.028	0.028	0.056	
All Periods						
	NGDP	PLT	DM	IT	AIT-4	AIT-10
NGDP						
PLT	0.639					
DM	1	1				
IT	1	1	0.472			
AIT-4	1	1	0.667	0.667		
AIT-10	1	1	0.278	0.278	0.111	

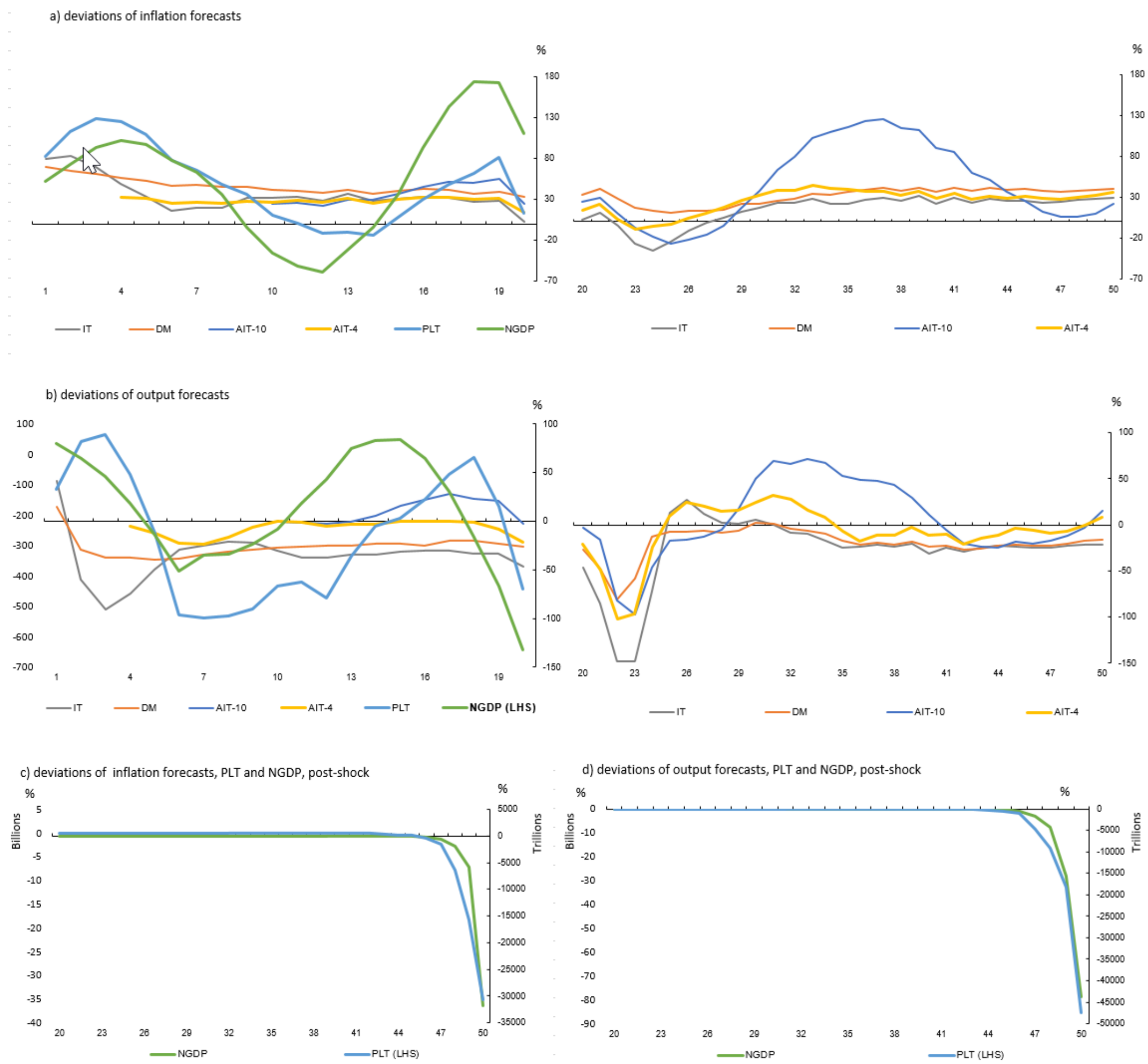
Results from Wilcoxon rank order tests based on the average losses from each of 6 sessions for all treatments. These results present the probability that losses in the treatments listed in the rows are less than the losses in the treatments listed in the columns, in accordance with hypothesis H1 presented in Section **??**.

Table E3: Summary statistics about participants' forecasts

	Deviation from REE		Interquartile Range	
	$E_{i,t}\pi_{t+1}$	$E_{i,t}x_{t+1}$	$E_{i,t}\pi_{t+1}$	$E_{i,t}x_{t+1}$
Periods 1-19				
AIT-4	29.12	-2.57	17.94	27.26
	(33.59)	(16.5t)	(30.57)	(37.76)
AIT-10	37.81	13.06	16.35	18.73
	(16.89)	(28.22)	(22.99)	(32.52)
DM	46.79	-26.53	12.15	18.76
	(17.67)	(17.33)	(18.05)	(36.20)
IT	37.11	-34.48	22.93	20.93
	(31.70)	(38.28)	(37.11)	(22.32)
NGDP	51.21	-153.5	33.97	59.54
	(84.32)	(190.2)	(28.91)	(38.90)
PLT	53.19	-16.07	49.98	63.06
	(66.38)	(87.53)	(49.4)	(51.08)
Periods 20-50				
DM	31.56	-20.21	8.483	9.906
	(13.17)	(17.91)	(17.58)	(12.65)
IT	14.89	-27.7	11.66	17.45
	(19.69)	(41.03)	(11.64)	(19.52)
AIT-4	25.71	-8.463	10.77	22.77
	(15.63)	(31.41)	(9.791)	(22.01)
AIT-10	44.55	6.18	14.74	22.6
	(49.99)	(43.14)	(10.49)	(17.28)
NGDP	-5.4e+14	-8.9e+14	1.24E+25	1.24E+25
	(5.3e+15)	(7.6e+15)	(2.14E+26)	(2.14E+26)
PLT	-1.6e+09	-3.4e+09	1.79E+09	2.75E+09
	(1.4e+10)	(3.2e+10)	(2.18E+10)	(2.23E+10)

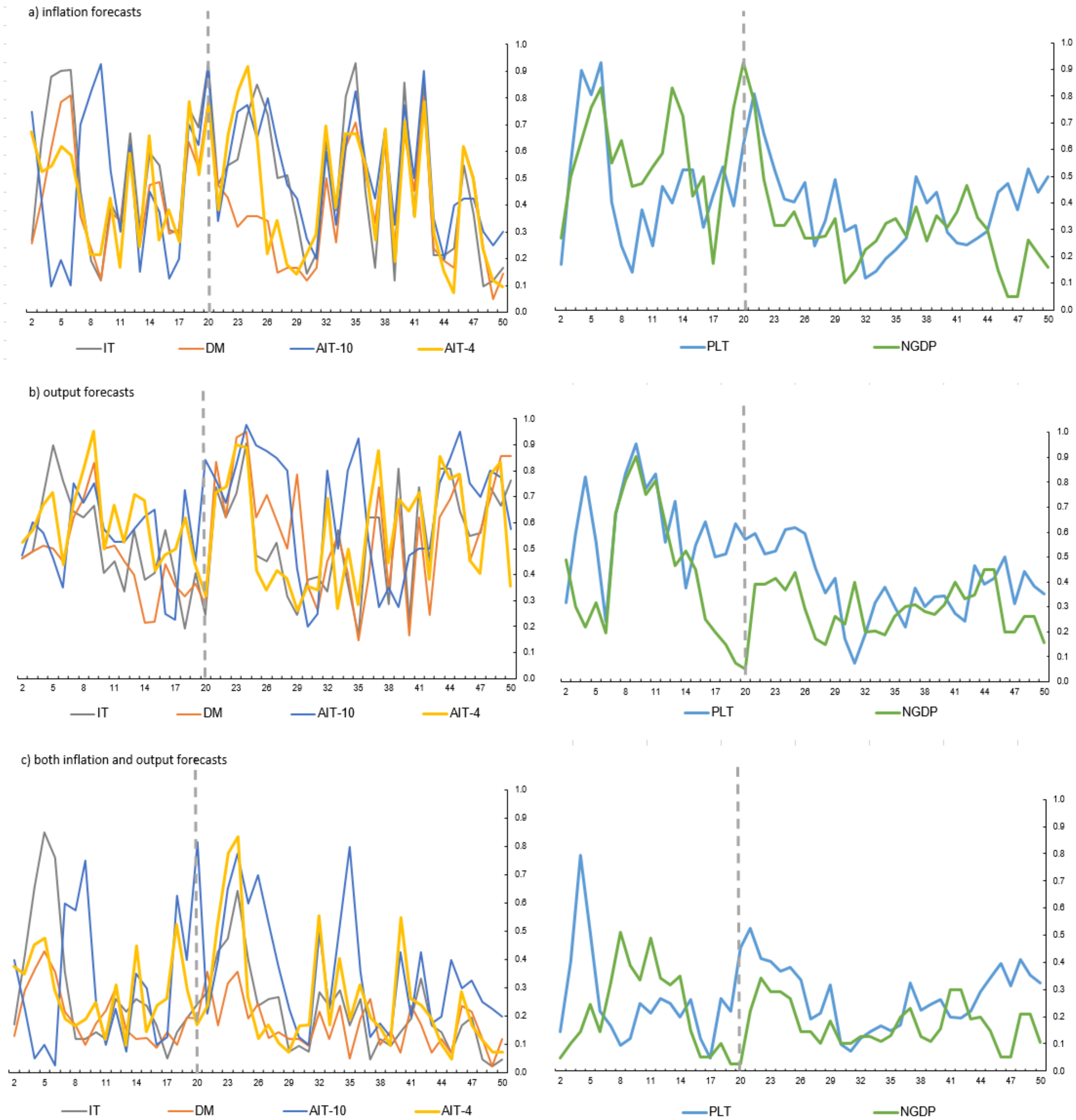
This table presents inflation and output gap forecast statistics. Columns (1) and (2) present the average across sessions of the median absolute forecast deviations from the REE solution. Columns (3) and (4) present the interquartile range of forecasts. Standard deviations are presented in parentheses.

Figure E1: Deviations of inflation and output forecasts from REE forecasts



This figure presents the median deviations of inflation and output forecasts from REE, averaged for each period across all sessions for each treatment.

Figure E2: Share of participants exhibiting basic rationality in inflation and output forecasts



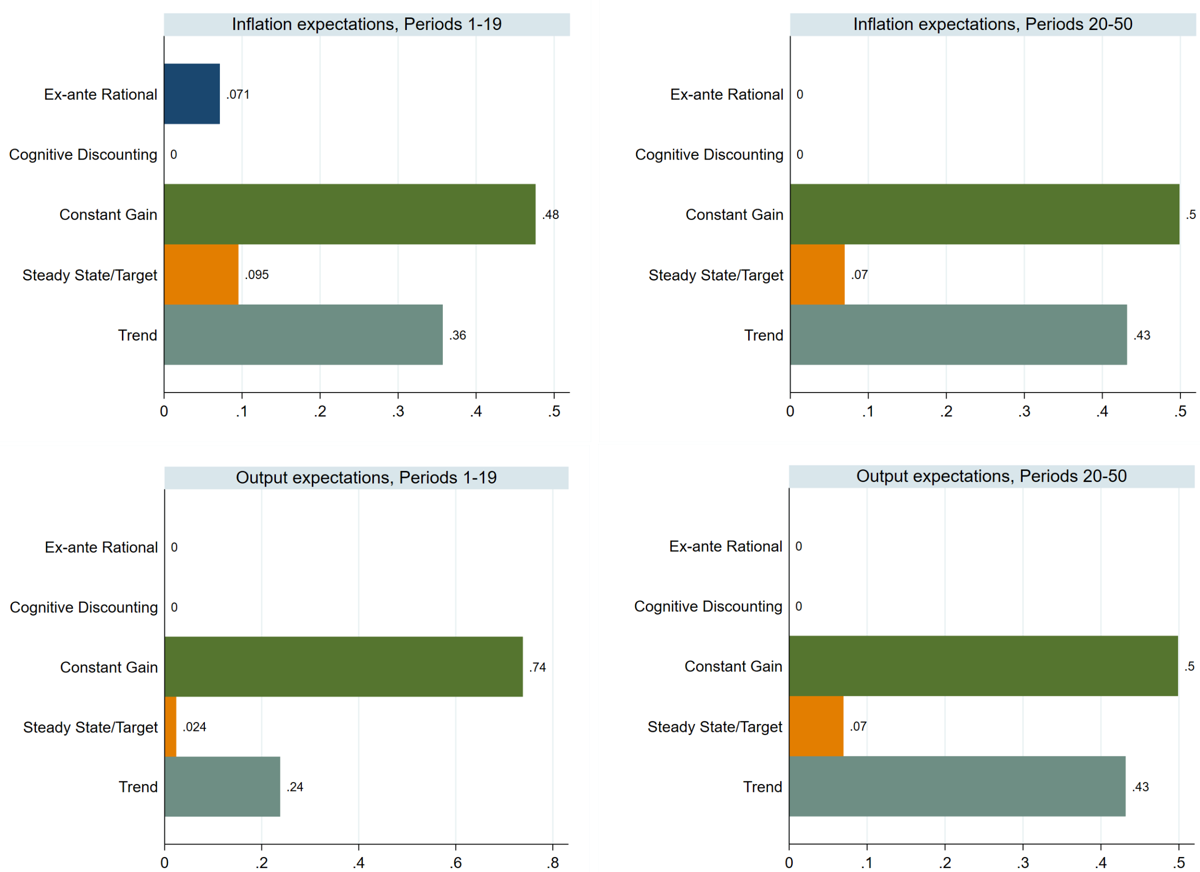
This figure presents the shares of participants whose forecasts satisfy the definition of basic rationality as forecasting in the correct direction. Panel (a) presents share for inflation forecasts, panel (b) – share for output forecasts, and panel (c) – share for both inflation and output forecasts satisfying basic rationality.

F. Additional PLT Comm experimental results

In this section, we present individual-level findings from the PLT Comm treatment.
Figure F1 presents the distribution of forecasting heuristics for inflation and output gap forecasts in PLT Comm.

Figure F2 shows the distribution of losses in the PLT Comm treatment relative to the original set of treatments. Losses in PLT Comm are significantly lower than losses in PLT in both the pre-shock and post-shock phases (Wilcoxon rank sum test, $N = 6$ for each treatment, $p < 0.001$). PLT Comm produces the lowest average losses in both phases of the experiment, excluding unstable session 5.

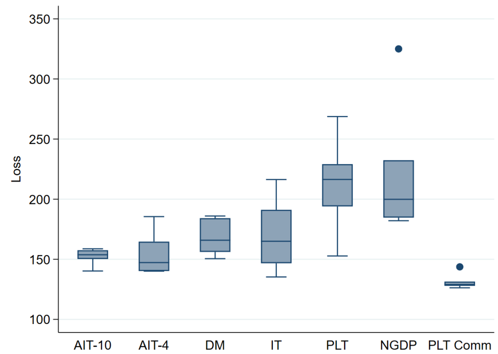
Figure F1: Distribution of forecasting heuristics for inflation and output gap forecasts in PLT Comm



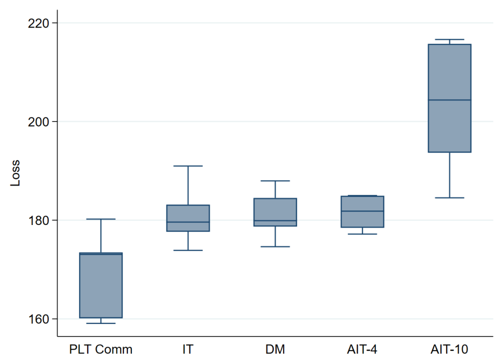
This figure presents the share of participants in each phase classified into a given heuristic.

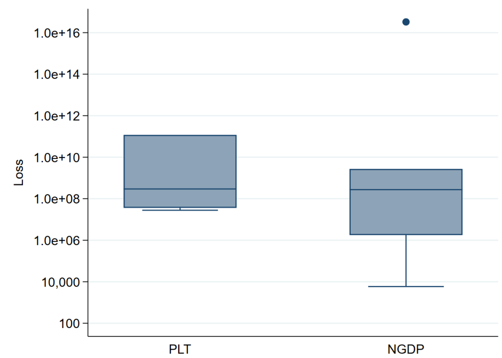
Figure F2: Distribution of losses including PLT Comm treatment

Periods 1-19



Periods 20-50





References

Mikhail Anufriev, Tiziana Assenza, Cars Hommes, and Domenico Massaro. Interest rate rules and macroeconomic stability under heterogeneous expectations. *Macroeconomic Dynamics*, 17(8):1574–1604, 2013.

Tiziana Assenza, Peter Heemeijer, Cars H Hommes, and Domenico Massaro. Managing self-organization of expectations through monetary policy: A macro experiment. *Journal of Monetary Economics*, 2019.

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

James Bullard and Kaushik Mitra. Learning about monetary policy rules. *Journal of Monetary Economics*, 49(6):1105–1129, 2002.

George Evans and Seppo Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton University Press, 2001.

Xavier Gabaix. A behavioral New Keynesian model. *American Economic Review*, 110(8):2271–2327, 2020.

Cars Hommes, Domenico Massaro, and Matthias Weber. Monetary policy under behavioral expectations: Theory and experiment. *European Economic Review*, 118:193–212, 2019.