History-dependent monetary policy – less is more*

Olena Kostyshyna
Bank of Canada

Luba Petersen
Simon Fraser University
NBER

Jing Yang
Bank of Canada

April 20, 2026

Abstract

History-dependent monetary policies (HDMP) can potentially address the challenges that standard monetary policy frameworks face, especially at the ELB. Using controlled laboratory experiments, we identify several vulnerabilities in HDMP. Our analysis of experimental data shows that a weakened expectations channel can undermine HDMP. Experimental participants' expectations feed off past data; they are not forward-looking and do not internalize the stabilization properties or future *make-up* strategies of HDMP. Conditional on these expectations, we highlight two risks associated with HDMP: inflation overshoots under average inflation targeting and deflationary spirals under price level targeting and nominal GDP level targeting. However, effective central bank communication through medium-term macroeconomic projections can improve the performance of price level targeting.

JEL classifications: C9, D84, E52, E58
Keywords: monetary policy, monetary policy frameworks, expectations, central bank communication, experimental macroeconomics

*Previous versions of this paper have been circulated under the titles "A horse race of monetary policy regimes: An experimental investigation" [Kostyshyna et al., 2022] and "Why Don't History-dependent Monetary Policies Work?"'[Kostyshyna et al., 2024]. We have benefited from useful discussions with Oleksiy Kryvtsov, Herbert Dawid, Chris Gibbs, Seppo Honkapohja, Cars Hommes, Chen Lian, Tomasz Makarewicz, Rhys Mendes, Emi Nakamura, Athanasios Orphanides, Isabelle Salle, Jón Steinsson, Alexander Ueberfeld, Yang Zhang and Joel Wagner, as well as seminar and conference participants at the Bank of Canada, International Monetary Fund, 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, and the NBER SI 2024 Monetary Economics. Finally, we thank Ekeme Afia, Michael Mirdamadi, Fatemeh Mokhtarzadeh, and Ryan Rholes for outstanding research assistance and Justin Baize for excellent programming support. This research was funded by the Bank of Canada. IRB approval was obtained from Simon Fraser University (#2012s0755). The views expressed herein are those of the authors and not necessarily those of the Bank of Canada.

† Kostyshyna and Yang: Bank of Canada, 234 Wellington Street Ottawa, ON, K1A 0G9; okostyshyna@bank-banque-canada.ca, jyang@bank-banque-canada.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 since the Global Financial Crisis and the pandemic, with frameworks emphasizing history-dependent mandates—such as price level targeting (PLT), nominal GDP level targeting, and average inflation targeting (AIT)—gaining prominence [Williams, 2017, Bernanke, 2017].These approaches potentially offer greater economic stabilization than traditional rate-targeting regimes. Their success hinges on the credibility of the regime and the strength of the expectations channel. The expectations channel operates through forward-looking expectations where agents internalize future make-up strategies inherent in history-dependent frameworks. However, the advantages of history-dependent regimes can diminish or even reverse if expectations are not fully rational or are backward-looking [Kryvtsov and Petersen, 2013]. Despite growing academic interest in these policies, real-world experience with them remains limited.[1] Moreover, evidence on public understanding of these frameworks remains scarce and mixed. Survey data suggests that U.S. households have difficulty grasping the implications of AIT for future inflation [Coibion et al., 2023], whereas German households appear to be more responsive to such strategies [Hoffmann et al., 2022].

To enhance our understanding of the performance of history-dependent monetary regimes in practice, and to address the issue of limited empirical data, we conduct a series of controlled macroeconomic laboratory experiments. These experiments systematically vary the degree of history-dependence in monetary policy across different treatments. At one end, we have policies with no history-dependence, such as inflation targeting (IT) and a dual mandate (DM). Moving toward greater history dependence, we include average inflation targeting mandates with short (4-quarter) and long-horizons (10-quarter) (AIT-4 and AIT-10), and at the greatest history-dependence, we examine price level targeting (PLT) and nominal-GDP level targeting (NGDP).

Given the central role of the expectations channel in the efficacy of history-dependent regimes, we study expectations of our experimental participants. In these experiments, participants are tasked with making incentivized forecasts of inflation and output during three distinct phases: (i) stable economic conditions, (ii) a temporary negative demand shock that drives the economy to its effective lower bound (ELB), and (iii) the subsequent recovery phase. Importantly, participants' expectations directly influence economic outcomes within

[1] Sweden briefly experimented with PLT in the 1930s [Berg and Jonung, 1999], and the U.S. implemented a flexible average inflation targeting framework in August 2020. To date, nominal GDP level targeting has yet to be applied in practice.

the experiment. The key advantage of our approach is that our findings are not contingent upon assumptions regarding any specific expectation formation model.

Our results demonstrate the potential vulnerabilities of history dependent monetary policies. Within our controlled laboratory setting, we observe the ability of monetary policy to stabilize the economy deteriorates with the degree of history-dependence, contrary to what is expected by theory. Economies with greater history-dependence in monetary policy exhibit significantly higher volatility in both inflation and output. The difference between the regimes is especially evident during brief but significant contractions in economic activity that drive the economies to their ELB. Frameworks with little to no history-dependence restore stability more quickly than level-targeting regimes such as PLT and NGDP. This outcome runs counter to the anticipated stabilization benefits that level-targeting regimes are expected to deliver during ELB episodes. Moreover, the extent of deterioration in the performance of history-dependent regimes far exceeds what is suggested by commonly used models incorporating bounded rationality.

Our findings about HDMP are strongly dependent on the expectations of participants given the design of our experiment. Within this context, we identify that the failure of the expectations channel plays a crucial role in undermining the performance of history dependent monetary policy frameworks. Our rich individual-level data enables us to uncover the sources of this failure. First, rather than relying on fundamentals and anticipated monetary policy actions, participants form their expectations based on recent historical experiences. Their expectations fail to internalize the stabilization properties of monetary policy, particularly the make-up strategies that are essential for history-dependent regimes. As a result, expectations are predominantly backward-looking, which weakens the forward-looking expectations channel central to history-dependent frameworks.

Second, the framing of PLT and NGDP in terms of levels of prices or nominal output adds cognitive complexity to the process of forming inflation expectations, as compared with monetary policies framed in terms of the inflation rate. We examine this effect by comparing forecast performance in level-targeting PLT and NGDP relative to rate-targeting regimes. Participants under the level targeting frameworks produce less accurate forecasts and exhibit greater disagreement in their expectations. Furthermore, we find some evidence of additional complexity in NGDP relative to PLT.

The evidence from our experiments points to two significant risks associated with monetary

policy frameworks that incorporate higher degrees of history dependence: inflation overshoot in AIT and deflationary spirals at ELB in PLT and NGDP. These risks arise from the interaction of history dependent, backward-looking monetary policy and agents' own backward-looking inflation expectations.

The first risk is inflation overshooting under AIT. The inertia in monetary policy meets inertia in expectations formation resulting in more volatility and higher persistence of inflation and output gap, with these effects intensifying as the AIT horizon lengthens. Following an ELB shock, the resulting deflation is deeper and more prolonged than theoretical models predict because the expectations channel does not provide the expected lift, while monetary policy is constrained at the ELB. During the subsequent recovery, the combination of "low-for-longer" policy and rising inflation expectations generates positive inflation overshooting.

The second risk is the development of deflationary spirals following the ELB shock under PLT or NGDP level targeting. Concerns about deflation have been prominent since the U.S. experienced a significant deflationary episode during the Great Depression. These concerns resurfaced after the Great Financial Crisis as monetary policy faced constraints imposed by the ELB. Our findings indicate that PLT and NGDP, with their higher degree of history dependence compared with AIT, results in higher volatility and stronger persistence in inflation during stable, pre-shock periods. Experimental economies with PLT or NGDP experience deflationary episodes even *before* the ELB shock occurs. The deflationary episode after the ELB shock further reinforces deflationary expectations, and deflationary expectations become entrenched. As monetary policy remains constrained at ELB, expectations fail to provide the anticipated uplift, sending the economy into a deflationary spiral. As the spiral unfolds, participants increasingly engage in trend-chasing behavior, which further entrenches deflationary expectations. While some experimental studies suggest that PLT may deliver stability [Hommes and Makarewicz, 2021a, Salle, 2023, Arifovic et al., 2023], our results indicate that PLT poses risks of deflationary spirals at the ELB, in line with Arifovic and Petersen [2017]. The risk of deflationary spirals in NGDP has not been documented yet (to the best of our knowledge).

Having observed the poor performance of PLT in our initial experiments, we conducted a follow-up treatment to explore whether supplementing policy with central bank communication of macroeconomic projections of inflation and output can enhance the effectiveness of this level targeting framework. Central bank communication has been shown to effectively guide macroeconomic expectations [Coibion et al., 2022, Cornand and M'baye, 2018]. By

introducing communication about future inflation rates and output, we aimed to reduce the complexity associated with the PLT regime [Mokhtarzadeh and Petersen, 2020]. The results from this follow-up experiment are very encouraging: five out of six sessions produced highly stable economic outcomes. PLT with communication proved even more effective at stabilizing the economy than rate-targeting frameworks like IT and DM. This improvement arises as participants coordinate their expectations around the central bank's projections, reducing their reliance on backward-looking dynamics and restoring the functioning of the expectation channel. Our findings highlight the critical role of central bank communication in implementing more complex mandates like PLT. Communicating inflation projections helps manage inflation expectations.

Our research contributes to the literature on monetary policy design and to an understanding of the effects of history-dependence in monetary policy, a concept widely acclaimed for its potential to stimulate economic activity when policy rates are at their effective lower bound [Svensson, 2002, Eggertsson and Woodford, 2003, Wolman, 2005]. While traditional models under rational expectations have underscored the potency of history-dependence through the expectations channel [Woodford, 2003, Vestin, 2006], recent work incorporating bounded rationality [Honkapohja and Mitra, 2014, 2020, Amano et al., 2020, Wagner et al., 2023] highlights important limitations. Our findings align with the bounded rationality literature, though we find even worse performance of history-dependent frameworks. Our work is also related to the growing experimental literature that examines various monetary policy regimes [Pfajfar and Žakelj, 2014, 2016, Hommes and Makarewicz, 2021a, Hommes et al., 2019b, Assenza et al., 2021, Mauersberger, 2021, Kryvtsov and Petersen, 2013, 2021, Kronick and Petersen, 2022].

Our work contributes to the literature by providing new evidence about the reasons *why* history-dependence in monetary policy does not work as intended – backward-looking expectations undermining the expectations channel and the cognitive complexity associated with framing PLT in terms of the price level. We have also identified inflation overshoot in AIT and deflationary spirals in PLT and NGDP at ELB as two main risks in these regimes. There is currently no evidence on nominal GDP (NGDP) level targeting in either the empirical or experimental literature. While deflationary spirals under PLT and IT have been documented [Arifovic and Petersen, 2017, Hommes et al., 2019a], the possibility of spirals under NGDP targeting is new.

We provide empirical support for models of expectation formation used in the behavioral

macroeconomics literature. Bounded rationality manifests in our experimental subjects through extrapolative expectations and limited responses to economic fundamentals. This reflects both an under-reaction to policy changes and shocks, and an over-reaction to past trends, resonating with diagnostic expectations theory [Bordalo et al., 2020]. The heterogeneity in expectation formation supports the use of heterogeneous agent models in macroeconomic analysis [Brock and Hommes, 1997]. The majority of our participants are not forward-looking, instead relying on past experiences to form expectations [Malmendier and Nagel, 2016]. However, we also observe spikes in attention and rationality consistent with models of rational inattention [Sims, 2010, Mackowiak and Wiederholt, 2009]. Additionally, some participants anticipate policy directions but underestimate necessary adjustments, supporting bounded rationality models based on cognitive discounting and limited common knowledge [Gabaix, 2020, Angeletos and Lian, 2018].

Our study highlights the critical role of effective central bank communication in implementing history-dependent policies like price level targeting (PLT). Communication is a key tool in central banking [Blinder et al., 2022, Haldane and McMahon, 2018, Levin, 2014, Ehrmann et al., 2022]. We show that PLT and similar frameworks can be more effective when paired with well-designed communication strategies. Communicating inflation projections helps anchor expectations and mitigates the complexity of level-targeting frameworks.

2 Experimental Design

In this section we describe the implementation of the experimental macroeconomy, the different monetary policy frameworks studied, and the laboratory implementation.

2.1 Data-generating process

Our experimental environment is based on a standard three-equation New Keynesian (NK) model commonly used for monetary policy analysis in theory and in the lab [Woodford, 2003, Adam, 2007, Pfajfar and Žakelj, 2014, Assenza et al., 2021]. The economy evolves according to the following system of equations:

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

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

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

Here, π_t is inflation, x_t is the output gap, i_t is the nominal interest rate, and r_t^n is the natural rate of interest that keeps the economy at full employment with stable inflation. The natural rate follows an AR(1) process with demand shocks ϵ_t. Parameters are calibrated to Canadian data as in Dorich et al. [2021], Kryvtsov and Petersen [2021]: $\beta = 0.994$, $\sigma = 1$, $\rho = 0.8$, $\sigma_{rn} = 0.005$, $\kappa = 0.125$, $\pi^* = 0$, $x^* = 0$, $r_t^{n*} = 0$, $\bar{r} = i^* = 60$.

In the New Keynesian model, we work with the standard log-linearized representation around a zero-inflation steady state. This normalization sets the inflation target to 0% in the linearized system, which is conventional in the theoretical literature and simplifies the experimental environment. Using a 0% target also avoids anchoring participants toward a specific positive number (such as 2%), consistent with the experimental design rationale in Cornand and M'baye [2018]. Importantly, this normalization has no effect on the qualitative or comparative performance of the monetary policy regimes.

2.2 Monetary policy frameworks

All policy rules adjust the nominal interest rate based on deviations of key variables from their targets. The six monetary policy frameworks we study differ in targets, weights, or history dependence.

In both inflation-targeting (IT) and dual mandate (DM) frameworks , the central bank sets the nominal interest according to the following general policy rule:

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

This rule aims to minimize deviations of current inflation and the output gap from their target levels of zero. In other words, these rules exhibit no history-dependence. The parameters ϕ_π and ϕ_x govern the reactions of the central bank to deviations of inflation and output gap from their targeted rates. The key distinction between IT and DM lies in the weight given to the output gap, In the DM framework, the weight on the output gap, ϕ_x, is equal to that on inflation, ϕ_π, and is generally higher. In the IT regime, $\phi_\pi = 5.5$ and $\phi_x = 3.0$.

In contrast, under the DM regime, both are set at $\phi_\pi = \phi_x = 4.5$.

In average inflation targeting (AIT) regimes, the central bank adjusts the nominal interest rate to minimize deviations of inflation from its target based on the recent average inflation rate. The central bank also places some weight on the current output gap when making its policy decisions.The degree of history-dependence is determined by how far back the central bank looks when calculating average inflation. We examine two horizons for average inflation: a short horizon of 4 quarters and a long horizon of 10 quarters. Our reasons for studying two horizons in AIT is twofold. First, it allows us to evaluate how different levels of history-dependence affect AIT performance, guiding policymakers on optimal horizon selection. Second, theory predicts that as the averaging horizon extends indefinitely, AIT approximately price level targeting (PLT). Thus, a longer AIT horizon may achieve outcomes similar to PLT, without its practical challenges [Amano et al., 2020]. We intend to compare the performance of PLT and AIT-10. The two AIT policy rules – AIT-4 and AIT-10 – are defined as follows:

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

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

where the policy coefficients are the same as in IT: $\phi_\pi = 5.5$, $\phi_x = 3.0$.

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

Policy coefficients are chosen to minimize the following central bank loss function: Dorich et al. [2021]:

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

The loss function in Equation 9 is a second-order approximation to a social welfare function derived from the expected utility of a representative household [Woodford, 2003, Rotemberg and Woodford, 1997]. As such, it reflects structural welfare-relevant objectives—stabilization of inflation, the output gap, and interest rate variability—rather than the objectives embedded in any particular policy regime.

For this reason, the welfare-based loss function does not change across monetary policy frameworks. Even in an IT regime, the central bank's welfare objective includes output gap stabilization (through its effect on household consumption–leisure tradeoffs) and interest rate smoothing (to limit unnecessary volatility and financial dislocations), consistent with Rudebusch and Svensson [1999].

The policy rules in each regime are therefore evaluated against the same welfare metric, allowing for a meaningful comparison of their stabilization properties. To ensure comparability, the coefficients in each rule are calibrated by minimizing this welfare-based loss function. This places all regimes on equal footing and aligns with the theoretical "horse race" methodology in Dorich et al. [2021].

2.3 Experimental implementation

The design of our experiments is based on the New Keynesian learning-to-forecast framework [Adam, 2007, Arifovic and Petersen, 2017, Hommes et al., 2019a]. In each period, subject j submitted forecasts for next period's inflation and the output gap – $E_{j,t}\pi_{t+1}$ and $E_{j,t}x_{t+1}$. Realized values for π_t, x_t, and i_t were determined using the current period's shock ϵ_t and the median expected inflation and output gap for $t+1$ based on Equation 1-Equation 3 with one of the policy rules in Equation 4-Equation 8. The median expectations are used to reduce the influence of outliers and stabilize aggregate dynamics.

In our experimental design, participants report one-period-ahead expectations for inflation and output. This choice is consistent with the structure of the log-linear New Keynesian

model that underlies our data-generating process, where only next-period expectations enter the linearized Euler equation and New Keynesian Phillips Curve. In this environment, one-step-ahead expectations are the relevant variables for equilibrium dynamics, unlike in frameworks where non-linear aggregation yields long-horizon forecast terms [Preston, 2005]. There are several advantages to focusing on short-term expectations. First, the theoretical gains from history-dependent mandates derive primarily from stabilizing short-run expectations [Walsh, 1998], not from altering agents' long-horizon beliefs. Second, eliciting only next-period forecasts reduces the cognitive burden on participants and allows us to collect long time series of data -— critical for evaluating history-dependent rules -—without overloading subjects.

Finally, eliciting multi-period expectations risks introducing forward-looking behavior that is not empirically realistic. Experimental evidence shows that eliciting long-horizon forecasts induces substantially more forward-looking expectations [Evans et al., 2022]. In contrast, survey evidence indicates that short-term inflation expectations, not long-run expectations, are most predictive of price-setting, wage-setting, and spending behavior [Abberger et al., 2024, Gautier et al., 2025, Coibion et al., 2018, 2020, 2022]. Thus, relying on one-period-ahead forecasts preserves empirical realism and avoids unintentionally biasing participants toward rational or near-rational long-horizon forecasting.

Participants received detailed information about the economy's data-generating process in the instructions, including clear descriptions and explicit equations of how the central bank sets monetary policy and its impact on the economy. We also explained how their forecasts would translate into points and payoffs at the end of the experiment. The experimental instructions are provided in Online Appendix A. Participants did not have information about each others' expectations as is standard in learning-to-forecast experiments, and is consistent with empirical evidence on the formation of expectations in the literature (Coibion et al. [2022]). [2]This design decision also avoids strategic coordination on each other's forecasts and on the steady state.

Treatments We implemented six treatments: each treatment corresponds to one of monetary policy regimes presented in Equation 4-Equation 8. Thus, our treatments are IT, DM, AIT-4, AIT-10, PLT, and NGDP.

[2]Survey respondents revise their views about inflation in response to the provision of publicly available information about forecasts of the Federal Reserve indicating that they are not aware of such information.

2.4 Procedures

Each treatment was run in six independent sessions with seven inexperienced participants per session.[3] The experiment lasted 50 periods, with a pre-drawn exogenous shock sequence creating two distinct phases: a stable initial phase (periods 1–19) and a post-shock phase (periods 20–50) including an effective lower bound (ELB) episode. In period 20, a large negative demand shock brought the economy to the ELB, quickly dissipating to a steady-state level of zero by period 23. The remainder of the post-shock phase, lasting 27 periods, allowed us to study recovery and responses across monetary policy mandates following ELB episodes. This design enables us to test the stabilization properties of monetary policy regimes during both stable and volatile periods, including those at the ELB.

Participants' earnings during the experiment were based on the accuracy of their inflation and output gap forecasts. Subject j's points in period t depended on the absolute distance between their forecasts made in period $t-1$ and realized inflation and output in period t as in [Kryvtsov and Petersen, 2021]:

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

Participants' total payoffs over all the forecasting periods were converted to Canadian dollars at an exchange rate of \$1 per point. If participants earned below a total of \$10, a flat \$10 payment was awarded to these participants.

Participants viewed real-time information about their experimental economy on a computer screen, as shown in Figure B1 in Online Appendix B. The screen displayed four charts tracking shocks and interest rates, inflation (including the target and the participant's forecast), output (with the participant's forecast), nominal GDP and price levels, and central bank targets. Targets were presented as horizontal lines extending the entire duration of the experiment: zero for inflation and output gap, and 1000 for price-level and nominal output targets.

On the left side of the screen, participants had two input windows to submit their one-

[3]The size of the group can play an important role in driving aggregate dynamics, especially in settings with a high degree of strategic complementarities. Group size can influence aggregate dynamics, especially where strategic complementarities are strong. Hommes et al. [2021] show that in asset pricing experiments with positive feedback (as in our environment at the ELB), increasing the group size from six participants to 90 to 100 participants does not significantly change pricing. In other work, Bao et al. [2020] find that increasing from six to 31-32 participants can speed up extrapolative pricing and deviations from equilibrium predictions. Our decision to keep the group size relatively small likely reduced the coordination on extrapolative expectation models.

period-ahead forecasts for inflation and output in basis points. They had 75 seconds to enter forecasts during the first 10 periods and 50 seconds for the remaining periods. If a forecast was not submitted on time, the experiment proceeded to the next round, with zero points awarded for the missed entry. The median forecast was then calculated from submitted forecasts. Participants could enter any number positive, negative, or zero without upper or lower bounds on their forecasts.

The experiments were conducted online via Zoom with 252 undergraduate students from Simon Fraser University in Canada and Texas A&M University in the United States, spanning May to July 2020 and May to June 2021. Online sessions were required due to pandemic-related health restrictions and lab closures.

Participants were recruited through the SONA and ORSEE systems [Greiner, 2015], with sessions for each treatment equally divided between the two institutions. Each session lasted about two hours, including 40 minutes for instructions and Q&A and 10 minutes for four practice rounds with the experimental interface. A web-hosted PDF of the instructions was shared via Zoom at the start of each session and accessible throughout the experiment. Participants could ask questions throughout the session.

Participants at both laboratories received a $7 (CAD or USD) show-up fee plus performance-based earnings designed to average approximately $25. Payments were made by e-Transfer at SFU and via Venmo at TAMU. Ex post, average total earnings were $28.08 at SFU (range: $12–$35) and $26.67 at TAMU (range: $10–$35), inclusive of the show-up fee.[4] These differences are not statistically significant at the 10% level, and we do not find any systematic behavioral differences across the two samples.

3 Performance under rational and non-rational expectations

Owing to the distinct nature of each policy mandate, the policy regimes are expected to produce noticeably distinct aggregate dynamics. Figure 1 presents the rational expectation (RE) equilibrium solutions for inflation, output gap, and the nominal interest rate under each monetary policy mandate associated with our pre-selected shock sequence. As has been established in the literature, inflation deviates significantly less from the steady state

[4] Any participant earning below $10, inclusive of the show-up fee, received a flat $10 payment.

under history-dependent frameworks like NGDP and PLT than under IT.

We derive theoretical predictions regarding the stabilization performance of the different monetary policy regimes based on the loss function specified in Equation 9. Our model is simulated with each monetary policy regime under rational expectations, applying the same sequence of demand shocks as in the experiment. For each regime, we calculate the average total loss as the square root of total loss (Equation 9) divided by 50 periods. The results are summarized in Table 1. This table shows that more history dependent monetary policy regimes are more stabilizing, i.e. losses are lower.

Given our simulated sequence of shocks, the overall total loss (as well as the loss associated with inflation) is predicted to be the 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 all these six regimes are quite close under rational expectations.

The lower losses under PLT and NGDP targeting are a direct consequence of the expectations channel under rational expectations. Because both frameworks are history-dependent, the central bank's credible make-up strategy yields strongly anchored expectations: agents internalize that past deviations from target will be offset by future policy actions. This commitment feature stabilizes expectations of future inflation (in PLT) and future nominal income (in NGDP), which in turn dampens the response of current inflation and output to shocks.

Under the common quadratic (Equation 9), lower volatility in inflation and the output gap maps directly into lower losses. Rate-targeting rules such as IT and DM lack this commitment mechanism, so *rational* expectations respond more strongly to shocks, leading to more persistent deviations of inflation and output. This explains why the ordering of losses in Table 1 places PLT and NGDP ahead of IT and DM.

These mechanisms are well established in the theoretical literature: price-level targeting anchors inflation expectations and improves stabilization by incorporating bygones into policy [Vestin, 2006, Woodford, 2003, Svensson, 1999, Ambler, 2009]. Similarly, NGDP-level targeting improves stability by anchoring expectations of future nominal income paths [Garín et al., 2015, McCallum, 2015, Sumner, 2014].

Simulations with non-rational expectations

Experimental studies of monetary policy regimes show that participants' expectations are mostly non-rational [Anufriev et al., 2013, Assenza et al., 2021, Hommes and Makarewicz, 2021b]. Motivated by this evidence, we consider several expectation models commonly identified in surveys and laboratory experiments and applied in the estimation of DSGE models [Milani, 2012]: cognitive discounting [Gabaix, 2020], constant-gain learning [Branch and Evans, 2006], anchoring on targets [Coibion et al., 2018], extrapolative trend-chasing [Frankel and Froot, 1990, Bordalo et al., 2020] and naive expectations (details in Table C1, Appendix C). Figures C1–C3 in Appendix C present the results across all monetary policy regimes, highlighting the impact of different expectations, including rational expectations, on economic stability.

Cognitive discounting and steady-state expectations yield the most stable economic outcomes in inflation and output, which are more stable than in the economy with rational expectations. This happens as cognitive discounting expectations are a fraction α (ranging from 0 to 1, and 0.5 in our simulation) of rational expectations making cognitive discounting expectations less volatile than rational. And agents with steady state expectations always expect economy to be at the steady state, which stabilizes the economy. The dynamics in inflation and output in the simulations with constant gain learning are close to rational expectations outcomes, with different volatility depending on the monetary policy regime. In contrast, naive expectations and trend-chasing expectations generate the highest volatility in both inflation and output compared with the economies with all other expectations. The volatility of the economies with naive and trend-chasing expectations increases with the history dependence of monetary policy, with the level-targeting regimes delivering the least stabilization. And PLT with naive and trend chasing expectations results in the most volatile inflation and output among all regimes.

In Figure 2 we evaluate the loss function (Equation 9) for the model with rational expectations (RE) and for the models with other expectations. This figure illustrates an interesting observation – the ranking of the monetary policy regimes changes depending on the type of expectations. Under cognitive discounting and steady state expectations, all monetary policy regimes fare similarly well and better than under rational expectations. And while regimes with higher history dependence are more stabilizing under rational expectations (first set of bars in Figure 2, and also Table 1), the ability of history dependent regimes to deliver economic stabilization deteriorates when expectations are backward-looking. Under naive and trend-chasing expectations, the ranking of regimes reverses relative to that with

rational expectations, with losses increasing in the following order: IT, DM, AIT-4, AIT-10, NGDP, PLT. The reversal is strongest under naive and trend-chasing expectations. When agents forecast future inflation and output based on last-period values (naive expectations) or extrapolate a past trend into the future (trend chasing), level-targeting regimes become less effective at stabilizing the economy. Note that these simulations are performed with a relatively small value for the trend-chasing parameter ($\tau = 0.1$), and additional simulations indicate that economic volatility increases when τ increases. Figures C4 and C5 in Appendix C evaluate the robustness of the rankings to alternative policy rule parameterizations. The baseline ordering is preserved when policy responsiveness is reduced by 25 percent, but changes when responsiveness is increased by 25 percent. This asymmetry reflects that stronger responses yield larger gains under rate-targeting regimes, which are further from optimal, while level-targeting frameworks (e.g., PLT and NGDP targeting) are closer to optimal and thus offer less scope for improvement.

4 Aggregate dynamics and policy performance

We present the time series of inflation, output, and interest rate across six sessions for each of our six treatments: DM and IT (Figure 5), AIT-4 and AIT-10 (Figure 6), PLT (Figure 7), and NGDP (Figure 8). For reference, the rational expectations (RE) solutions is shown in red.

Inflation and output dynamics in IT and DM are remarkably consistent across sessions and closely align with the RE model's predictions. This consistency reflects a shared understanding among median forecasters of how aggregate shocks and monetary policy affect the economy. Both treatments exhibit stable inflation and output during early periods (1–19), followed by a brief episode at the ELB during the large demand shock. Recovery from the ELB occurs within 3-4 periods in DM and 4-5 periods in IT, slightly slower than the RE solution's 2-period recovery.

The stability of IT and DM in our experiments partly stems from the relatively high policy responsiveness to both output and inflation. The coefficients in our DM treatment are the strongest considered in the literature. For instance, 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 as high as $\phi_x = 1$. Hommes et al. [2019b] demonstrate that when expectations are backward-looking, a strong response to output is crucial for sta-

bilizing both output and inflation.

The dynamics of inflation and output show that AIT-4 is capable of stabilizing the economy similarly to IT and DM, while AIT-10 delivers less stability and displays less consistency across sessions than IT, DM, and AIT-4. Inflation in AIT-10 overshoots more than in AIT-4, resulting in greater volatility and persistence. Following the ELB shock, deflation is stronger and more prolonged as the degree of history dependence increases, with AIT-10 experiencing longer deflation than AIT-4 or IT. During recovery from the ELB episode, AIT-10 exhibits a significant positive inflation overshoot, with inflation rising much higher and remaining elevated for longer than in AIT-4, unlike IT, which shows no such overshoot. The "low for longer" mechanism inherent in AIT with extended history dependence contributes to this inflation overshoot. We will explore the role of the expectations channel in contributing to these dynamics in the next section.

Increasing history dependence to its maximum under PLT results in even less stability than in AIT with long horizon . Even during the initially stable periods, economies under PLT exhibit greater volatility in inflation and output compared with IT and AIT with both short or long horizons. Some PLT economies experienced deflationary episodes even before the large negative demand shock hit in period 20. Following the shock, all experimental sessions under PLT spiraled into deflationary episodes as monetary policy remained constrained by the ELB. Only one of the six sessions showed any attempt at recovery, though it was ultimately unsuccessful.

NGDP, another level-targeting regime, produces outcomes as unstable as those under PLT. Inflation and output exhibit high volatility and persistence, with deflationary episodes even before the ELB shock. After the ELB shock, all NGDP sessions unravel into spiraling deflation and declining output. Although one session shows a brief lift-off from the ELB, it eventually falls back into an expectations-driven recession with monetary policy constrained at the ELB. Similar deflationary spirals under the ELB have been observed in experimental studies of IT and PLT [Arifovic and Petersen, 2017] and IT [Hommes et al., 2019a, Assenza et al., 2021].

We evaluate the stabilization performance of the six monetary policy regimes in terms of their ability to stabilize inflation, output, and interest rate using the loss function (Equation 9). Table 2 and Figure 9 present the losses by treatment and phase, revealing a distinct ranking between rate-targeting and level-targeting regimes. The rankings differ before and

after the ELB shock. During the stable pre-shock period (periods 1-19), AIT-4 and AIT-10 outperforms DM and IT, followed by NGDP and PLT. After the ELB shock, the performance of AIT-4 and AIT-10 deteriorates, with DM and IT surpassing them in terms of stabilization, while NGDP and PLT perform the worst. Overall, stabilization declines with increasing history dependence after the ELB shock. Wilcoxon rank-order tests (Table 3) show that losses in DM are statistically significantly different from those in AIT-10, AIT-4, NGDP, and PLT at the 1% and 5% levels. Similarly, losses in AIT-4 and AIT-10 differ significantly from those in NGDP and PLT at the 1% and 5% levels. However, differences in losses within the rate-targeting treatments and within the level-targeting treatments are not statistically significant.

The evidence from our experiments leads us to reject the rational expectations prediction of the relatively superior performance of history-dependent monetary policies, outlined in Section 3. In our experiments, regimes responding to inflation and output, such as IT, DM, AIT-4, and AIT-10, outperform the most history-dependent level-targeting regimes, PLT and NGDP. AIT is more effective with a shorter horizon: AIT-4 outperforms AIT-10. *Less* history dependence *is more* stabilizing. This aligns with the models by Amano et al. [2020] and Wagner et al. [2023], who show that AIT with shorter horizons can better stabilize economies with boundedly rational agents. We further examine participants' expectations in the following section.

Our experimental results highlight two key risks regarding history-dependent monetary policy frameworks. First, AIT with longer history dependence can lead to inflation overshooting. And second, deflationary spirals emerge in PLT and NGDP when monetary policy is constrained by the ELB. These findings arise in experimental economies based on a standard New Keynesian model, differing only in that expectations are formed by experimental participants. Next, we analyze how participants form expectations and how the interaction between history-dependent monetary policy and these expectations contributes to the observed instability.

5 The role expectations in history-dependent monetary policy

5.1 Conditional responses of expectations to demand shocks

We begin by examining how participants' expectations respond to shocks in our experimental economies. A key advantage of our experimental framework is that the exogenous demand shock process, r_t^n, is observed by the experimenter. This enables us to estimate the responses of endogenous variables as functions of the sequences of ϵ_t. Let $X_{k,t}$ denote individual i's forecast in period t. Using the local projections method [Jordà, 2005], we estimate the following empirical specification for the change in $X_{i,t}$ over h periods for each treatment:

$$X_{i,t+h} - X_{i,t-1} = c^h + \sum_{l=0}^{L} \beta_l^h + \epsilon_{t-l} + \sum_{n=0}^{N} \delta_n^h X_{i,t-n} + D_s + S_i + error_{kst}^h. \tag{11}$$

The specification in Equation 11 conditions on the history of shocks ϵ_{t-l} and lags of endogenous aggregate variables $X_{i,t-n} \in \{x_{i,t-n}, \pi_{i,t-1}, i_{i,t-1}\}$ where $L = N = 2$. We estimate the pre-shock and post-shock periods separately using panel regressions that include session dummies, D_s, and subject fixed effects S_i. Standard errors for the estimated coefficients are clustered at the session level. The estimated responses of forecasts to a +1% aggregate demand shock are presented in Figure 10, while estimated responses of aggregate variables can be found in the Online Appendix D. Estimations are conducted separately for pre-shock (blue line) and post-shock (green line) periods to illustrate the impact of the ELB episodes and learning during the experimental session.[5] Predictions of the rational expectations model are included as red dashed lines for comparison.

In the rational expectations model, a positive and persistent demand shock increases expectations of output gap and inflation on impact (red lines in Figure 10). This response drives an increase in aggregate variables, which are subsequently stabilized through both the direct effects of monetary policy and the expectations channel. While history-dependent monetary policy may not reduce the immediate volatility of macroeconomic variables, it accelerates the economy's return to steady state, ultimately resulting in smaller deviations from the steady state over time.

Based on Figure 10, we observe several notable differences between the expectations of our

[5]Only pre-shock estimations are plotted for the level-targeting mandates as post-shock dynamics are explosive and highly noisy.

experimental participants and those under rational expectations. Participants' expectations (blue and green lines) respond differently to aggregate demand shocks than rational expectations (red lines), both before and after the ELB shock.

The responsiveness of expectations to shocks varies across monetary policy regimes with different levels of history dependence. During the pre-shock period, participants' expectations for inflation and output are generally less responsive to demand shocks than rational expectations in rate-targeting regimes such as DM, AIT-4, and AIT-10. This suggests that participants may not fully grasp the model dynamics, including the role of demand shocks and monetary policy responses. However, in level-targeting regimes like PLT and NGDP, inflation expectations overreact to demand shocks, consistent with the higher volatility in inflation and output observed in these regimes, as discussed in Section 4. After the ELB shock, expectations in IT, DM, and AIT with a short horizon continue to be less responsive than rational expectations. In contrast, AIT with a long horizon (AIT-10) exhibits overreaction in inflation expectations, consistent with the inflation overshooting observed during recovery from the ELB episode, as shown in Figure 6.

And finally, most of the impulse response functions for inflation and output expectations exhibit oscillatory behaviour indicative of backward-looking expectations [Pfajfar and Žakelj, 2014, Kryvtsov and Petersen, 2013]. Experimental expectations do not respond monotonically to demand shock as rational expectations do, suggesting important deviations from rationality.

5.2 The direction of expectations

Next, we examine whether participants form expectations in a manner consistent with the underlying structure of the economy and monetary policy framework, and, in particular, whether they forecast in a policy-consistent direction. Using participant-level data, we classify a participant i's forecast in each period t as policy-consistent if it changes in the same direction as the rational forecast for inflation or output relative to the previous period. The shares of policy-consistent inflation and output forecasts for each treatment are presented in Table 4.

During the pre-shock phase, 50–60% of participants display policy-consistent expectations for either inflation or output, with approximately 30% exhibiting policy consistency in both forecasts. There is little difference in the prevalence of policy-consistency across different

monetary policy regimes. Following the negative demand shock in periods 20 and 21, the share of policy-consistent inflation forecasts increases sharply. This suggests that large shocks temporarily reduce participants' inattention, consistent with the theory of rational inattention [Sims, 2010]. The observed spike in attention during the transition to the ELB aligns with Mackowiak and Wiederholt [2009], who argue that increased uncertainty enhances rationally inattentive agents' focus on shocks.

As the shock dissipates and fundamentals revert to the steady state, the share of policy-consistent expectations declines significantly across most regimes. This decline is particularly pronounced under level-targeting frameworks such as NGDP and PLT, where only 18% of NGDP participants and 26% of PLT participants form policy-consistent forecasts in the post-shock phase. We attribute this larger drop in the policy consistency to the greater complexity of level-targeting regimes compared with rate-targeting frameworks. We discuss this in further detail in subsection 5.4.

We next assess how the likelihood of forming policy-consistent inflation expectations depends on the central bank's recent performance. Performance is measured as the absolute deviation of the targeted variable from its target (e.g., inflation in IT, DM, AIT-4, and AIT-10; price level in PLT; and nominal GDP level in NGDP). Participants' previous policy-consistency is also included as a control.

During the pre-shock period, the likelihood of participants forecasting inflation in a policy-consistent direction increases with larger deviations of inflation from its target across all rate-targeting regimes (IT, DM, AIT-4, and AIT-10), suggesting a general understanding of monetary policy mechanisms (Table 5, panel A). However, this relationship disappears after the ELB shock (Table 5, panel B). In level-targeting regimes (PLT and NGDP), deviations from the target do not influence policy-consistent expectation formation in either the pre-shock or post-shock periods. This indicates that participants may struggle to translate deviations in the price level or nominal GDP level into inflation forecasts aligned with monetary policy intentions.

Across all regimes, there is notable persistence in forming policy-consistent expectations, indicating behavioral inertia rather than responsiveness to deviations from target. This inertia intensifies in the post-shock period, suggesting that prior belief formation plays a larger role in shaping expectations than contemporaneous central bank performance.

Even among the limited share of participants whose forecasts are policy-consistent, expectations do not adjust sufficiently to reflect current economic fundamentals. While these forecasts align more closely with the rational expectations benchmark than policy-inconsistent ones, they still fall short of what would be expected if participants had fully internalized the stabilization properties of monetary policy (Figure 11 and Figure 12).

Taken together, these findings indicate that the expectations channel of monetary policy is substantially weaker than assumed in models advocating history-dependent policy. Not only is the share of policy-consistent expectations small, but even these fail to capture the quantitative effects of future monetary policy actions.

5.3 Expectation formation under different policy regimes

Having established that only a small proportion of participants form expectations consistent with the model and monetary policy direction, we next use participant-level expectations data to explore *how* they form their expectations. First, we observe a substantial degree of heterogeneity in expectations among experimental participants, evidenced by the session-level interquartile range of inflation and output gap forecasts for each treatment (Figure 13, Table D1 in Appendix D). Higher dispersion indicates more disagreement and less common understanding among forecasters.

Dispersion in both inflation and output forecasts is consistently higher in level-targeting regimes compared with rate-targeting regimes, with similar levels of dispersion observed across IT, DM, AIT-4, and AIT-10. Forecast disagreement increases sharply following the large aggregate demand shock in all regimes. While disagreement diminishes in all rate-targeting regimes post-shock, it continues to grow in PLT and NGDP. This higher post-shock disagreement in level-targeting regimes suggests that these frameworks are more challenging for participants to understand and to use when forming forecasts, compared with rate-targeting regimes.

The heterogeneity in expectations encourages us to examine in detail how people form their expectations. We classify participants' expectations using six models of expectations considered earlier in Section 3 (and listed in Table C1). For each phase of the experiment, we assign the expectation model that best fits each participant's forecasting behavior. We then analyze the distribution of forecasting models across policy frameworks.[6]

[6]Further details on our approach and results are available in Online Appendix E.

We find that participants' expectations are rarely consistent with ex-ante rationality in their forecasts of either inflation or the output gap (Figures E1 and E2 in Online Appendix E). Fewer than 5% of participants in any treatment can be classified as ex-ante rational, and in some treatments, this share is nearly zero. This indicates that participants lack a sufficient understanding of how economic fundamentals and monetary policy influence aggregate dynamics.

Backward-looking expectations, particularly extrapolative and constant-gain learning, dominate forecasting behavior, accounting for over 90% of participants across all treatments. Trend-extrapolation is the most commonly used model, ranging from approximately 50% of participants in PLT to over 90% in DM, with similar distributions observed during both the pre-shock and post-shock periods.

We examine the degree of trend extrapolation across monetary policy frameworks during the pre-shock and post-shock periods by comparing the empirical cumulative distribution functions (CDFs) of the trend-extrapolation parameter (τ) assigned to extrapolative forecasters (Figure 14). In the pre-shock phase, there is little variation across treatments in the distribution of the degree of trend chasing, with the median participant having a parameter value between 0.1 and 0.4, depending on the treatment.

In the post-shock phase, however, we observe significant differences in the strength of trend extrapolation across policy frameworks. Participants in level-targeting regimes (PLT and NGDP) respond much more strongly to past trends in inflation and output compared with participants in rate-targeting regimes. The CDFs of trend chasing parameter for level-targeting regimes are located to the right of CDF's for rate-targeting regimes, indicating higher values of trend chasing in level-targeting than in rate-targeting frameworks. PLT and NGDP show very strong trend-chasing behavior, with values of trend chasing parameter τ close to or exceeding 1. In contrast, trend chasing in rate-targeting regimes is much weaker, characterized by a median τ close to zero, reflecting simple naïve forecasting.

Our results challenge the assumption that the formation of expectations is invariant to the policy framework. Instead, they demonstrate that different monetary policy regimes can lead to distinct compositions of expectations, consistent with findings by Pfajfar and Žakelj [2014] and Assenza et al. [2021].

In the experiments, the realized inflation and output are determined based on the *median* expectations for inflation and output of experimental participants. We have classified median expectations for each session in each of regime as one of 6 expectations types (more details are in Appendix E). Next, we implemented simulations with expectations following the heuristics describing median expectations observed in the experimental data for each monetary policy regime during pre-shock and post-shock periods. These simulations exhibit dynamics in inflation, output and interest rates similar to those observed in experimental data. Notably, the simulations with heuristics show increasing volatility with increases in history dependence of monetary policy regimes culminating in the explosive dynamics in level-targeting regimes such as PLT and NGDP (see Figures E1 to E4 in Appendix E).

5.4 Framing and complexity of history-dependent frameworks

In this section, we disentangle the role of history-dependence versus the role of framing in the design of monetary policy. The objective of monetary policy under PLT is formulated as a price level target, whereas AIT-10 specifies its objective in terms of the inflation rate. Theoretically, under rational expectations, the performance of AIT converges to that of PLT as the AIT horizon approaches infinity. Experimentally, AIT-10's sufficiently long reaction horizon allows it to serve as a useful comparison with PLT, helping to isolate the role of framing in monetary policy design.

Cognitive complexity is likely more pronounced in regimes with greater history-dependence. As the degree of history-dependence increases—from IT to AIT-4 and then to AIT-10—the forecasting task becomes more complex, requiring participants to incorporate a longer range of past information. In contrast, concurrent rate-targeting frameworks like IT involve simpler decision-making; an ex-ante rational forecaster in IT would only need to respond to expected fundamentals.

All participants report forecasts of inflation and the output gap, regardless of the monetary policy regime. This does not bias the comparison toward rate-based rules; rather, it reflects the structure of the New Keynesian model, in which inflation is the state variable entering both the Phillips curve and the policy rule, even under level-targeting frameworks.

What differs across regimes is not the elicitation format, but the mapping from the policy rule to the forecasted variable. Under IT and DM, this mapping is direct: both the policy rule and the forecasted variable are expressed in terms of inflation rates. Under PLT and

NGDP, subjects must translate level deviations into an implied inflation rate (and output levels into output gaps in the case of NGDP). This added step increases cognitive load and can potentially weaken the expectations channel in level-targeting regimes.

Importantly, this is not an artifact of the experimental design, but a central mechanism emphasized in the theoretical literature: level-targeting rules require the public to understand make-up strategies embedded in the level path. Our design therefore captures a genuine difference in the cognitive demands of the regimes rather than introducing an artificial bias.

At the same time, PLT can be less demanding than long-horizon AIT, since the current price-level gap summarizes the entire history of past inflation deviations, reducing the need for participants to track long sequences of past rates.

We evaluate the role of policy complexity by analyzing participants' forecast performance. Panel A of Table 6 reports the mean and standard deviations of participants' absolute forecast errors across treatments during each of the three experimental phases. The evolution of median forecast errors are plotted over time, across different levels of history-dependence in IT, AIT-4, AIT-10, and PLT in Figure 15.

To further quantify the effects of history-dependence and level targeting on forecast errors, we estimate the following linear regression for each phase of the experiment:

$$lnAbsFE\pi_{i,t} = \alpha + \beta HistoryDependence_{i,t} + \gamma LevelTarget_i + \epsilon_{i,t} \tag{12}$$

where $HistoryDependence_{i,t}$ is an integer value that denotes how many *past* periods of data the central bank responds to.[7] $LevelTarget_i$ is a binary indicator variable equal to 1 for the level-targeting treatments (PLT and NGDP) and 0 for rate-targeting treatments. We perform a similar analysis for output gap forecast errors. Note that PLT and NGDP are excluded from the Postshock analysis as forecast errors in these treatments become exponentially large. The estimation results are presented in Panel B of Table 6.

In the Preshock and Shock phases, we find minimal differences in forecast errors across rate-targeting treatments, regardless of the degree of history-dependence. History-dependence does not consistently affect forecast performance, except for a slight impact on Preshock

[7] At the beginning of the experiment, $HistoryDependence_t = t - 1$ for AIT-4, AIT-10, PLT and NGDP. AIT-4 reaches its maximum *HistoryDependence* of 4 starting in Period 5, while AIT-10 reaches its maximum *HistoryDependence* of 10 starting in Period 11.

output gap forecast errors. Each additional period of history-dependence increases output gap forecast errors by 1.6% ($100 \times (e^{0.016} - 1)$).

In contrast, level-targeting regimes are significantly more difficult for participants to understand, leading to much larger forecast errors. Inflation forecast errors increase by 285% ($100 \times (e^{1.349} - 1)$) and output forecast errors by 329% ($100 \times (e^{1.457} - 1)$) compared with concurrent inflation-targeting treatments.[8]

In the Postshock phase, participants in rate-targeting treatments produce significantly smaller forecast errors compared to the Shock and Preshock phases, reflecting evidence of learning (Panel A of Table 6). However, significant differences emerge as history-dependence increases within these treatments. Each additional period of history-dependence is associated with an estimated 10–11% increase in inflation and output forecast errors, indicating that forecasting becomes progressively more challenging for participants in AIT-4 and AIT-10 compared with IT and DM.

In level-targeting treatments, forecast errors grow exponentially over time as expectations become increasingly extrapolative. When restricting the analysis to level-targeting treatments (results not shown in Table 6), we find that forecast errors increase by an average of 68% per period as the degree of history-dependence rises. These results demonstrate that the complexity in forecasting can be further exacerbated with very inertial monetary policy.

The differences in forecast performance between PLT and AIT-10 are striking, even when the degree of history-dependence is the same during the first eleven periods of the experiment. We attribute the significantly larger forecast errors under PLT, along with the greater forecast dispersion and more extrapolative expectations (Figure 13 and Figure 14), to the additional cognitive burden of framing the policy target in terms of the price level. These disparities become more pronounced over time, further exacerbated by the increasing history-dependence inherent in PLT as the experiment progresses.

In Panel C, we further aim to disentangle the complexity of NGDP relative to PLT by comparing forecast errors under the two regimes. The results show that NGDP produces significantly larger output forecast errors during the Preshock phase and larger inflation forecast errors during the Shock phase than in PLT. We attribute these greater errors un-

[8]The large negative demand shock also contributes to significantly higher forecast errors across all treatments, as participants must contend with inactive monetary policy constrained by the effective lower bound.

der NGDP to its additional complexity compared with PLT, as participants must translate deviations of nominal GDP level from its target into forecasts of inflation rate *and* output gap. Unlike in PLT, NGDP does not have an output gap target of zero. This makes it more difficult for participants in the latter treatment to formulate their output gap forecasts.

Overall, the comparability across rate-targeting regimes with varying degrees of history-dependence suggests that increased history-dependence poses only a modest additional cognitive challenge. However, the stark differences between PLT and AIT-10 indicate that framing the target in terms of the price level is significantly more difficult for participants to process than framing it in terms of the inflation rate.

5.5 Expectations and history-dependence: two cautionary tales

Our findings lead to two cautionary conclusions about the risks posed by history-dependent monetary policy frameworks.

The first cautionary tale concerns inflation overshooting in regimes with average inflation targeting (AIT) with longer horizons. Greater history dependence amplifies volatility and persistence in inflation and output, resulting in deeper and longer deflationary or inflationary episodes. Following an ELB shock, deflation is both deeper and longer-lasting than predicted by theory, as the expectations channel fails to provide the anticipated lift while monetary policy remains constrained at the ELB. During the recovery, inflation overshoots as history-dependent monetary policy remains "low for longer," and inflation expectations feed off of rising inflation. The interaction of monetary policy inertia with inertia in expectations formation leads to greater volatility and more persistent fluctuations in inflation and the output gap.

The second cautionary tale involves the risk of deflationary spirals under PLT and NGDP level targeting following an ELB shock. Central banks have expressed concerns about deflationary spirals since the Great Financial Crisis, particularly when monetary policy is constrained by the ELB. In our experiments, the combination of over-reaction and delays in expectations to shocks (subsection 5.1) with high history dependence in monetary policy in PLT and NGDP results in increased volatility and persistence in inflation, even during the stable pre-shock period, with some economies experiencing deflationary episodes before the ELB shock (Section 4). These early deflationary experiences unanchor inflation expectations, setting the stage for further destabilization. When the ELB shock occurs, inflation

expectations – already deflationary – become further entrenched. With constrained monetary policy, the entrenched deflationary expectations cause the economy to unravel into a deflationary spiral.

PLT and NGDP level targeting exacerbate these risks by introducing additional complexity through framing the policy target in terms of the price level or nominal GDP level, which is cognitively more demanding than rate-targeting frameworks (subsection 5.4). This complexity leads participants to rely on simple trend-chasing heuristics, which further destabilize the economy. As the deflationary spiral develops, participants' trend-chasing expectations intensify, reinforcing the downward spiral (subsection 5.3).

6 The role of central bank communication in price level targeting

Our experiments highlight the significant challenges individuals face when forecasting in the economies with level-targeting mandates. Many participants fail to grasp that the central bank must return the price level to its target, and even fewer understand the extent of the make-up strategy required. Moreover, our experiments shows that PLT places greater cognitive demands on participants because they must translate price-level deviations into implied inflation rates.

One potential solution is to provide central bank guidance on the implied path of inflation. Recent studies have shown that precise, relevant, and well-communicated point projections can reduce forecasting complexity and better anchor expectations [Mokhtarzadeh and Petersen, 2020, Rholes and Petersen, 2021, Petersen and Rholes, 2022].

To test this, we introduced a new treatment, PLT Comm, in which the PLT framework was extended to include central bank communication of precise projected paths for inflation and the output gap. This communication aimed to simplify forecasting under a price-level targeting regime by providing both qualitative and quantitative information. Participants were informed whether the price level was above or below the target, that the central bank would respond by adjusting interest rates accordingly, and how such adjustments were expected to influence inflation and output.

Additionally, precise point projections were displayed on participants' screens as green fore-

casts extending five periods beyond the inflation and output gap time series. These projections were generated using the data-generating process, incorporating exogenous shocks and the recent price level, as explained to participants in the instructions. Six sessions of PLT Comm were conducted in July 2022, following the same protocols outlined in Section 2.

The aggregate dynamics from PLT Comm are presented in Figure 16. Across both phases of the experiment, the communicated projections lead to highly stable inflation and output gap dynamics, significantly outperforming all other treatments in minimizing aggregate losses (Table 2). Only one of the six sessions experiences a significant deflationary episode at the ELB that fails to recover. Session-level comparisons of aggregate losses between PLT Comm and other treatments show highly significant differences pre-shock (including all sessions) and post-shock (excluding the outlier session 5) based on pairwise Wilcoxon rank-sum tests ($p < 0.05$).

This improved economic stability is driven by a marked shift in the distribution of expectation models. Figure F1 in Online Appendix F illustrates these distributions. The share of participants using trend-extrapolative models for inflation forecasts drops from 55% in PLT (in both phases) to 36% in the pre-shock phase and 43% in the post-shock phase in PLT Comm. Additionally, the share of ex-ante rational forecasters increases from 2.2% in PLT to 7.1% in the pre-shock phase of PLT Comm.

The majority of PLT Comm participants are able to forecast both inflation and the output gap in the correct direction across all phases (Table 4). More than 60% of participants form policy-consistent expectations during the pre-shock phase in PLT Comm, compared with less than one-third in all other treatments. This increase in ex-ante rational and policy-consistent expectations can be attributed to the provided projections, which are based on the rational expectations solution and effectively incorporated by participants into their forecasts.

7 Conclusion

Our findings identify potential vulnerabilities of history-dependent monetary policies. We identify two key risks associated with inertial monetary frameworks in the presence of backward-looking expectations: (1) inflation overshooting under AIT with long horizons, and (2) deflationary spirals under PLT when monetary policy is constrained at the ELB. These outcomes run counter to the anticipated benefits of history dependence, particularly at the ELB. Using a standard New Keynesian model where expectations are provided by

experimental participants, we explore how the nature of expectations undermines the expectations channel.

First, we provide evidence that shorter history-dependence works better to manage the economy. While modest history-dependence can guide inflation expectations, excessive history-dependence can be detrimental. Participants rely heavily on recent trends and fail to adequately account for future monetary policy paths and fundamentals, undermining the make-up strategies critical to history-dependent frameworks. As a result, longer history dependence leads to monetary policy growing out of sync with the economy. A "low for longer" monetary policy can lead to inflation overshooting the central bank's target as the economy recovers from recession.

Second, participants' expectations primarily reflect trend-chasing behavior, propagating current dynamics into the future rather than counteracting them through anticipation of future policy actions. This contributes to the inflation overshoot in AIT-10 during economic recovery as expectations feed off increasing inflation and monetary policy stays low for longer. Trend-chasing expectations feed deflationary spirals in PLT when monetary policy is constrained by the ELB.

Third, the framing of monetary policy targets is critical for the success of history-dependent frameworks. Framing the target in terms of the inflation rate, rather than the price level, reduces cognitive complexity, leading to fewer trend-chasing expectations. With better managed expectations, monetary policy can be more effective at stabilizing the economy.

Finally, communication is a vital component of monetary policy. Providing *relevant* macroeconomic projections—such as inflation and output gap forecasts—helps simplify complex framework such as PLT and guides expectations effectively. However, care must be taken in selecting what information to communicate. Highlighting evolving targets can undermine credibility by drawing attention to the central bank's inability to meet them [Arifovic and Petersen, 2017]. Similarly, communicating the path of monetary policy can confuse forecasters who struggle to translate interest rate projections into inflation and output gap forecasts [Mokhtarzadeh and Petersen, 2020, Kryvtsov and Petersen, 2021].

Our framework opens avenues for further research. This experiment focused on aggregate demand shocks, where the "divine coincidence" may have simplified participants' understanding of monetary policy's dual stabilization role. Future work could explore monetary policy

under persistent cost-push pressures, which have gained renewed attention in the wake of the pandemic. Additionally, the framework could investigate the roles of cognitive complexity and limited common knowledge in expectation formation, particularly under heightened strategic complementarities. Our experiment focuses on communication under PLT. However, there is scope to examine the role of communication in other level-targeting regime such as NGDP. Moreover, our experiment examines the performance of a central bank in a contextless environment where it has no reputation and historical engagement with participants. In reality, central banks operate with sustained communication strategies, reputational capital, financial market discipline, and can learn over time. Future experiments should examine the role these institutional features play in shaping the performance of HDMP.

References

Klaus Abberger, Anne Kathrin Funk, Michael Lamla, Sarah Lein, and Stefanie Siegrist. The pass-through of inflation expectations into prices and wages: Evidence from an rct survey. Discussion Paper 19595, Centre for Economic Policy Research (CEPR), 2024. URL https://cepr.org/publications/dp19595. RCT survey of firms; finds higher pass-through for short-term expectations than long-term expectations.

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

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.

Steve Ambler. Price-level targeting and stabilisation policy: A survey. *Journal of Economic Surveys*, 23(5):974–997, 2009. doi: 10.1111/j.1467-6419.2009.00601.x.

George-Marios Angeletos and Chen Lian. Forward guidance without common knowledge. *American Economic Review*, 108(9):2477–2512, 2018.

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.

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

Jasmina Arifovic, Isabelle Salle, and Hung Truong. History-dependent monetary policy: A lab experiment and a henk model. *Available at SSRN 4586602*, 2023.

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*, 117:170–186, 2021.

Te Bao, Myrna Hennequin, Cars Hommes, and Domenico Massaro. Coordination on bubbles in large-group asset pricing experiments. *Journal of Economic Dynamics and Control*, 110: 103702, 2020.

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.

Pedro Bordalo, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–2782, 2020.

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

William A Brock and Cars H Hommes. A rational route to randomness. *Econometrica: Journal of the Econometric Society*, pages 1059–1095, 1997.

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, and Michael Weber. Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, (130), 2022.

Olivier Coibion, Yuriy Gorodnichenko, Edward S. II Knotek, and Raphael Schoenle. Average inflation targeting and household expectations. *Journal of Political Economy – Macroeconomics*, 1, 2023.

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

Jose Dorich, Rhys R. Mendes, and Yang Zhang. The Bank of Canada's "horse race" of alternative monetary policy frameworks: Some interim results from model simulations. *Bank of Canada Staff Discussion Paper*, 2021-13, 2021.

Gauti B. Eggertsson and Michael Woodford. The zero lower bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity, Economic Studies Program*, 34(1):139–235, 2003.

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.

George W Evans, Cars Hommes, Bruce McGough, and Isabelle Salle. Are long-horizon expectations (de-) stabilizing? theory and experiments. *Journal of Monetary Economics*, 132:44–63, 2022.

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.

Julio Garín, Robert Lester, and Eric Sims. On the desirability of nominal gdp targeting. Technical Report w21420, NBER, 2015. URL http://www.nber.org/papers/w21420.

Erwan Gautier, Frédérique Savignac, and Olivier Coibion. Firms' inflation and wage expectations during the inflation surge. Working Paper 33799, National Bureau of Economic Research, 2025. URL https://www.nber.org/papers/w33799.

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*, 129: S52–S66, 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, 2021a. ISSN 0167-2681.

Cars Hommes and Tomasz Makarewicz. Experimental study on the effect of nominal price level versus inflation targeting with and without guidance. *Journal of Economic Behavior and Organization*, 2021b.

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.

Cars Hommes, Anita Kopányi-Peuker, and Joep Sonnemans. Bubbles, crashes and information contagion in large-group asset market experiments. *Experimental Economics*, 24(2): 414–433, 2021.

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.

Òscar Jordà. Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182, 2005.

Olena Kostyshyna, Luba Petersen, and Jing Yang. A horse race of monetary policy regimes: An experimental investigation. Technical report, National Bureau of Economic Research, 2022.

Olena Kostyshyna, Luba Petersen, and Jing Yang. Why don't history-dependent monetary policies work? 2024.

Jeremy Kronick and Luba Petersen. Crossed wires: Does fiscal and monetary policy coordination matter? *CD Howe Institute*, 633, 2022.

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.

Felix Mauersberger. Monetary policy rules in a non-rational world: A macroeconomic experiment. *Journal of Economic Theory*, 197:105203, 2021.

Bennett T. McCallum. Nominal gdp targeting: Policy rule or discretionary splurge? *Journal / policy forum (various versions available)*, 2015. Discussion/overview of NGDP targeting literature and implementation issues.

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.

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

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

Damjan Pfajfar and Blaž Žakelj. Uncertainty in forecasting inflation and monetary policy design: Evidence from the laboratory. *International Journal of Forecasting*, 32(3):849–864, 2016.

Bruce J Preston. Learning about monetary policy rules when long-horizon expectations matter. *International Journal of Central Banking*, September, 2005.

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

Julio J. Rotemberg and Michael Woodford. An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual*, 12:297–346, 1997.

Glenn D. Rudebusch and Lars E. O. Svensson. Policy rules for inflation targeting. *Monetary Policy Rules*, pages 203–262, 1999.

Isabelle Salle. What to target? Insights from theory and lab experiments. *Journal of Economic Behaviour and Organization*, pages 514–533, 2023.

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

Scott B. Sumner. Nominal gdp targeting: A simple rule to improve fed performance. *Cato Journal*, 34(2):269–292, 2014. URL https://ciaotest.cc.columbia.edu/journals/cato/v34i2/f_0031475_25523.pdf.

Lars E. O. Svensson. Price-level targeting versus inflation targeting: A free lunch? *Journal of Money, Credit and Banking*, 31(3):277–295, 1999. Classic comparison highlighting commitment/expectation channels.

Lars EO Svensson. Escaping from a liquidity trap and deflation: The foolproof way and others. *Journal of Economic Perspectives*, 17(4):145–166, 2002.

David Vestin. Price-level versus inflation targeting. *Journal of Monetary Economics*, 53(7): 1361–1376, 2006.

Joel Wagner, Tudor Schlanger, and Yang Zhang. A horse race of alternative monetary policy regimes under bounded rationality. *Journal of Economic Dynamics and Control*, 154, 2023.

Carl E Walsh. *Monetary theory and policy*. MIT press, 1998.

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.

Alexander L Wolman. Real implications of the zero bound on nominal interest rates. *Journal of Money, Credit and Banking*, pages 273–296, 2005.

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

Figure 1: Simulations with rational expectations

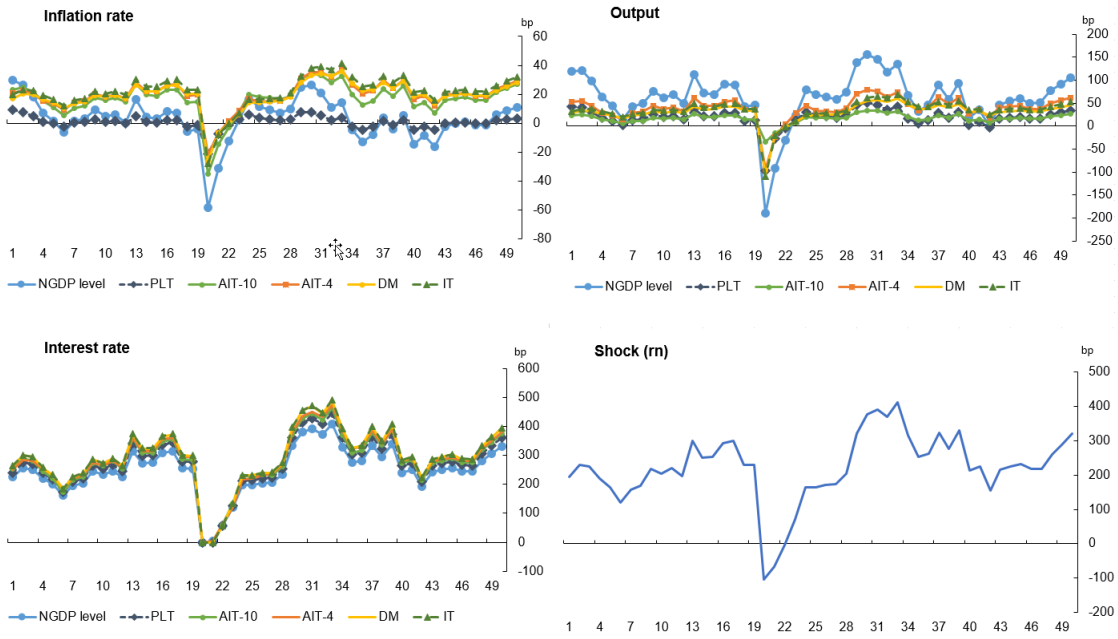
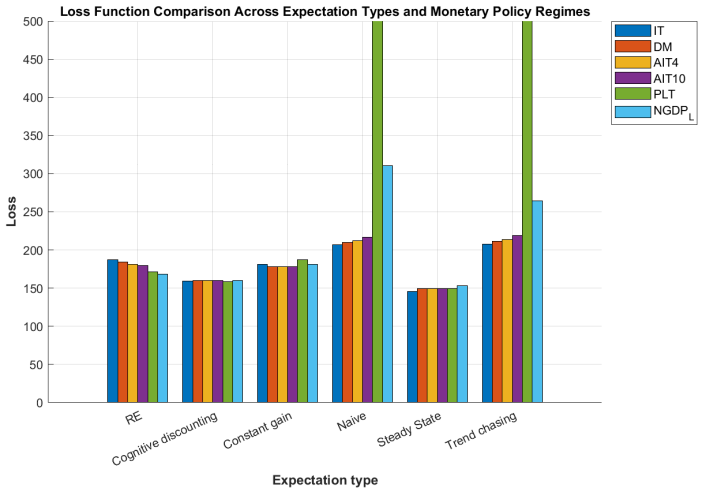
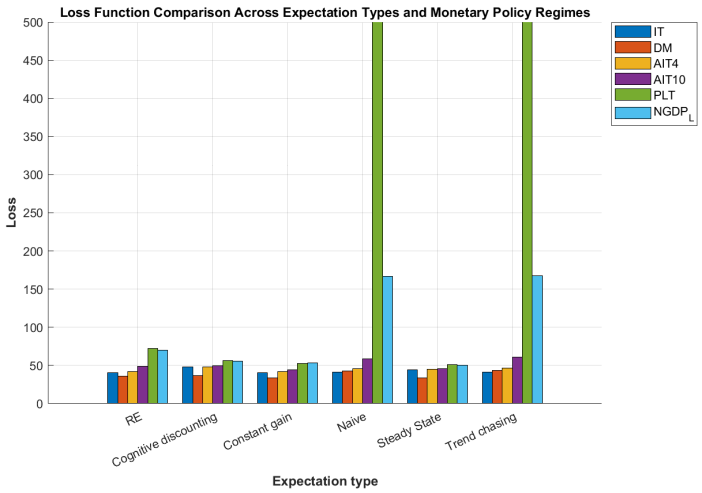


Figure 2: Summary of losses from the simulations



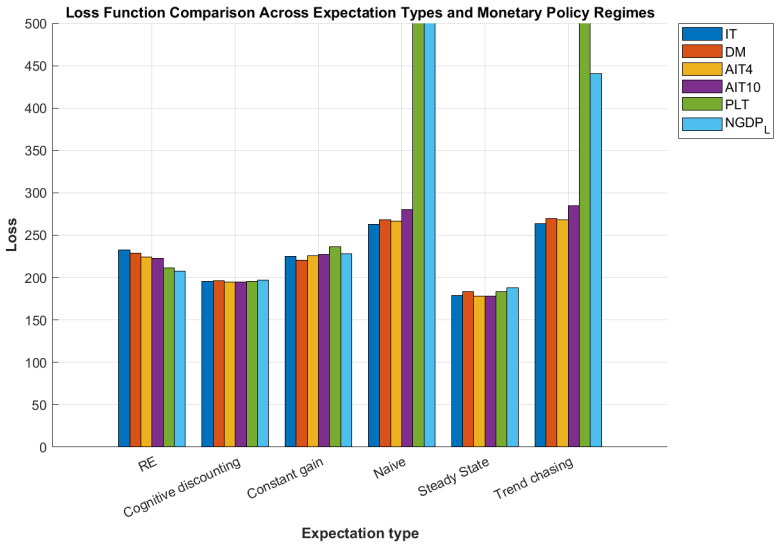
This figure presents the losses (equation 9) from the simulations of our model with different types of expectations over periods 1–50. For the ease of comparison of losses across regimes and expectations, the maximum value for y-axis was set at 500. The losses for PLT with naive expectations are 1,182; and for PLT with trend chasing expectations are 393,597, both values exceed the maximum of y-axis.

Figure 3: Summary of losses from the simulations with higher policy coefficients



This figure presents the losses (equation 9) from the simulations of our model with different types of expectations over periods 1–50, **with policy coefficients 25% higher than in the baseline rules**. For the ease of comparison of losses across regimes and expectations, the maximum value for y-axis was set at 500. The losses for PLT with naive expectations are 615; for PLT with trend expectations are 628.

Figure 4: Summary of losses from the simulations with lower policy coefficients



This figure presents the losses (equation 9) from the simulations of our model with different types of expectations over periods 1–50, **with policy coefficients 25% lower than in the baseline rules**. For the ease of comparison of losses across regimes and expectations, the maximum value for y-axis was set at 500. The losses for PLT with naive expectations are 1,403,891; for PLT with trend expectations are 703,419; and for NGDP with naive expectations are 3,266, which exceed the maximum of y-axis.

Figure 5: Aggregate dynamics of inflation, output and interest rate in dual mandate (DM) and inflation targeting (IT) treatments

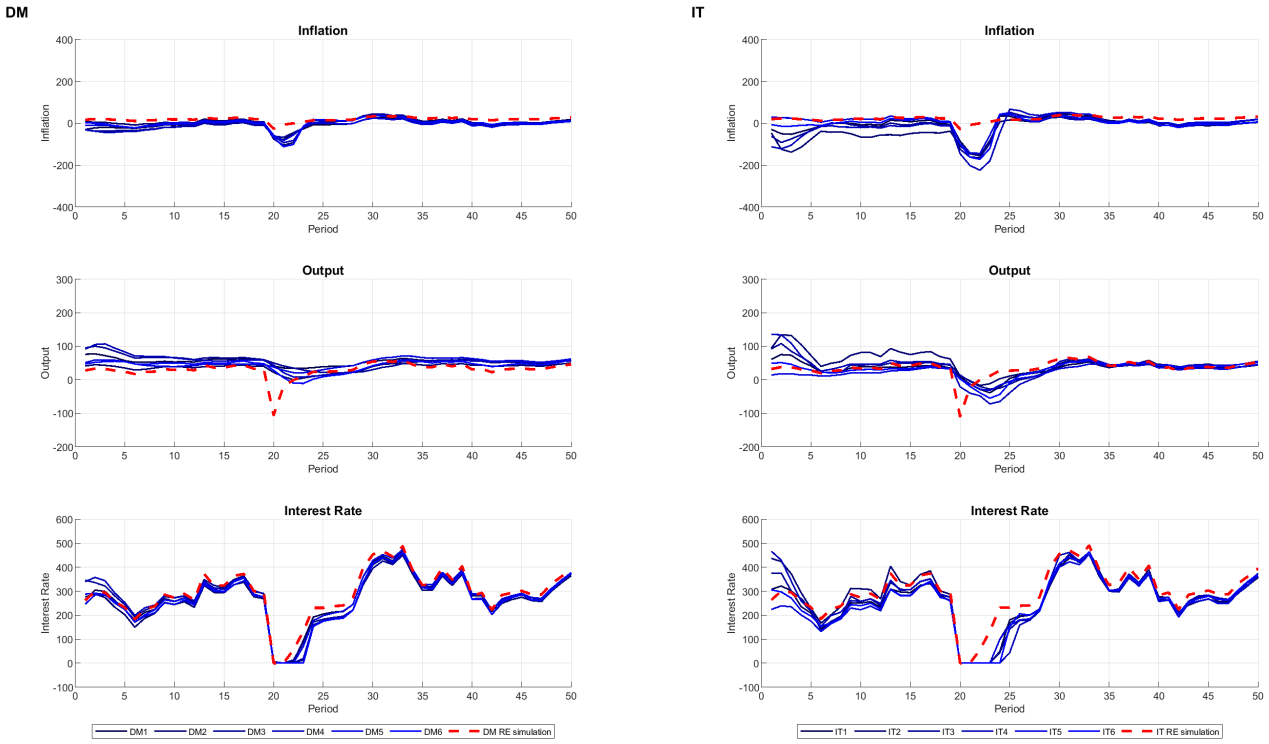


Figure 6: Aggregate dynamics of inflation, output and interest rate in AIT-4 and AIT-10 treatments

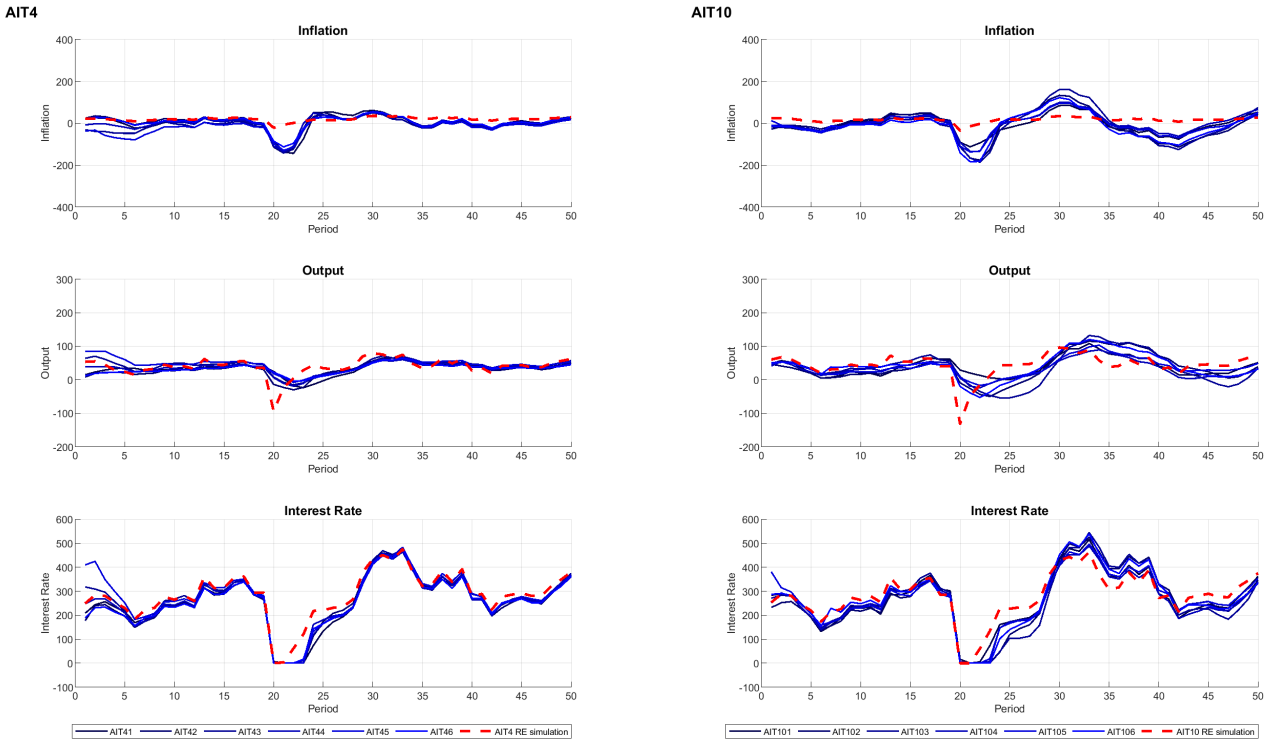


Figure 7: Aggregate dynamics of inflation, output, and interest rate in price-level targeting (PLT) treatment

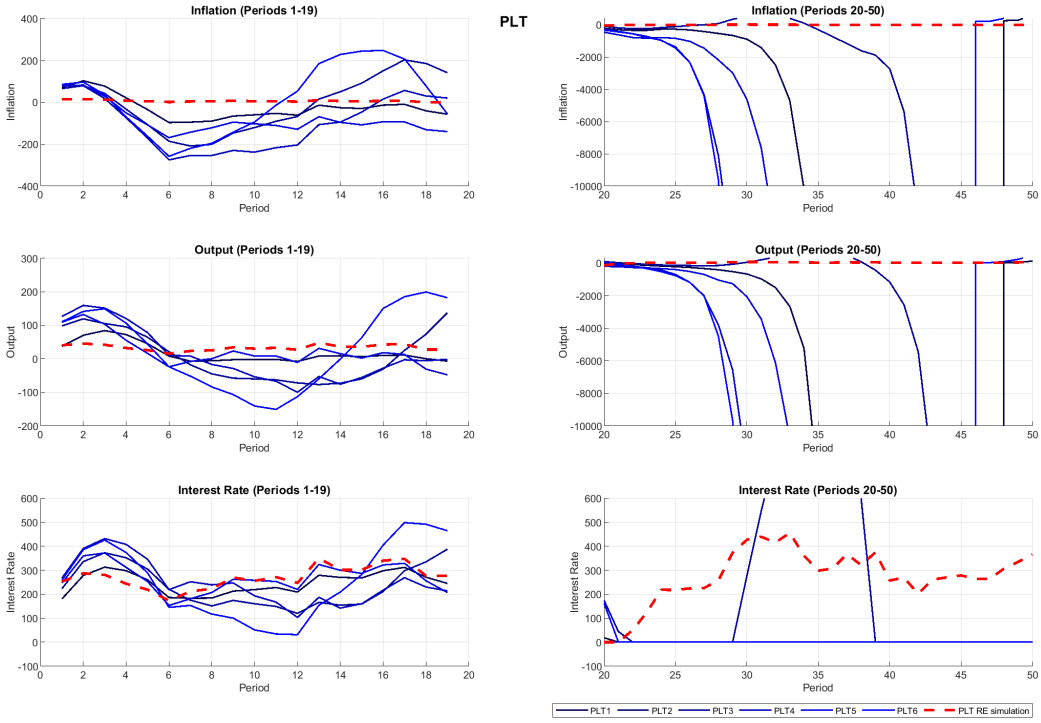


Figure 8: Aggregate dynamics of inflation, output, and interest rate in NGDP level targeting (NGDP) treatment

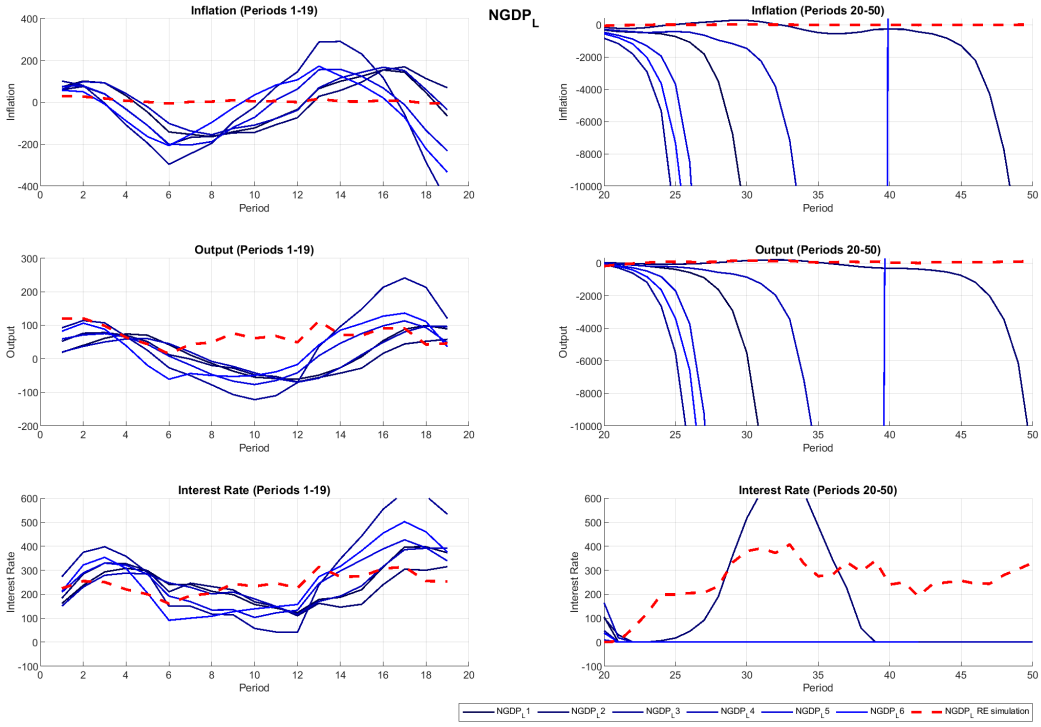
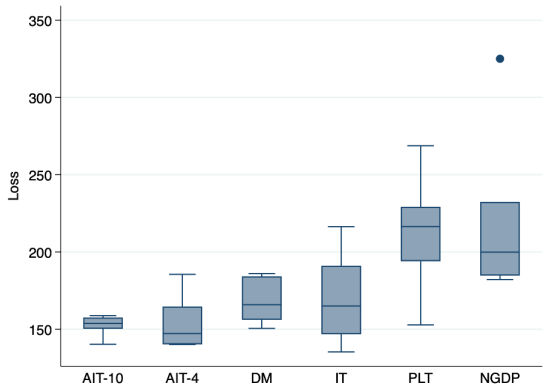
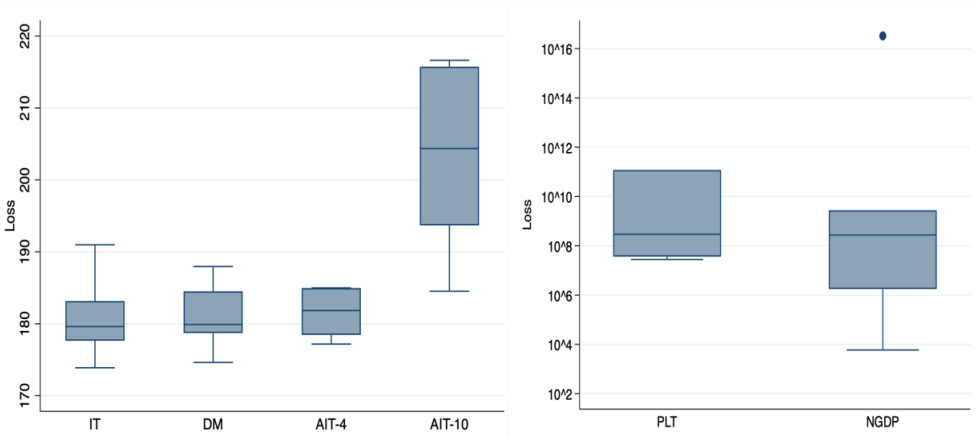


Figure 9: Distribution of session losses, by phase

(a) Pre-shock

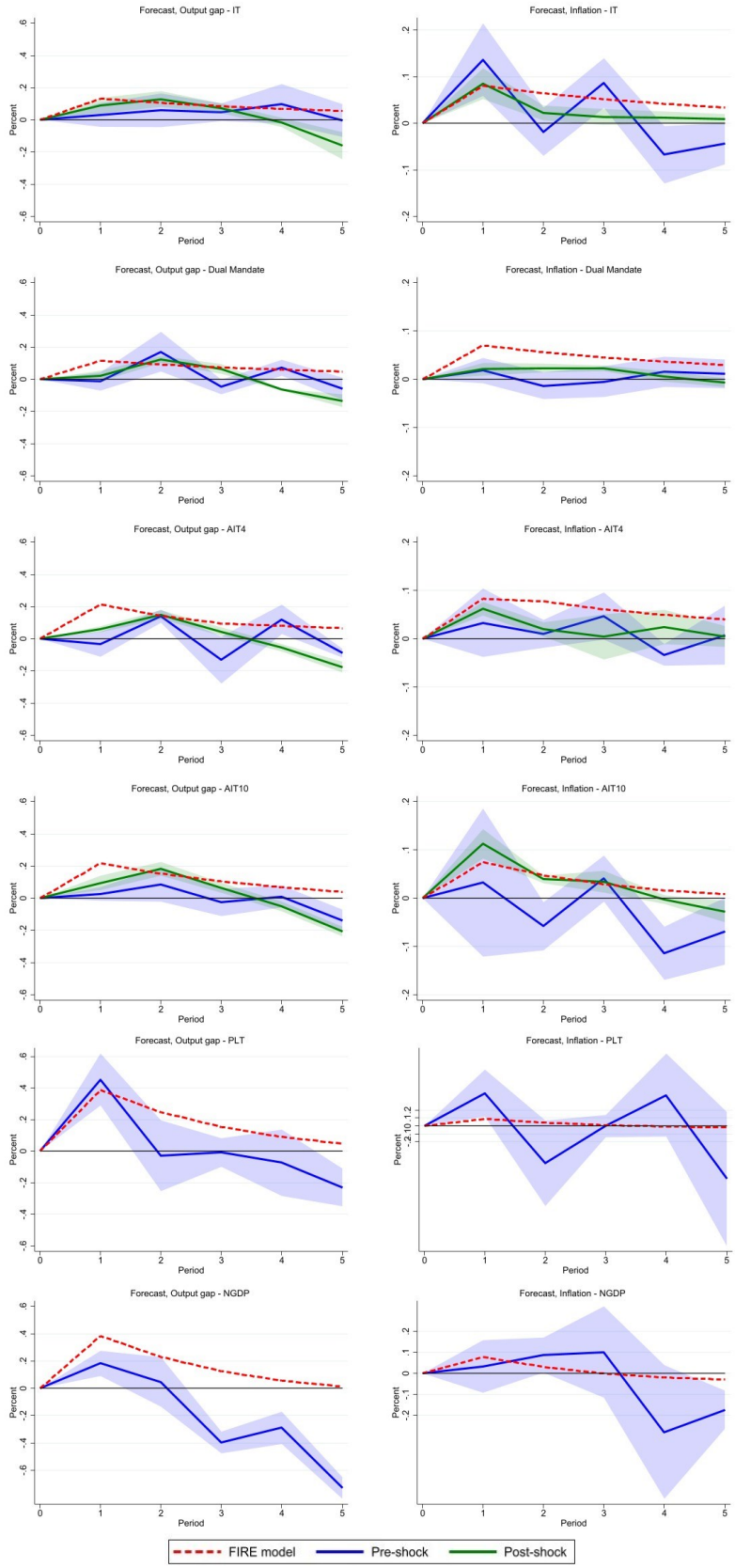


(b) Post-shock



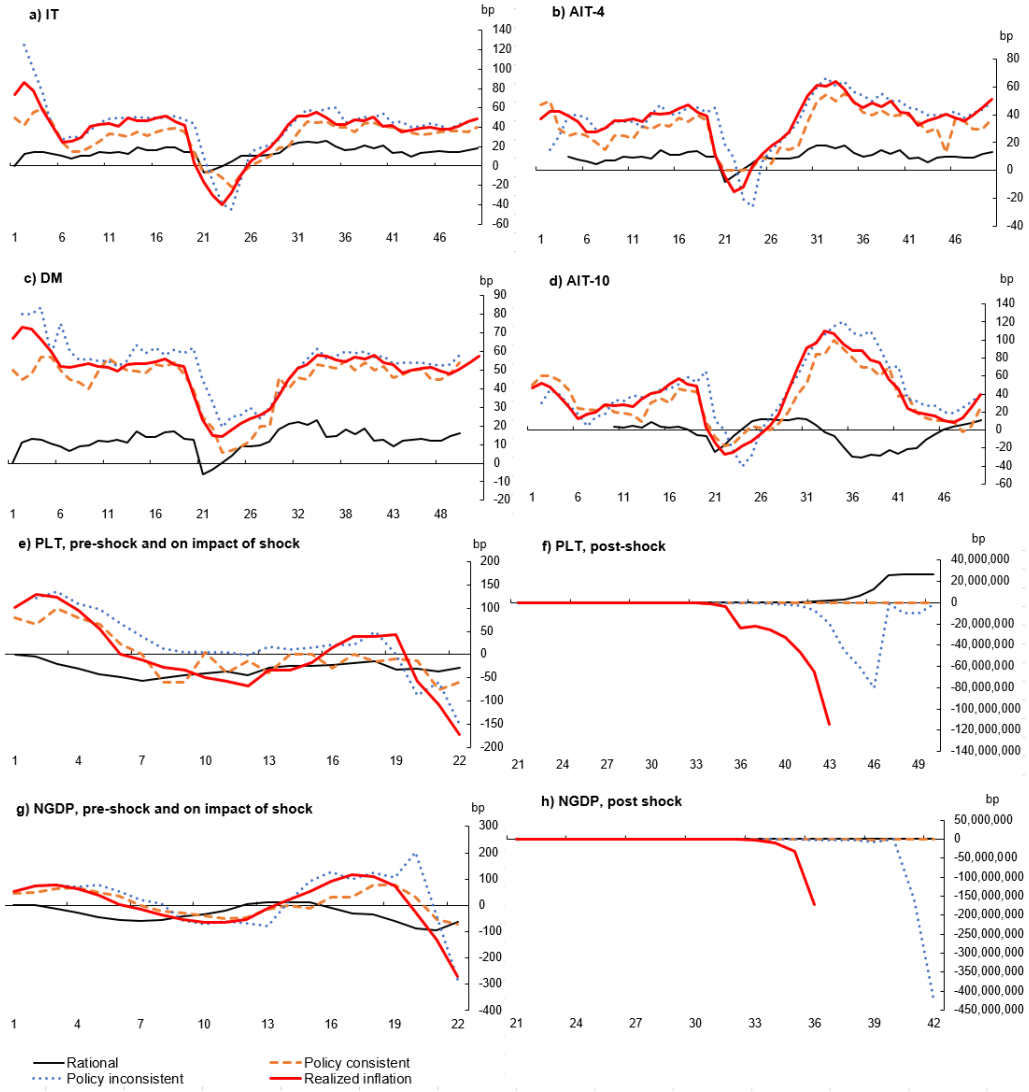
Dots in the figures represent values outside upper adjacent value (upper quartile $\pm$ 3/2 times the interquartile range). The y-axis in Panel (b) for PLT and NGDP is a logarithmic scale.

Figure 10: Response of forecasts to a +1% demand impulse.



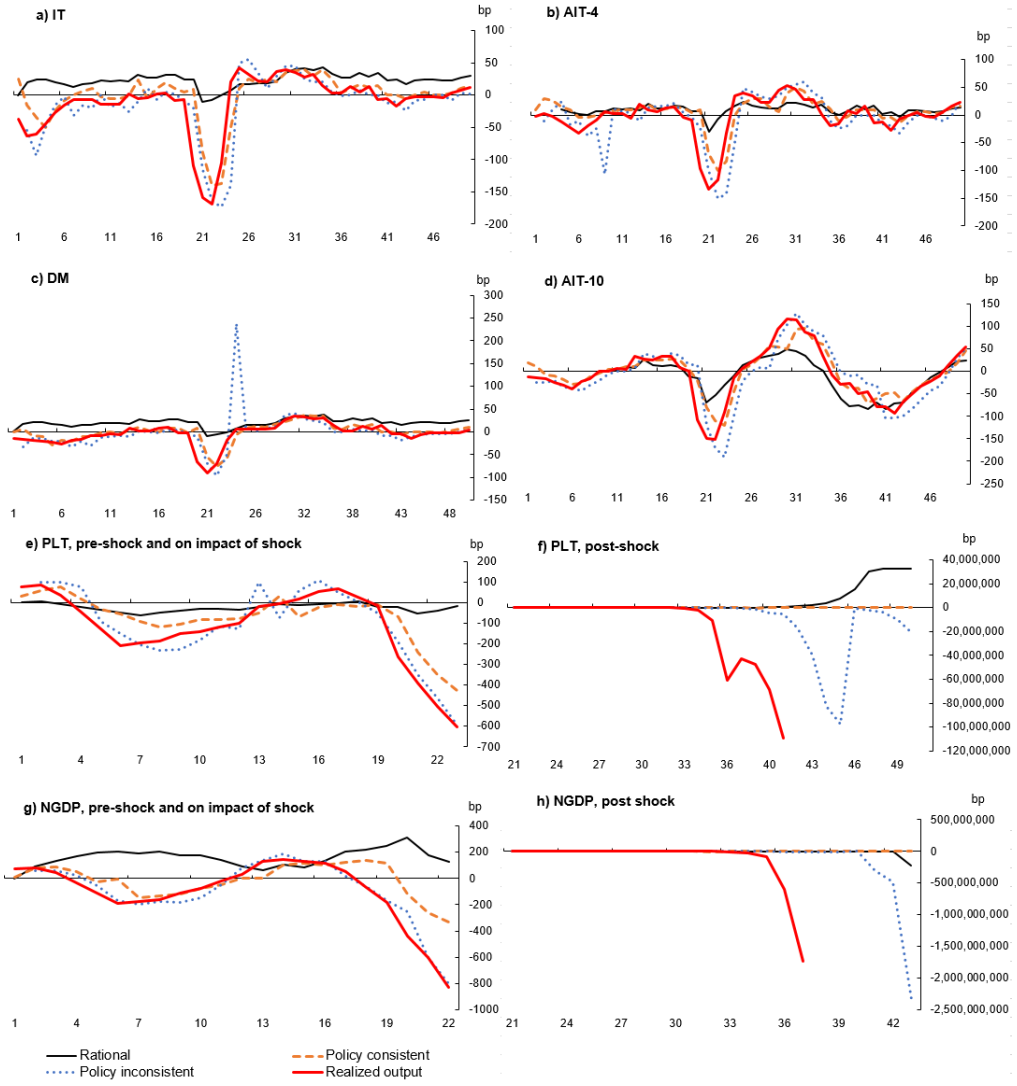
Notes: This figure presents IRFs estimated by local projections using Equation 11. Shaded areas denote the 90% confidence interval. IRFs in the FIRE model are the simulated responses to a one-standard deviation shock to aggregate demand.

Figure 11: Policy consistent and policy inconsistent inflation forecasts



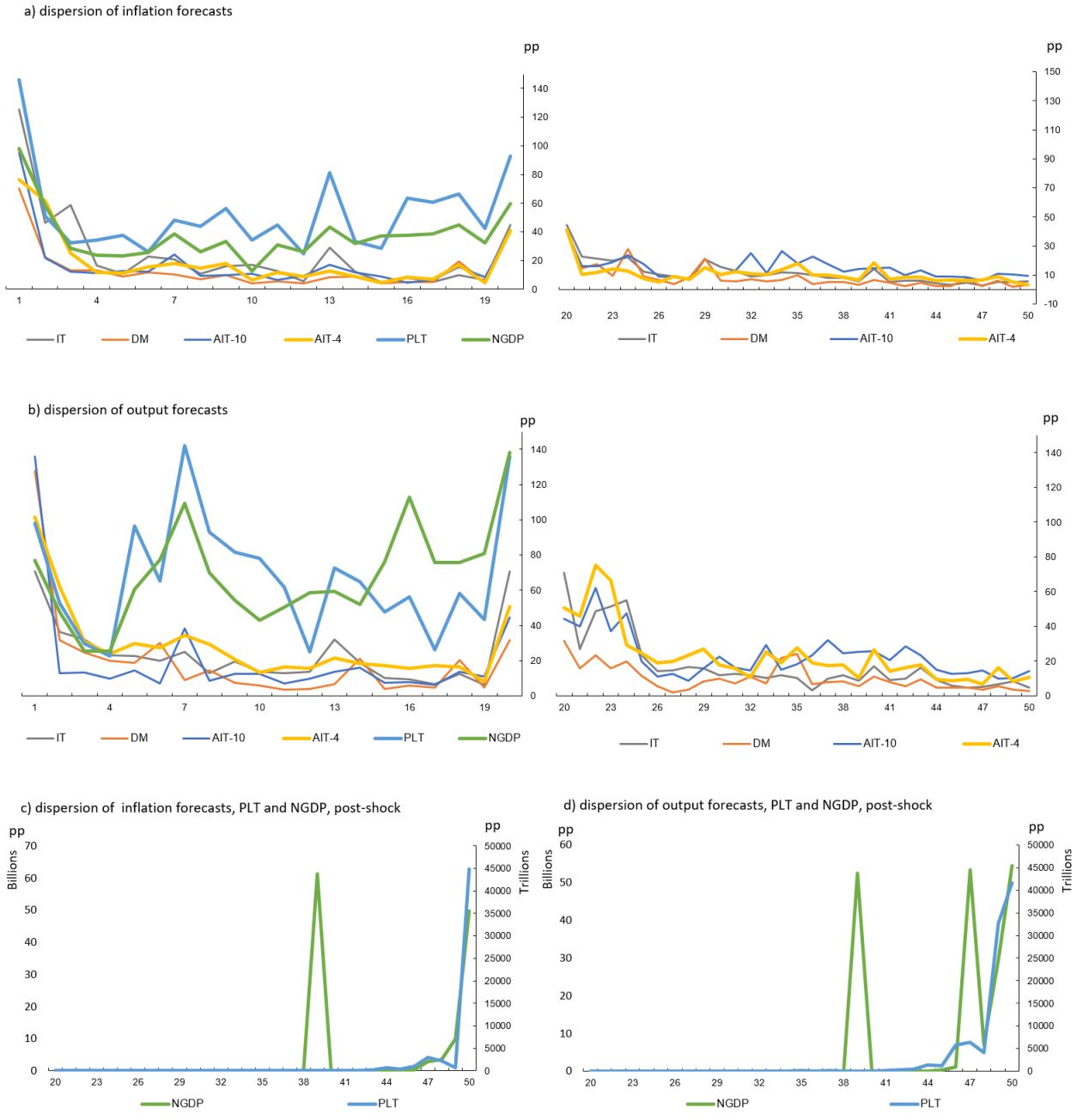
This figure presents the median policy consistent and policy inconsistent forecasts, averaged across all sessions of each treatment.

Figure 12: Policy consistent and policy inconsistent output forecasts



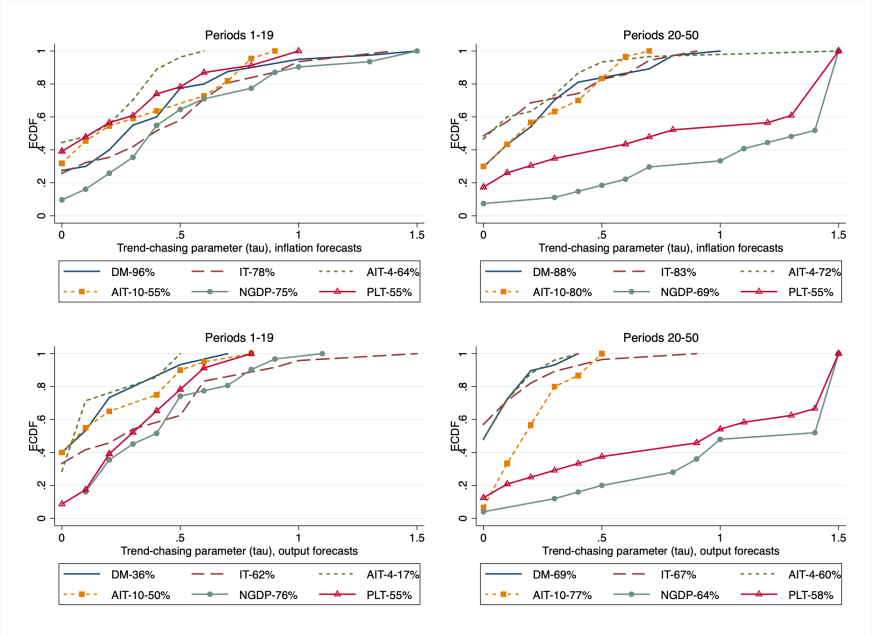
This figure presents the median policy consistent and policy inconsistent forecasts, averaged across all sessions of each treatment.

Figure 13: Dispersion of inflation and output forecasts



Notes: 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.

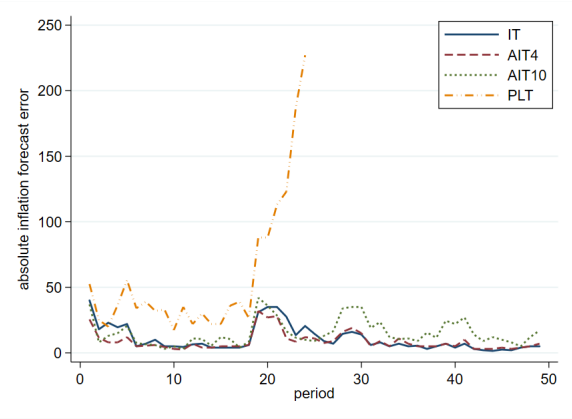
Figure 14: Distribution of trend-chasing parameter τ in inflation and output forecasts by phase



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.

Figure 15: Median absolute forecast errors, by treatment

(a) Inflation



(b) Output gap

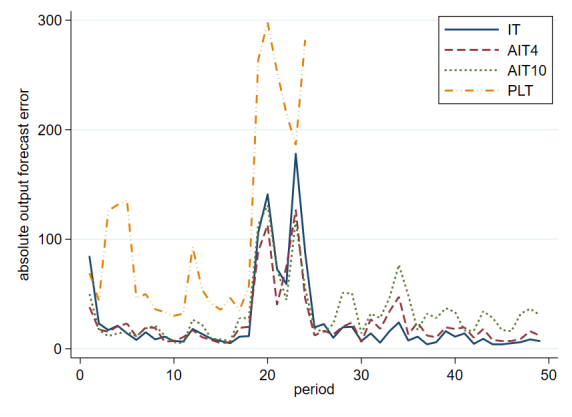


Figure 16: Aggregate dynamics of inflation, output, and interest rate in PLT Comm treatment

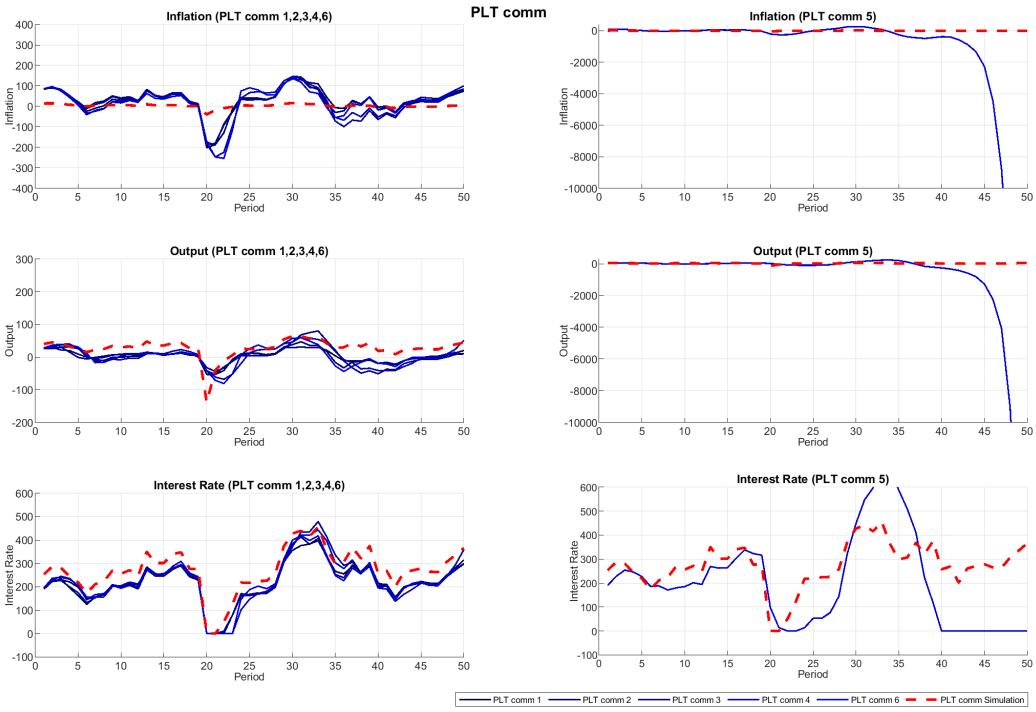


Table 1: Losses from simulations with rational expectations

Regime	Periods 1-50	Periods 1-19	Periods 20-50
NGDP	168.2	153.8	176.4
PLT	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

Table 2: Losses in laboratory experiments

Regime	Periods 1-50	Periods 1-19	Periods 20-50
AIT-4	172	154	182
	(3.2)	(7.3)	(1.3)
DM	176	168	181
	(2.6)	(6)	(1.9)
IT	177	170	181
	(5.5)	(12.3)	(2.4)
AIT-10	186	152	203
	(4)	(2.7)	(5.7)
PLT	31×10^9	213	39×10^9
	(19×10^9)	(16)	(25×10^9)
NGDP	4×10^{15}	221	5×10^{15}
	(4×10^{15})	(22.3)	(5×10^{14})
PLT Comm	2723 [155*]	131	3435 [169*]
	(2568)[6.7*]	(2.6)	(3265.9)[9.2*]

Notes: This table presents losses averaged across all sessions of each treatment with standard errors in parentheses.
* Values in square brackets exclude single outlier session.

Table 3: Wilcoxon rank order test, statistical significance

Periods 1-19					
	NGDP	PLT	DM	IT	AIT-4
PLT	0.435				
DM	0.008	0.019			
IT	0.039	0.027	0.436		
AIT-4	0.005	0.005	0.055	0.168	
AIT-10	0.002	0.008	0.027	0.212	0.261
Periods 20-50					
	NGDP	PLT	DM	IT	AIT-4
PLT	0.211				
DM	0.002	0.002			
IT	0.002	0.002	0.316		
AIT-4	0.002	0.002	0.374	0.316	
AIT-10	0.002	0.002	0.003	0.003	0.005
All periods					
	NGDP	PLT	DM	IT	AIT-4
PLT	0.211				
DM	0.002	0.002			
IT	0.002	0.002	0.436		
AIT-4	0.002	0.002	0.168	0.168	
AIT-10	0.002	0.002	0.100	0.100	0.013

Notes: 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 4: Share of forecasts exhibiting policy consistent expectations

	Preshock (Periods 1-19)			Shock (Periods 20-21)			Postshock (Periods 22-50)		
	Inflation	Output	Both	Inflation	Output	Both	Inflation	Output	Both
NGDP	0.59	0.48	0.27	0.85	0.22	0.12	0.29	0.29	0.18
	(0.02)	(0.02)	(0.02)	(0.04)	(0.05)	(0.04)	(0.01)	(0.01)	(0.01)
PLT	0.49	0.64	0.29	0.71	0.58	0.49	0.36	0.38	0.26
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)
DM	0.47	0.49	0.25	0.63	0.56	0.27	0.36	0.57	0.16
	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)
IT	0.54	0.54	0.33	0.70	0.49	0.27	0.46	0.54	0.21
	(0.02)	(0.02)	(0.02)	(0.05)	(0.06)	(0.05)	(0.01)	(0.01)	(0.01)
AIT-4	0.48	0.63	0.33	0.58	0.52	0.20	0.43	0.56	0.25
	(0.02)	(0.02)	(0.02)	(0.05)	(0.06)	(0.04)	(0.01)	(0.01)	(0.01)
AIT-10	0.48	0.57	0.33	0.63	0.80	0.51	0.51	0.65	0.36
	(0.02)	(0.02)	(0.02)	(0.06)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)
PLT Comm	0.71	0.74	0.62	0.92	0.85	0.82	0.66	0.62	0.50
	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.01)	(0.01)	(0.01)

Notes: This table presents the session-level mean share of forecasts in the direction of the rational expectations equilibrium solution. Standard errors are presented in parentheses.

Table 5: Evolution of policy-consistent expectations

Dep. Var. Pre-shock: Periods 1-19 $\mathbb{1}_{i,t}^{PolicyConsistent}$	IT	DM	AIT-4	AIT-10	PLT	NGDP	PLT Comm	
$\mathbb{1}_{i,t-1}^{PolicyConsistent}$	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 $\mathbb{1}_{i,t}^{PolicyConsistent}$	IT	DM	AIT-4	AIT-10	PLT	NGDP	PLT Comm	
$\mathbb{1}_{i,t-1}^{PolicyConsistent}$	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

Notes: This table presents estimation results from logit panel regressions: $\mathbb{1}_{i,t}^{PolicyConsistent} = \alpha + \beta_1 \mathbb{1}_{i,t-1}^{PolicyConsistent} + \beta_2 AbsDevFromTarget_{t-1} + \beta_4 \mu_i + \epsilon_{i,t}$ where $\mathbb{1}_{i,t}^{PolicyConsistent}$ is equal 1 if inflation expectations of participant i in period t are classified as policy consistent. $AbsDevFromTarget_{t-1}$ is one of the following variables: $|\pi_{t-1} - \pi^*|$ is deviation of last-period inflation from target for IT and DM; $|\frac{1}{4}\sum_{t-4}^{t-1}\pi_j - \pi^*|$ is the deviation of last 4-period average inflation from target for AIT-4; $|\frac{1}{10}\sum_{t-4}^{t-1}\pi_j - \pi^*|$ is the deviation of last 10-period average inflation from target in AIT-10; $|P_{t-1} - P^*|$ is the deviation of price level from its target; $|NGDP_{t-1} - NGDP^*|$ is the deviation of nominal GDP from its target. μ_i is the subject fixed effect and $\epsilon_{i,t}$ are robust standard errors. The dependent variable is an indicator variable that takes the value of 1 if participant i in period t inflation exhibits policy consistent expectations. 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.

Table 6: History-dependence and complexity

Panel A: Forecast accuracy

	Preshock (Periods 1-19)		Shock (Periods 20-21)		Postshock (Periods 22-50)	
Abs.FE	Inflation	Output	Inflation	Output	Inflation	Output
IT	16.99	25.45	33.78	106.63	9.88	24.43
	(26.11)	(37.21)	(14.24)	(55.49)	(24.43)	(44.47)
DM	17.80	25.07	28.40	58.12	7.16	18.36
	(106.16)	(99.14)	(30.75)	(47.61)	(21.17)	(48.24)
AIT-4	21.10	33.47	28.53	81.87	16.30	27.53
	(55.13)	(77.97)	(14.96)	(53.12)	(77.97)	(38.39)
AIT-10	25.09	26.90	33.42	98.39	20.99	37.96
	(218.47)	(40.61)	(18.22)	(51.35)	(24.46)	(30.05)
NGDP	47.95	105.79	191.30	354.45	3.80E+23	3.80E+23
	(53.34)	(85.79)	(132.53)	(239.12)	(1.15E+25)	(1.15E+25)
PLT	54.99	86.92	108.60	266.84	8.53E+13	6.94E+12
	(119.57)	(96.09)	(58.39)	(122.08)	(2.83E+14)	(2.30E+14)
PLT Comm	17.64	39.31	50.34	130.12	19.10	40.66
	(45.68)	(68.39)	(132.62)	(142.61)	(34.70)	(48.40)

Panel B: History dependence and level targeting on forecast errors

	Preshock (Periods 1-19)		Shock (Periods 20-21)		Postshock (Periods 22-50)	
ln(Abs.FE)	Inflation	Output	Inflation	Output	Inflation	Output
HistoryDependence	0.009	0.016**	0.014	0.032	0.096***	0.102***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
LevelTarget	1.349***	1.457***	1.311***	0.843**		
	(0.11)	(0.09)	(0.27)	(0.36)		
Constant	2.015***	2.512***	3.189***	4.041***	1.633***	2.318***
	(0.09)	(0.07)	(0.08)	(0.14)	(0.08)	(0.06)
N	4443	4508	487	489	4344	4461
F	75.83	146.3	40.81	27.84	47.91	80.22

Panel C: History dependence and level targeting on forecast errors - PLT vs. NGDP

	Preshock (Periods 1-19)		Shock (Periods 20-21)	
ln(Abs.FE)	Inflation	Output	Inflation	Output
HistoryDependence	0.009	0.016**	0.014	0.032
	(0.01)	(0.01)	(0.01)	(0.02)
LevelTarget	1.298***	1.289***	1.054***	0.755**
	(0.19)	(0.14)	(0.26)	(0.37)
NGDP	0.103	0.337*	0.518*	0.178
	(0.21)	(0.17)	(0.28)	(0.31)
Constant	2.015***	2.511***	3.189***	4.041***
	(0.09)	(0.07)	(0.08)	(0.14)
N	4443	4508	487	489
F	77.19	104.8	37.24	21.28

Panel A presents means (standard deviations) of absolute inflation and output gap forecast errors across phases. All values are expressed in terms of basis points. Panel B presents results from OLS regressions evaluating the natural log of absolute forecast errors. $HistoryDependence_t$ indicates the number of past periods the central bank's policy is responding to. LevelTargeting is an indicator variable for the PLT and NGDP treatments. Level targeting treatments are excluded in the Postshock specifications because of instability at the ELB. Panel C compares forecast errors in the level targeting frameworks, where PLT is the baseline treatment and NGDP is an indicator variable for the NGDP treatment.

Online Appendix for History-dependent monetary policy – less is more [*]

NOT FOR PUBLICATION

Olena Kostyshyna
Bank of Canada

Luba Petersen
Simon Fraser University
NBER

Jing Yang
Bank of Canada

December 13, 2025

This appendix provides the following supplementary materials.

A. Instructions
B. Screenshot
C. Simulations under rational and naive expectations
D. Tables and figures
E. Heterogeneity in forecasting models
F. Additional experimental results in PLT Comm

[*] † Kostyshyna and Yang: Bank of Canada, 234 Wellington Street Ottawa, ON, K1A 0G9; okostyshyna@bank-banque-canada.ca, jyang@bankofcanada.ca; Petersen: Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada, luba_petersen@sfu.ca. The views expressed herein are those of the authors and not necessarily those of the Bank of Canada.

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:

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

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:

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

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:

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

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:

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

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:

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

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.)

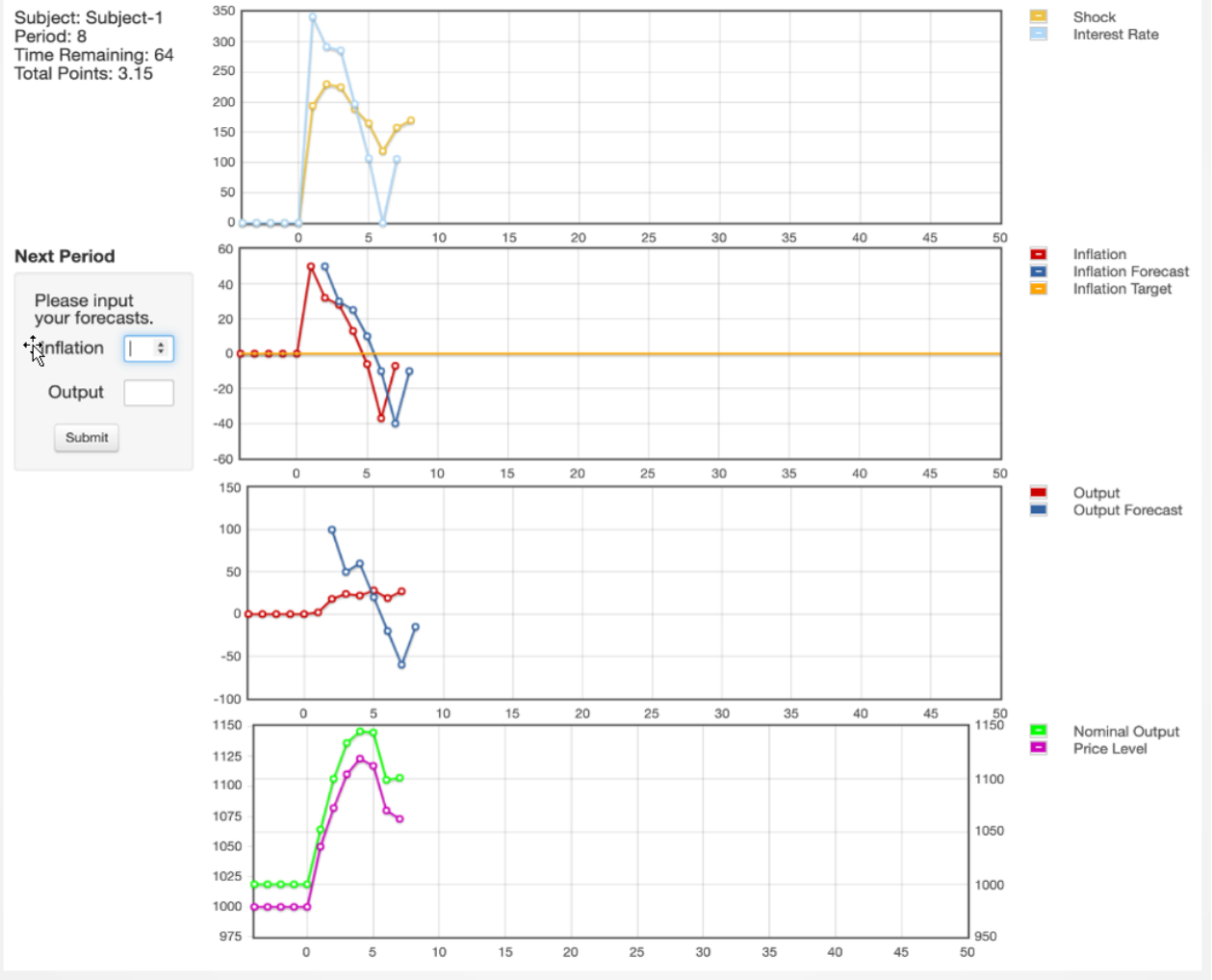
You may submit positive, negative or zero forecasts. Please use whole numbers. Please review your forecasts before pressing the SUBMIT button. Once the SUBMIT button has been clicked, you will not be able to revise your forecasts until the next period. You will earn zero points if you do not submit both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

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 non-rational expectations

Expectations of experimental participants have been shown to follow various mechanisms [**?**]. We examine the dynamics of a small New-Keynesian model under 6 most common forms of expectation formation listed in Table C1. In our simulations, we use the following parameter values: cognitive discounting $\alpha = 0.5$, constant gain $\gamma = 0.1$, and trend chasing $\tau = 0.1$.

Table C1: Models of expectations as functions of exogenous or historical data.

M1	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	$E_{i,t}x_{t+1} = x_{t-1} + \tau(x_{t-1} - x_{t-2})$
		$E_{i,t}\pi_{t+1} = \pi_{t-1} + \tau(\pi_{t-1} - \pi_{t-2})$
M6	Naive if $\tau = 0$	$E_{i,t}x_{t+1} = x_{t-1}$
	in trend chasing	$E_{i,t}\pi_{t+1} = \pi_{t-1}$

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

Figure C1: Simulations with rational expectations and heuristics in dual mandate (DM) and inflation targeting (IT).

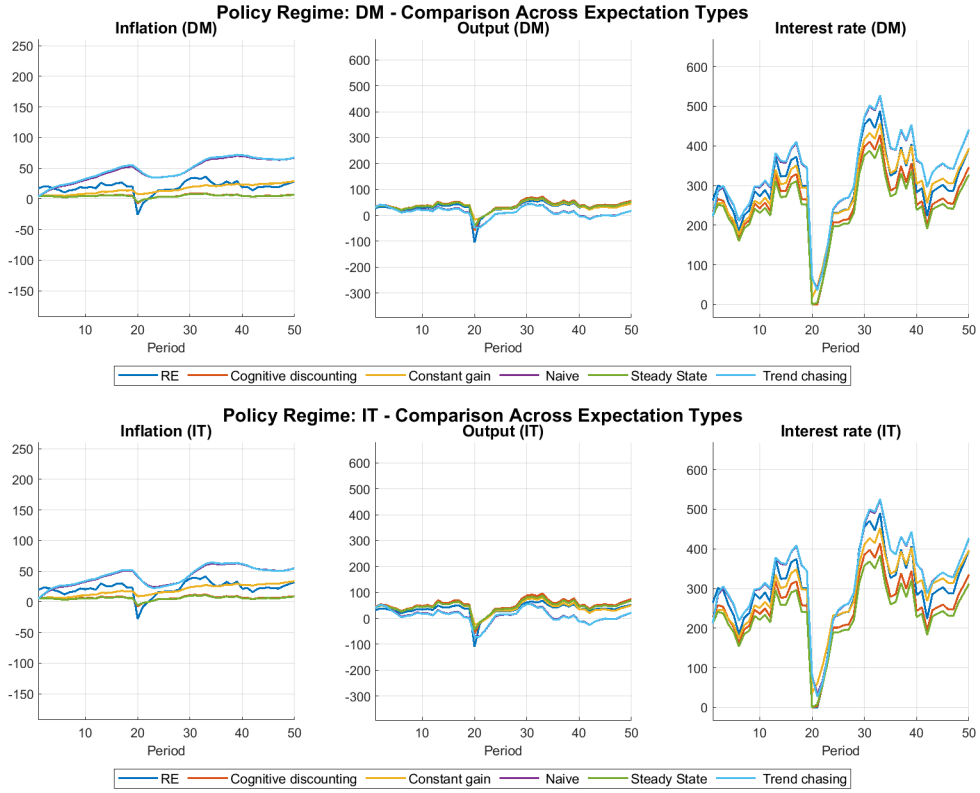


Figure C2: Simulations with rational expectations and heuristics in AIT-4 and AIT-10.

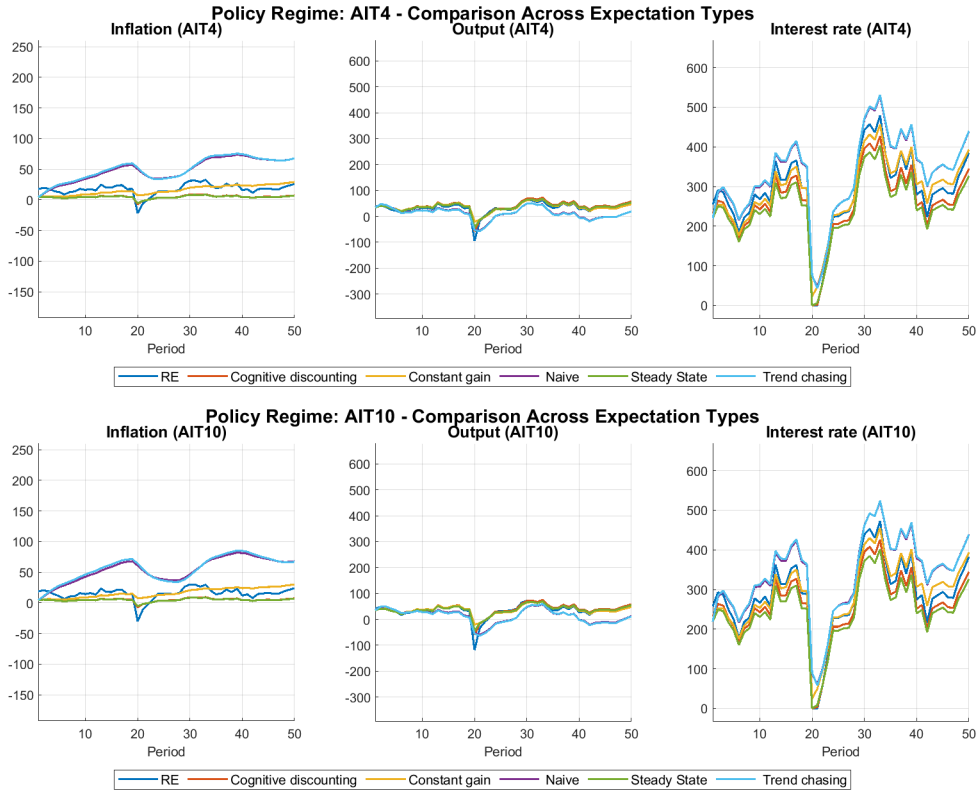
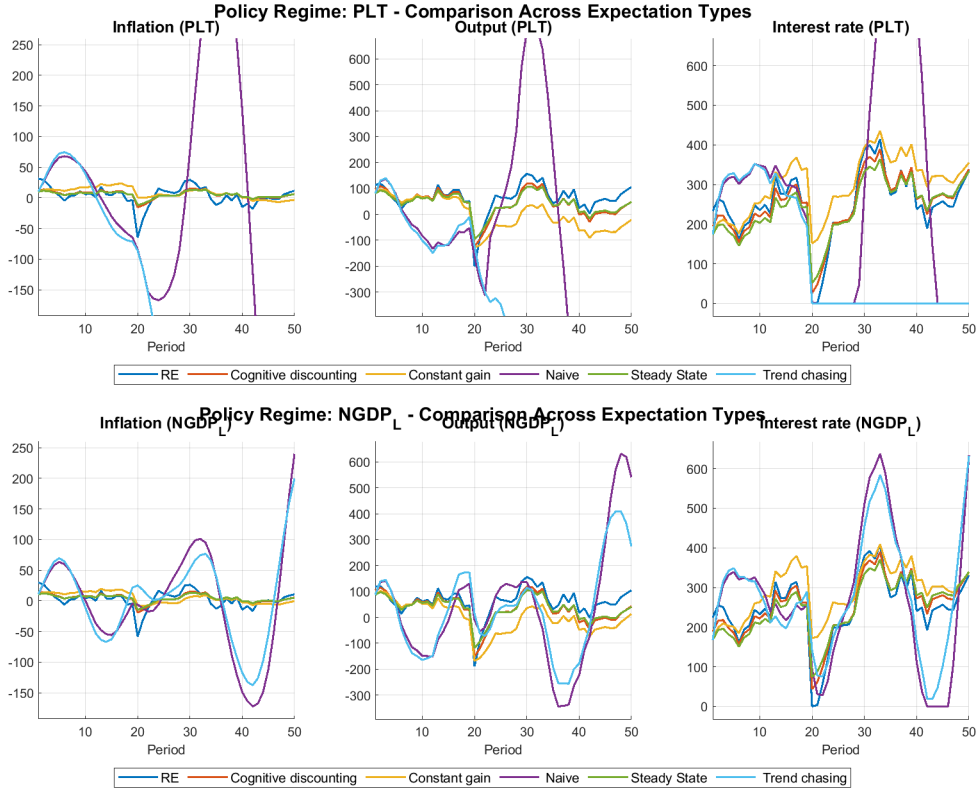


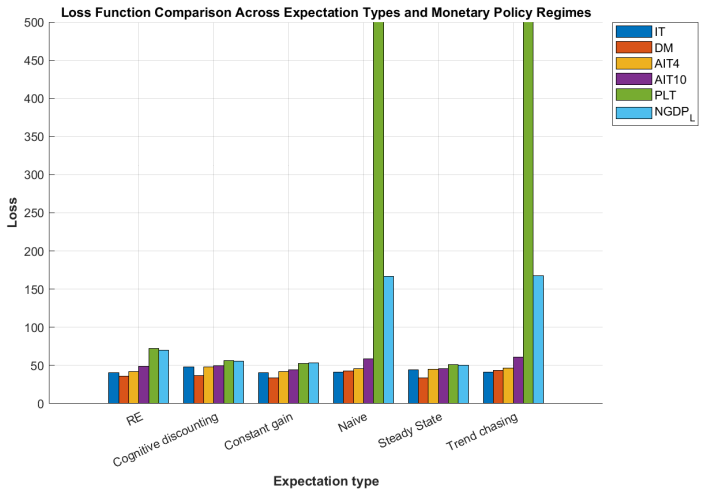
Figure C3: Simulations with rational expectations and heuristics in PLT and NGPP level targeting



We systematically evaluate the robustness of the ordering of treatment welfare losses to alternative parameterizations of the monetary policy rule. The results show that the baseline ranking is preserved when policy responsiveness is reduced by 25 percent (Figure C4), indicating that our main findings are not driven by an overly aggressive calibration of the policy coefficients.

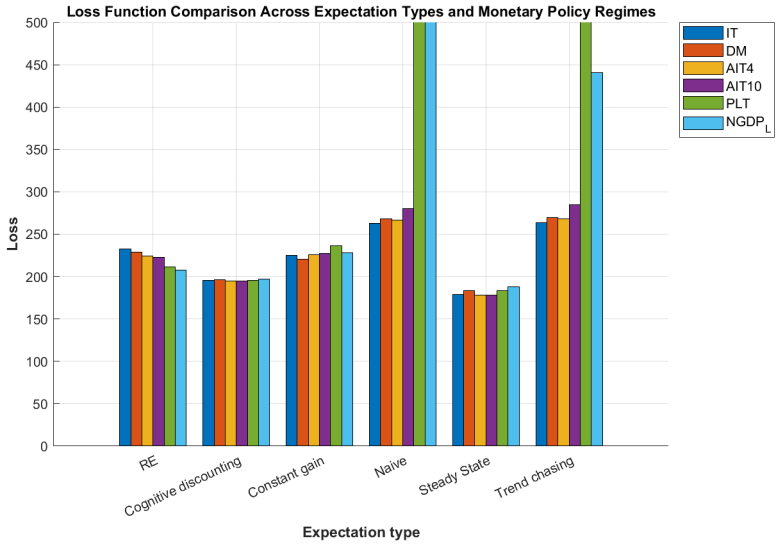
In contrast, increasing policy responsiveness by 25 percent (Figure C5) leads to a reordering of welfare outcomes across treatments. This asymmetry is intuitive. More aggressive policy responses disproportionately enhance stabilization under rate-targeting regimes, where monetary policy is not fully optimal and therefore has substantial scope for improvement, particularly in the presence of demand-driven fluctuations. By comparison, level-targeting frameworks such as PLT and NGDP targeting already approximate optimal policy by construction, implying that the marginal gains from further increases in responsiveness are necessarily limited. As a result, stronger policy coefficients compress welfare differences across treatments and can alter the ranking. Importantly, this does not reflect a lack of robustness, but rather highlights a predictable interaction between policy aggressiveness and the degree of policy optimality embedded in each regime.

Figure C4: Summary of losses from the simulations with higher policy coefficients



This figure presents the losses (equation 9) from the simulations of our model with different types of expectations over periods 1–50, **with policy coefficients 25% higher than in the baseline rules**. For the ease of comparison of losses across regimes and expectations, the maximum value for y-axis was set at 500. The losses for PLT with naive expectations are 615; for PLT with trend expectations are 628.

Figure C5: Summary of losses from the simulations with lower policy coefficients



This figure presents the losses (equation 9) from the simulations of our model with different types of expectations over periods 1–50, **with policy coefficients 25% lower than in the baseline rules**. For the ease of comparison of losses across regimes and expectations, the maximum value for y-axis was set at 500. The losses for PLT with naive expectations are 1,403,891; for PLT with trend expectations are 703,419; and for NGDP with naive expectations are 3,266, which exceed the maximum of y-axis.

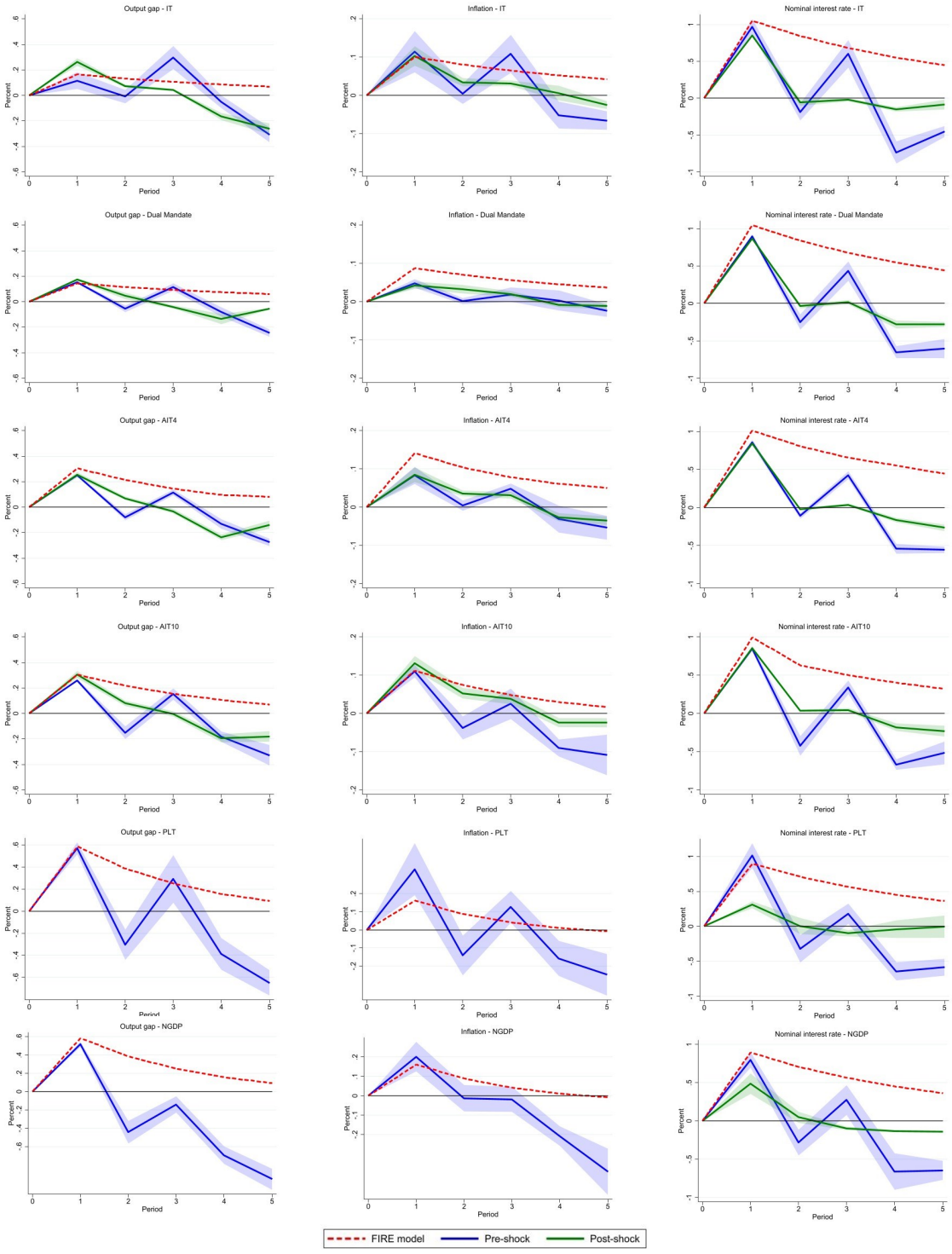
D. Tables and Figures

Table D1: 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 D1: Response of aggregate variables to a +1% demand impulse.

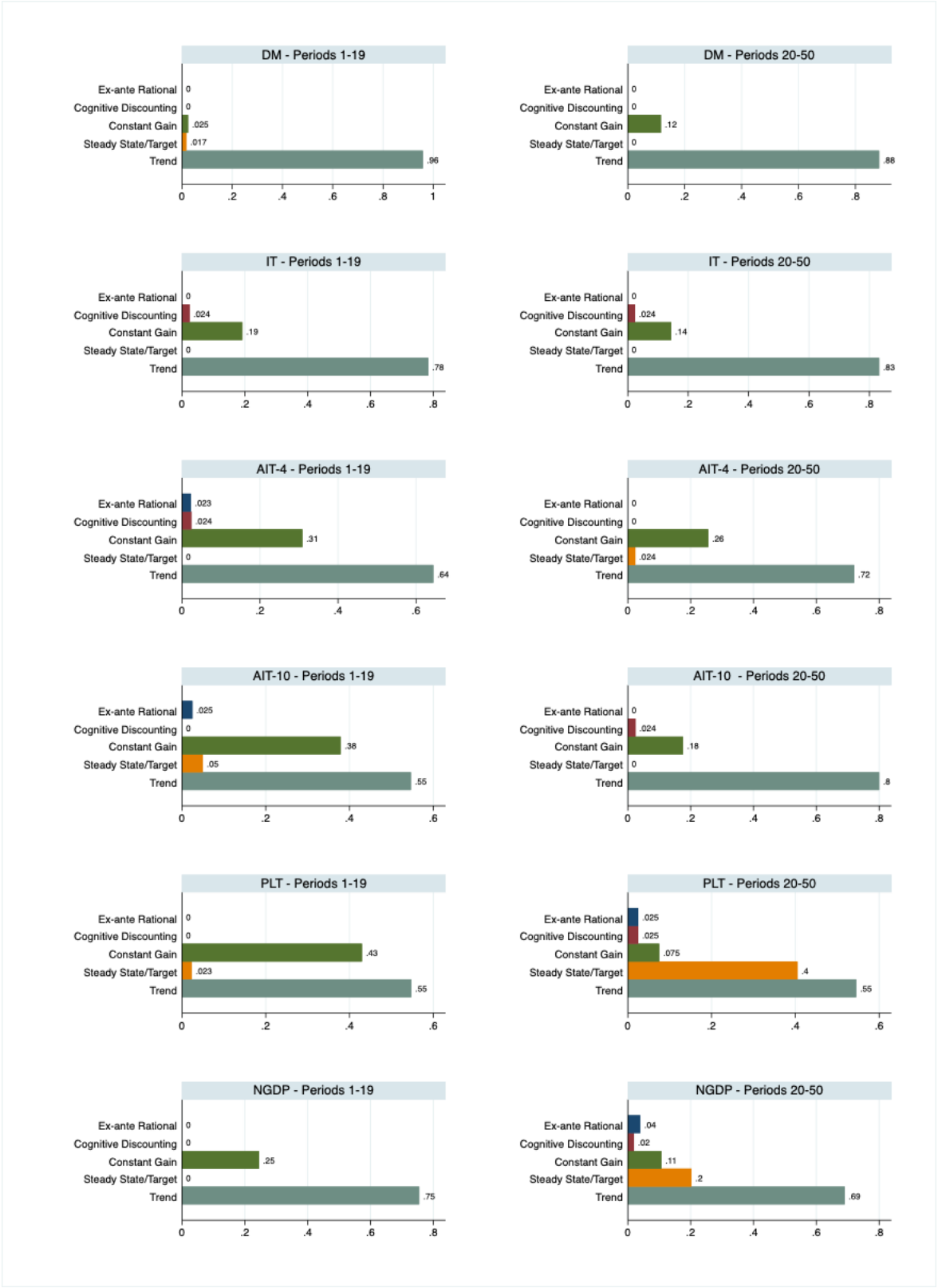


Notes: IRFs in the data are estimated by local projections. Shaded areas denote the 90% confidence interval. IRFs in the FIRE model are the simulated responses to a one-standard deviation shock.

E. Heterogeneity in forecasting models

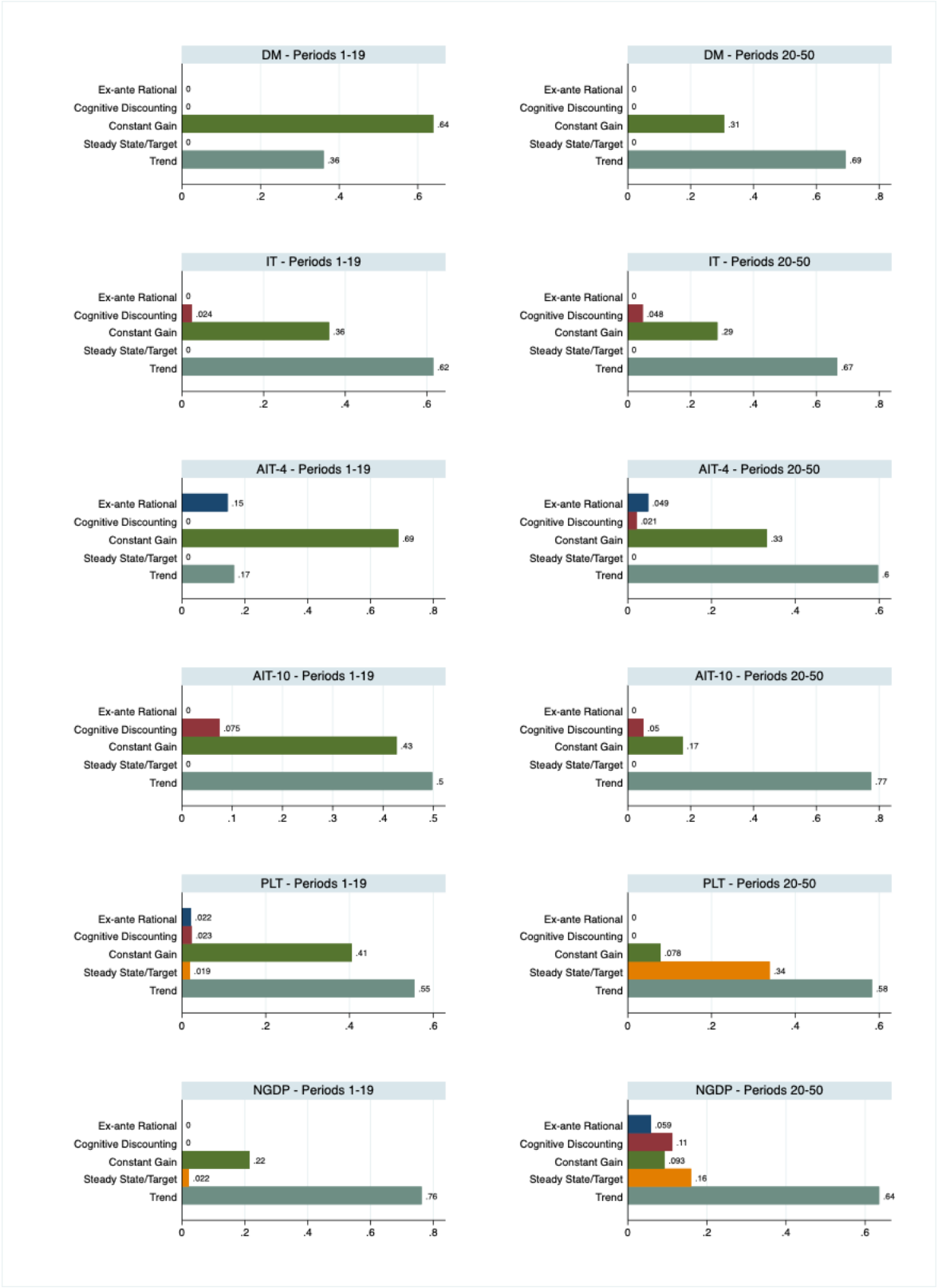
We analyze participants' forecast data to identify how they form expectations. Each participant is categorized based on the expectation formation type that best explains their forecasting behavior. We determine the forecasting model for each participant that best fits their forecasting behavior during 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 models presented in Table C1. For some models such as M2 Cognitive Discounting, M4 Constant Gain, 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 $[0, 1.5]$. Note that and M6 Naive expectations is a special case of trend chasing expectations when trend chasing parameter $\tau = 0$. We consider values of these parameters from these ranges with an increment of 0.1. We assign each participant the model and its parameter value (if applicable) that produces the lowest mean absolute error. Participants' forecasts are classified for two different phases of the experiment. The prevalence of the assigned models, by treatment and phase, are presented in Figure E1 for inflation forecasts and Figure E2 for output forecasts.

Figure E1: Distribution of forecasting models for inflation forecasts, by treatment and phase



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

Figure E2: Distribution of forecasting models for output expectations, by treatment and phase



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

i. Simulations with heuristics describing median expectations

Realized inflation and output in the experiments are determined by the *median* inflation and output expectations of participants (Section 2.2). For each session and regime, we classify median expectations in the pre-shock and post-shock periods according to one of the six heuristics listed in Table C1. The resulting classifications, reported in Table E1, reveal heterogeneity in heuristics across monetary policy regimes and some variation across the six sessions within each regime. For each regime, we summarize this distribution by identifying a *typical* heuristic—close to the median and/or mode across sessions—which is highlighted in bold in the table.

Using the typical heuristic and its estimated parameter values for the pre-shock and post-shock periods in each regime, we simulate the model and overlay the results (orange dotted lines) on the experimental data in Figures E3, E4, E5, and E6. These simulations reproduce key features of the experimental dynamics in inflation, output, and interest rates. In particular, they generate increasing volatility as monetary policy becomes more history-dependent, culminating in explosive dynamics under level-targeting regimes such as PLT and NGDP.

Although the simulated dynamics closely mirror those observed in the experiments, the levels of inflation, output, and interest rates can differ. These differences arise from the classification procedure: median expectations are assigned to the heuristic that minimizes the absolute distance between observed and predicted values. While this heuristic provides the best fit among the six considered, the match need not be exact, leading to level differences between simulated and experimental outcomes.

Table E1: Classification of median expectations

		Pre-shock (Periods 1-19)				Post-shock (Periods 20-50)			
		Inflation		Output		Inflation		Output	
Treatment			Parameter		Parameter		Parameter		Parameter
IT	**S1**	Trend	0.4	Trend	0.3	Trend	0.2	Trend	0.1
	S2	Trend	0.4	Trend	0.2	CG	1.1	Naïve	
	S3	Naïve		Trend	0.3	Naïve		Naïve	
	S4	Trend	0.6	Trend	0.7	Naïve		Naïve	
	S5	CG	1.1	CG	0.9	Naïve		CG	0.9
	S6	Trend	0.5	Trend	0.1	Trend	0.1	Naïve	
DM	S1	Trend	0.2	Naïve		Naïve		Naïve	
	S2	Trend	0.3	CG	0.9	Naïve		Naïve	
	S3	Trend	0.5	Naïve		Trend	0.3	Trend	0.1
	S4	Trend	0.4	Trend	0.1	Trend	0.2	Trend	0.1
	S5	Trend	0.3	CG	1.1	Trend	0.3	Naïve	
	S6	Naïve		Trend	0.2	Trend	0.3	Trend	0.1
AIT-4	S1	Naïve		CG	1.1	Trend	0.2	Naïve	
	S2	Naïve		CG	0.9	CG	1.1	CG	0.9
	S3	Trend	0.4	CG	0.9	Trend	0.1	CG	0.9
	S4	Naïve		CG	0.9	Naïve		Naïve	
	S5	Naïve		CG	0.9	Naïve		Naïve	
	S6	Naïve		CG	0.6	Trend	0.1	CG	0.9
					0.9				
AIT-10	S1	CG	0.9	CG	1.1	Naïve		Trend	0.3
	S2	CG	0.9	Trend	0.1	Naïve		Trend	0.2
	S3	Trend	0.4	Naïve		Trend	0.2	Trend	0.3
	S4	Trend	0.4	Trend	0.1	Trend	0.3	Trend	0.2
	S5	Naïve		Naïve		Trend	0.1	Trend	0.2
	S6	Naïve		CG	0.9	Naïve		Trend	0.3
	S6	Trend	0.5	Trend	0.4	CG	1.1	Trend	0.3
PLT	S1	CG	0.1	CG	0.9	CG	0.1	Target	
	S2	CG	1.1	CG	0.9	Naïve		Target	
	S3	Trend	0.4	Trend	0.3	Trend	1.5	Trend	1.5
	S4	Naïve		Trend	0.2	Trend	0.7	Trend	1.3
	S5	Naïve		Naïve		Naïve		CG	0.1
	S6	Trend	0.3	Trend	0.5	CG	0.4	CG	0.3
NGDP	S1	Trend	0.3	Trend	0.4	CG	0.9	Trend	0.9
	S2	Naïve		Trend	0.1	Trend	0.7	Trend	0.5
	S3	Trend	0.4	Trend	0.4	Trend	1.5	Trend	1.5
	S4	Trend	0.4	Trend	0.3	Trend	1.5	Trend	1.5
	S5	Trend	0.4	Trend	0.5	CG	1.2	CG	0.9

This table presents the results of classifying median inflation and output expectations for each session in each monetary policy regime.

Figure E3: Aggregate dynamics of inflation, output and interest rate with simulation using median expectation heuristic in dual mandate (DM) and inflation targeting (IT) treatments

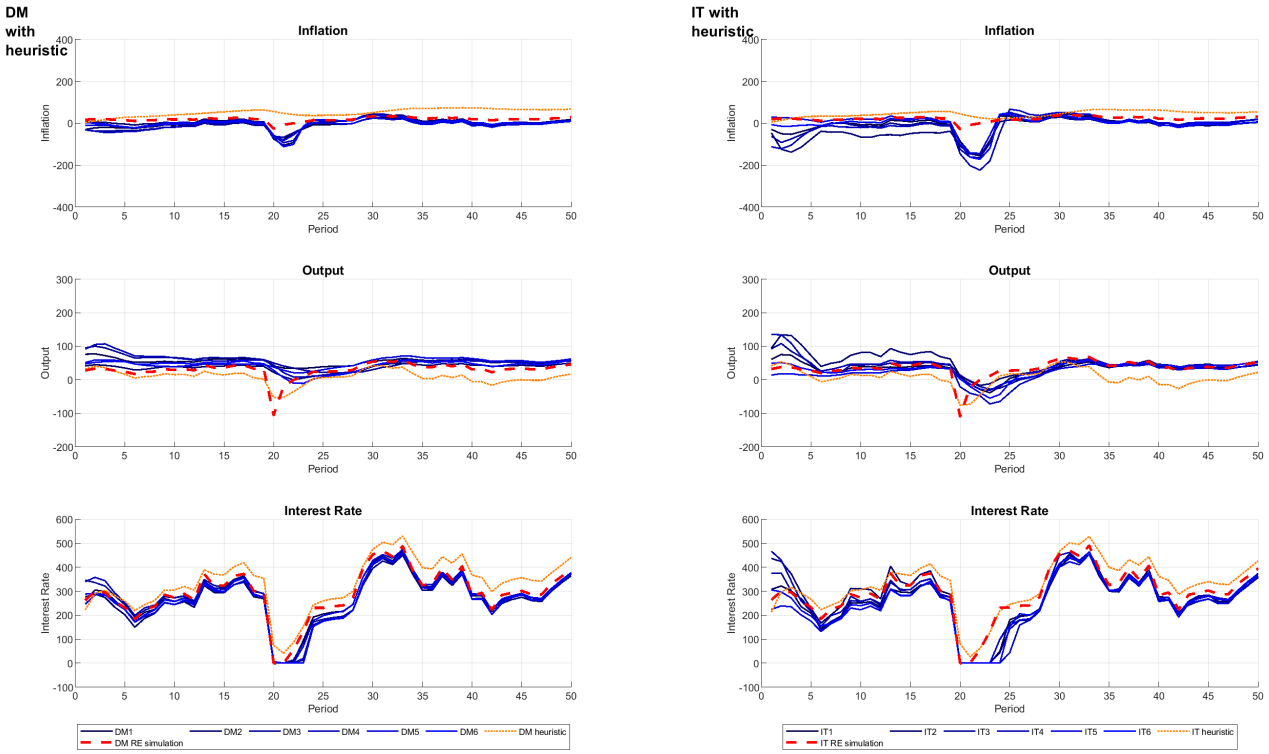


Figure E4: Aggregate dynamics of inflation, output and interest rate with simulation using median expectation heuristic in AIT-4 and AIT-10 treatments

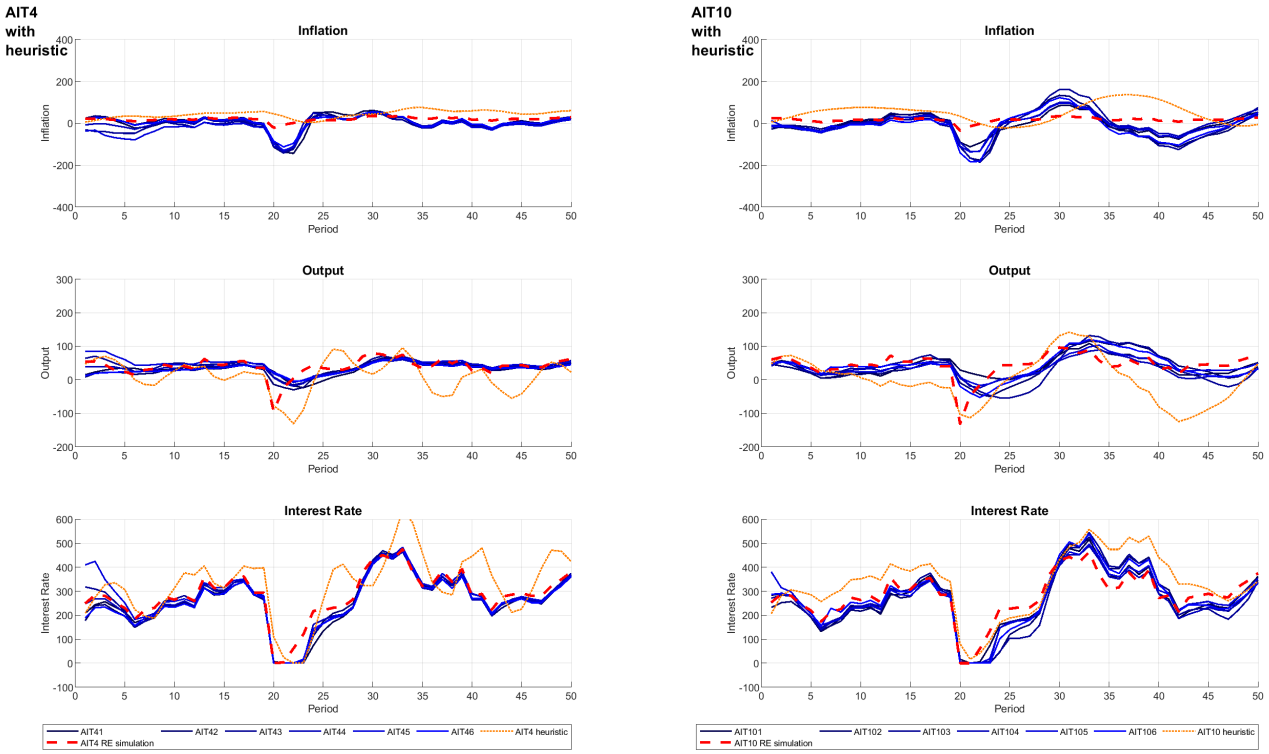


Figure E5: Aggregate dynamics of inflation, output, and interest rate with simulation using median expectation heuristic in price-level targeting (PLT) treatment

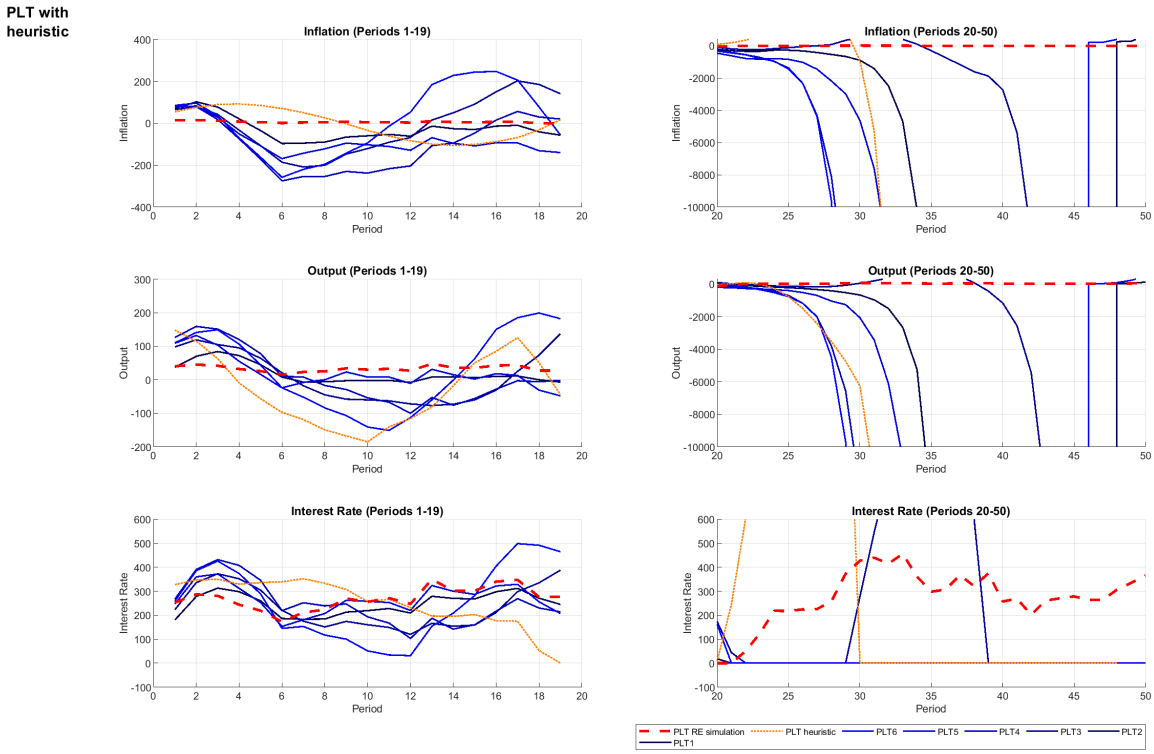
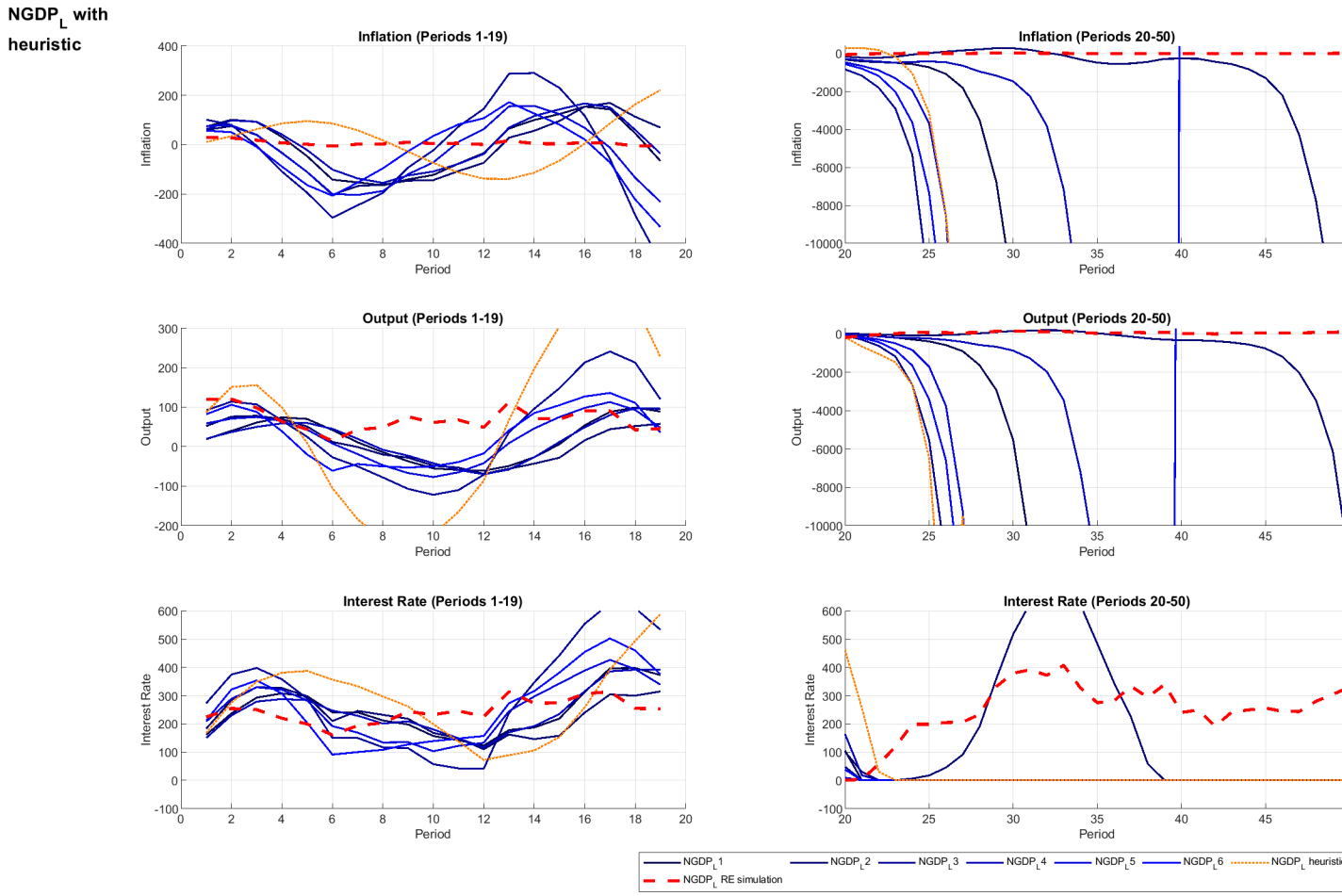


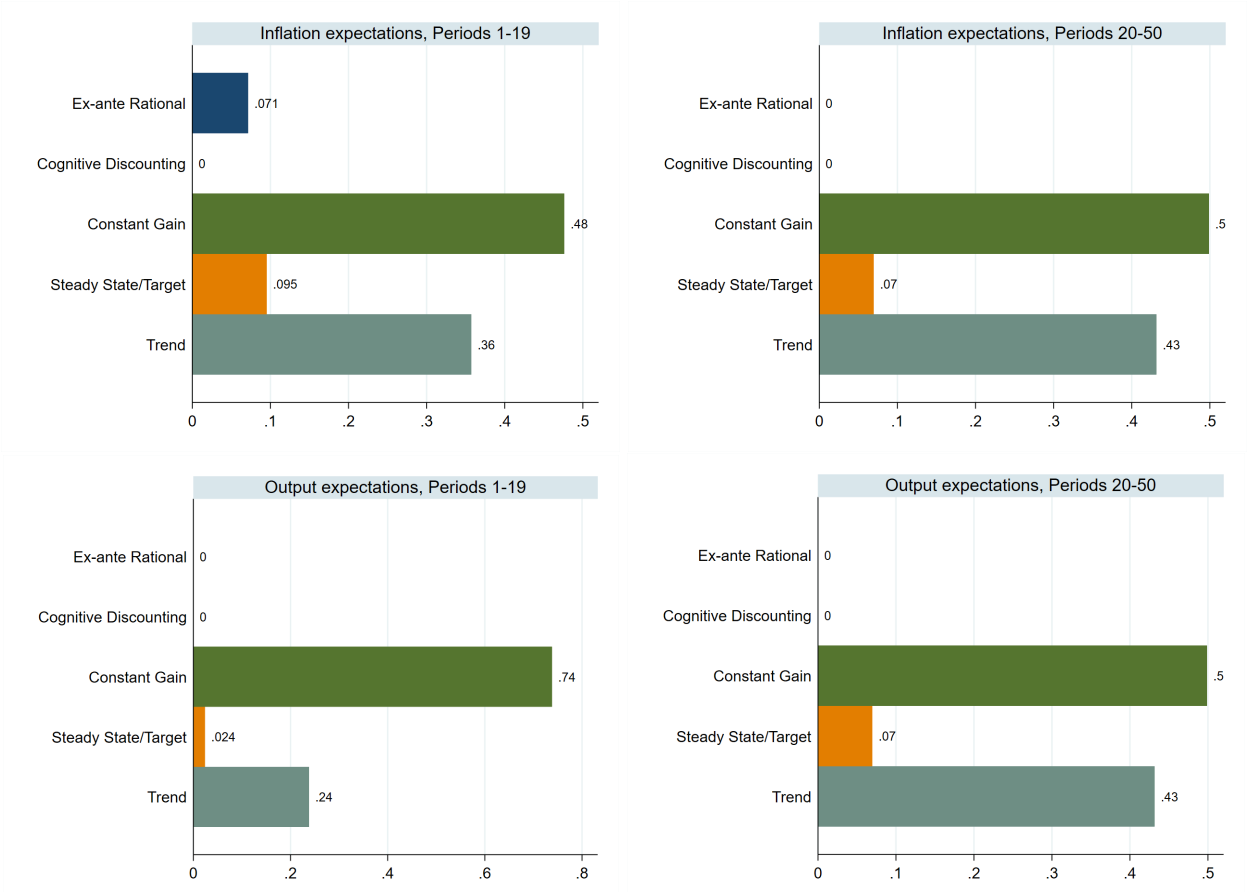
Figure E6: Aggregate dynamics of inflation, output, and interest rate with simulation using median expectation heuristic in NGDP level targeting (NGDP) treatment



F. Additional experimental results in PLT Comm

Figure F1 presents the distribution of forecasting models for inflation and output gap forecasts in PLT Comm.

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

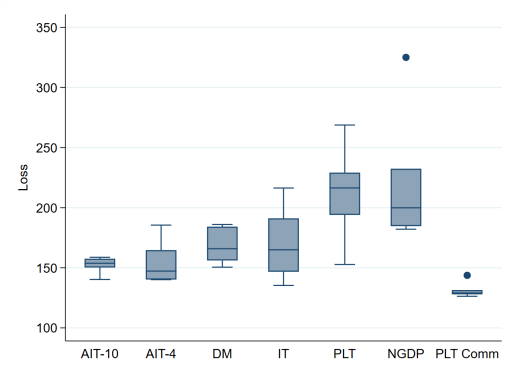
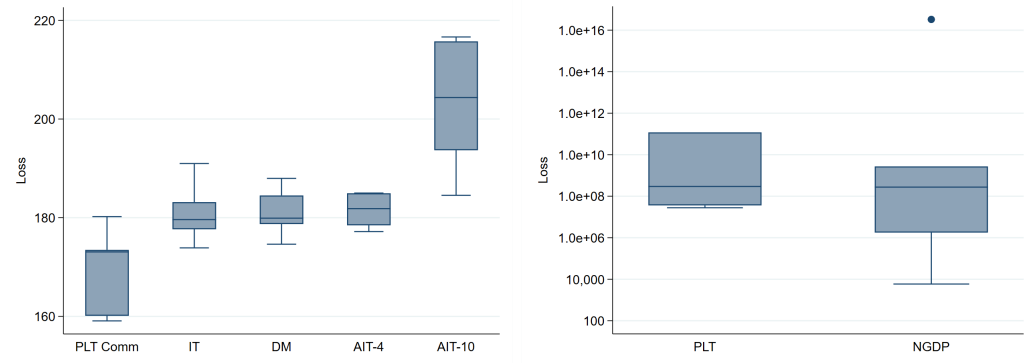


This figure presents the share of participants in each phase classified into a given model.

Figure F2 shows losses in the PLT Comm treatment relative to the original set of treatments. Losses in PLT Comm, excluding unstable session 5, 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 F2: Losses including PLT Comm treatment

Periods 1-19



Periods 20-50