History-dependent monetary policy – less is more*

Olena Kostyshyna
Bank of Canada

Luba Petersen
Simon Fraser University
NBER

Jing Yang
Bank of Canada

January 6, 2025

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 show that monetary policy with less history dependence is more stabilizing contrary to what is expected in theory. Our analysis of experimental data identifies that a weakened expectations channel undermines HDMP. Expectations of experimental participants feed off the past data, they are not forward-looking and do not internalize the stabilization properties and future *make-up* strategy of HDMP. We identify two risks associated with HDMP: inflation overshoots in average inflation targeting and deflationary spirals in 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 frameworks.

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

* 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 key role of expectations channel in the efficacy of history-dependent regimes, we study expectations of our experimental participants. In these experiments, participants are tasked with repeatedly making 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 the experiment. The

[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.

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

Our results demonstrate that 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. While frameworks with little to no history-dependence—such as inflation targeting—restore stability relatively quickly, level-targeting regimes like price-level targeting (PLT) and nominal-GDP targeting (NGDP) tend to spiral into deflationary trends at the ELB. 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.

We identify that the failure of the expectations channel plays a crucial role in undermining the performance of history dependent monetary policy frameworks. In our experimental economies, expectations come directly from the participants, allowing us to use rich individual-level data to uncover the sources of this failure. First, rather than relying on fundamentals and anticipated monetary policy actions, participants in our experiments tend to 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. Instead, expectations are predominantly backward-looking and trend-chasing, driven by observed dynamics rather than forward-looking assessments. This backward-looking nature of expectations weakens the expectations channel in history-dependent frameworks by amplifying current trends rather than counteracting them through forward-looking behavior.

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 present 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. Inertial monetary policy, combined with backward-looking expectations, leads to higher volatility and stronger persistence in both inflation and output, with these effects intensifying as the AIT horizon lengthens. As expectations feed off observed economic dynamics, they prolong the impacts of both negative and positive shocks. 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, as monetary policy remains "low for longer",and inflation expectations respond to rising inflation, positive inflation overshooting occurs. The inertia in monetary policy meets inertia in expectations formation resulting in more volatility and higher persistence of inflation and output gap. This positive overshooting, following a period of below-target inflation, may be seen as consistent with the post-pandemic inflation surge in the United States during 2021–2022, following the introduction of an average inflation targeting mandate in 2020, although supply and demand factors, including supply chain disruptions, have played a role [Gordon and Clark, 2023].

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. As expectations react to observed deflation, expectations become deflationary and unanchored. The deflationary episode after 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 this spiral unfolds, participants increasingly engage in trend-chasing behavior, further reinforcing deflationary expectations and worsening the spiral. 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 such a 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 and addresses the cognitive challenge posed by framing PLT in terms of the price level.

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] suggest limitations to these frameworks. Our findings align with the bounded rationality literature, though our results indicate 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. Additionally, we identify two risks in the implementation of history-dependent monetary policy regimes: 1) inflation overshooting under AIT, a phenomenon that has not been discussed in the literature, and 2) deflationary spirals in PLT and NGDP level targeting at the ELB. 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 previously documented by Arifovic and Petersen [2017] and Hommes et al. [2019a], the possibility of deflationary spirals under NGDP level targeting is a novel finding in the literature.

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.

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 designed around a simple New Keynesian model that is commonly used for monetary policy analysis. We construct a laboratory economy that follows a data-generating process based on the canonical model [Woodford, 2003]. Similar monetary policy environments have been used in Adam [2007], Pfajfar and Žakelj [2014], and Assenza et al. [2021].

The economy is described by the following system of equations:

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

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

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

Equation 1 describes the evolution of inflation in period t, π_t, in response to aggregate one-period-ahead inflation expectations, $\mathbb{E}_t \pi_{t+1}$, and the output gap, deviation of output from its steady state level, x_t. The output gap, given by Equation 2, is a function of aggregate expectations of one-period-ahead inflation and output gap expectations, $\mathbb{E}_t x_{t+1}$, as well as the deviations of the nominal interest rate, i_t, from the natural rate of interest, r_t^n. The natural rate of interest, described by Equation 3, is the rate of interest that keeps the economy at full employment while keeping inflation constant. The natural rate of interest is assumed to follow an AR(1) process and is subject to a sequence of demand shocks, ϵ_t. Parameters in our model are calibrated to quarterly data, as in Dorich et al. [2021] and Kryvtsov and Petersen [2021], and are consistent with Canadian data: β =0.994, $\sigma = 1$, $\rho = 0.8$, $\sigma_{rn} = 0.005$, $\kappa = 0.125$, π^* =0, x^* =0, r_t^{n*} =0, $\bar{r} = i^* = 60$.

We have taken measures to simplify the experimental macroeconomy for participants through our 1) choice of data generating process (DGP); 2) choice of shocks; and 3) linearization of the economy around a zero-inflation steady state. When we use the simple three-equation NK model as our experimental DGP, some of the assumptions required for the log-linear approximation given in Equation 1-Equation 3 may not always hold in the experiment. Expectational errors may not be small and unbiased, and as emphasized by Preston [2005], the micro-founded New Keynesian framework would produce meaningfully different dynamics with heterogeneous expectations. However, the version of the NK model with heterogeneous expectations has been shown to be less stable in the laboratory [Mauersberger, 2021] than the standard model. Thus, our choice in favour of the standard model was made with this in mind.

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

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

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

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 reason 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

Parameters in the policy rules are obtained from optimizing the following loss function as implemented by Dorich et al. [2021]:

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

While this loss function is ad hoc, it provides a realistic description of the goals pursued by a central bank. Central banks are concerned not only about inflation and output gap stabilization but also interest rate variation. The coefficients in the policy rules in different monetary policy regimes were chosen to minimize this loss function while putting the frameworks on comparable footing, and are in line with the theoretical horse race conducted in Dorich et al. [2021].

2.2 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. We used the median, rather than the mean, to capture aggregate expectations and reduce the influence of outliers, helping to stabilize dynamics in the presence of an ELB and potentially boundedly rational participants.

Participants provided one-period-ahead expectations for inflation and output. Focusing on short-term expectations has several advantages. First, the largest predicted gains from history-dependent mandates come from stabilizing short-term expectations [Walsh, 1998]. Short-term expectations instead provide insight into the strength of the expectations channel. Second, this approach simplifies the forecasting task (requiring repeated forecasts for only two variable), while maximizing the number of forecasting periods and allowing ample time to process information across different types of scenarios.

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. As is standard in learning-to-forecast experiments, participants did not have information about each others' expectations, and 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.

2.3 Procedures

We implement six independent sessions for each monetary policy treatment. For each session, we invited a group of seven inexperienced participants to play the roles of professional forecasters tasked with making forecasts in 50 sequential periods.[3]

The exogenous shocks in the experimental economy were pre-drawn, as explained in the participant instructions. The shock sequence was chosen to create two distinct phases in the experiment. Each session began with a stable phase (periods 1–19), allowing us to evaluate different policy mandates away from the effective lower bound (ELB). 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 50 cents per point.

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 exper-

[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.

iment: 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-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 and Texas A&M University, 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.[4]

3 Performance under rational and naive 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 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

[4] 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. They received a $7 show-up fee plus performance-based pay, averaging $25 total, with payments via e-Transfer in Canada and Venmo in the U.S.

total loss as the square root of total loss (Equation 9) divided by 50 periods. The results are summarized in Table 1. A detailed breakdown of the total loss, showing the individual contributions from deviations in inflation, output, and interest rates from the steady state, is provided in Table C1 in Online Appendix C.

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.

Other experimental studies of monetary policy regimes illustrate that participants' expectations are mostly non-rational [Anufriev et al., 2013, Assenza et al., 2021, Hommes and Makarewicz, 2021b]. Given this evidence, we introduce a very simple form of adaptive expectations – naïve expectations – into our model to understand the implications for stabilization properties of different monetary policy regimes. Naïve expectations are set as $E_t\pi_{t+1} = \pi_{t-1}$ and $E_t x_{t+1} = x_{t-1}$. We find that the presence of naïve agents can be disruptive to economies with certain monetary policy regimes. Level-targeting regimes such as PLT and NGDP can break down for certain shares of naïve agents. The threshold share of naïve agents is 33% in the PLT regime and 45% in NGDP; economies become unstable in these regimes with shares of naïve agents above the threshold level. IT, DM, and AIT tolerate 100% of naïve agents, remaining stable. In other words, PLT and NGDP are the least robust to the presence of naïve expectations.

We have simulated our model with different shares of naïve expectations using a sequence of demand shocks implemented in the experiment. We have evaluated the loss function (Equation 9) for the model with rational expectations (REE) and for the models with adaptive expectations which are presented on Figure 2 and in Table 1 (more details are discussed in Appendix C).

For a small share of naïve agents (33%), NGDP level targeting performs better than other regimes (Figure 2). However, as the share increases to 45% , IT and DM turn out to be more effective and more robust to the presence of non-rational expectations in their ability to stabilize the economy and perform better than history-dependent regimes NGDP, AIT-10 and AIT-4. With 100% of adaptive expectations, IT and DM perform better than AIT-10 and and AIT-4.

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 3), AIT-4 and AIT-10 (Figure 4), PLT (Figure 5), and NGDP (Figure 6). 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 stabilizing 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 (Figure 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 (Figure 5). 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 (Figure 6). 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 7 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:

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

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 8, 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 8). 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 8, we observe several notable differences between the expectations of our 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 4.

And finally, most of impulse response functions for inflation and output expectations exhibit

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

oscillatory behaviour indicative of backward-looking expectations, i.e. expectations the feed off the observed dynamics [Pfajfar and Žakelj, 2014, Kryvtsov and Petersen, 2013]. In none of the monetary policy framework, expectations respond monotonically to the demand shock as rational expectations do, suggesting important deviations from rationality in expectations of our experimental participants.

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 6.

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 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, suggesting behavioral inertia rather than responsiveness to deviations from target. This inertia intensifies in the post-shock period, implying that participants' prior belief formation plays a more significant role than the central bank's performance in shaping expectations.

In addition to the limited share of participants forming policy-consistent forecasts, even these forecasts do not adjust sufficiently to reflect current economic fundamentals. While policy-consistent expectations are closer to the rational expectations solution than policy-inconsistent ones, they still fall short of fully internalizing the stabilization properties of monetary policy (Figure 9 and Figure 10).

For the limited share of participants forecasting in a policy-consistent manner, their expectations still do not adjust sufficiently to reflect current economic fundamentals. While these expectations align more closely with the rational expectations solution than policy-inconsistent ones, they fall short of what would be expected if participants had fully internalized the stabilization properties of monetary policy (Figure 9 and Figure 10).

These characteristics of participants' expectations suggest that the expectations channel of monetary policy is significantly weaker than what is assumed by models that advocate for history-dependence in monetary policy. Not only is the share of model- and policy-consistent expectations small, but even these fail to fully capture the quantitative effects of monetary policy. In other words, participants' expectations do not adequately internalize future monetary policy actions or their anticipated impact.

5.3 Mechanisms of expectations formation

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 find 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 11, 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 several expectation models commonly identified in surveys and applied in the estimation of DSGE models [Milani, 2012]: ex-ante rational or model-consistent expectations [Muth, 1961, Sargent and Wallace, 1975], cognitive discounting [Gabaix, 2020], constant-gain learning [Branch and Evans, 2006], anchoring on targets [Coibion et al., 2018], and extrapolative trend-chasing [Frankel and Froot, 1990, Bordalo et al., 2020]. 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]

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

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

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 12). 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].

As previously discussed, following the large demand shock, participants in PLT and NGDP regimes are less likely to form expectations in the direction intended by monetary policy. Instead of relying on the make-up strategies inherent in level-targeting frameworks, participants respond more strongly to recent trends, attempting to catch up with observed economic dynamics. This behavior accelerates the unraveling of deflationary spirals. Strong extrapolation of past deflationary trends into future further destabilizes these economies, leading to explosive deflationary dynamics in PLT and NGDP regimes when monetary policy is constrained by the ELB.

Standard macroeconomic theory suggests that the expectations channel in PLT should provide the necessary lift during ELB episodes by raising inflation expectations in anticipation of higher future inflation delivered through the make-up strategy embedded in the framework. However, our experimental findings reveal that participants' expectations fail to internalize this forward-looking anticipation. Instead, participants exhibit *backward-looking* and *trend-chasing* behavior, extrapolating observed dynamics into the future. This reactive nature of expectations undermines the expectations channel in history-dependent frameworks, as it amplifies current trends rather than counteracting them through forward-looking behavior.

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.

In PLT, the most history-dependent regime, participants face an additional challenge: translating deviations of the price level from its target into the inflation rate required to return to the target. However, in terms of the historical dimension, PLT may be less cognitively demanding than AIT-10 because the most recent price level deviation encapsulates all past inflation deviations from the target. This reduces the need for participants to review an extensive history of past inflation 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 13.

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 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

[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.

[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.

(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 11 and Figure 12), 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 under 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 strong response of 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. The combination of constrained monetary policy and entrenched deflationary expectations causes 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. 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 forecasts 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 14. 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 6). 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 challenge the previously suggested potency 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 history-dependent frameworks. Providing *relevant* macroeconomic projections—such as inflation and output gap forecasts—helps simplify complex frameworks like 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.

References

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.

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 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.

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.

Matthew V. Gordon and Todd E. Clark. The impacts of supply chain disruptions on inflation. *Economic Commentary, FRB of Cleveland*, (2023-08), May 2023. doi: 10.26509/frbc-ec-202308.

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.

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.

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

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

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

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.

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

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

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

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

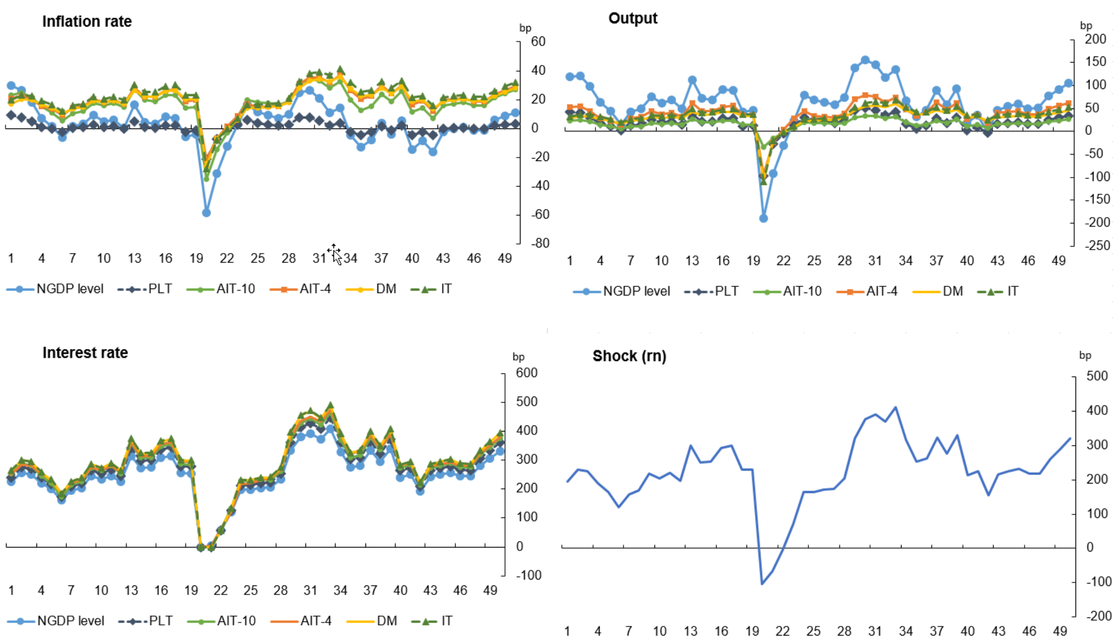
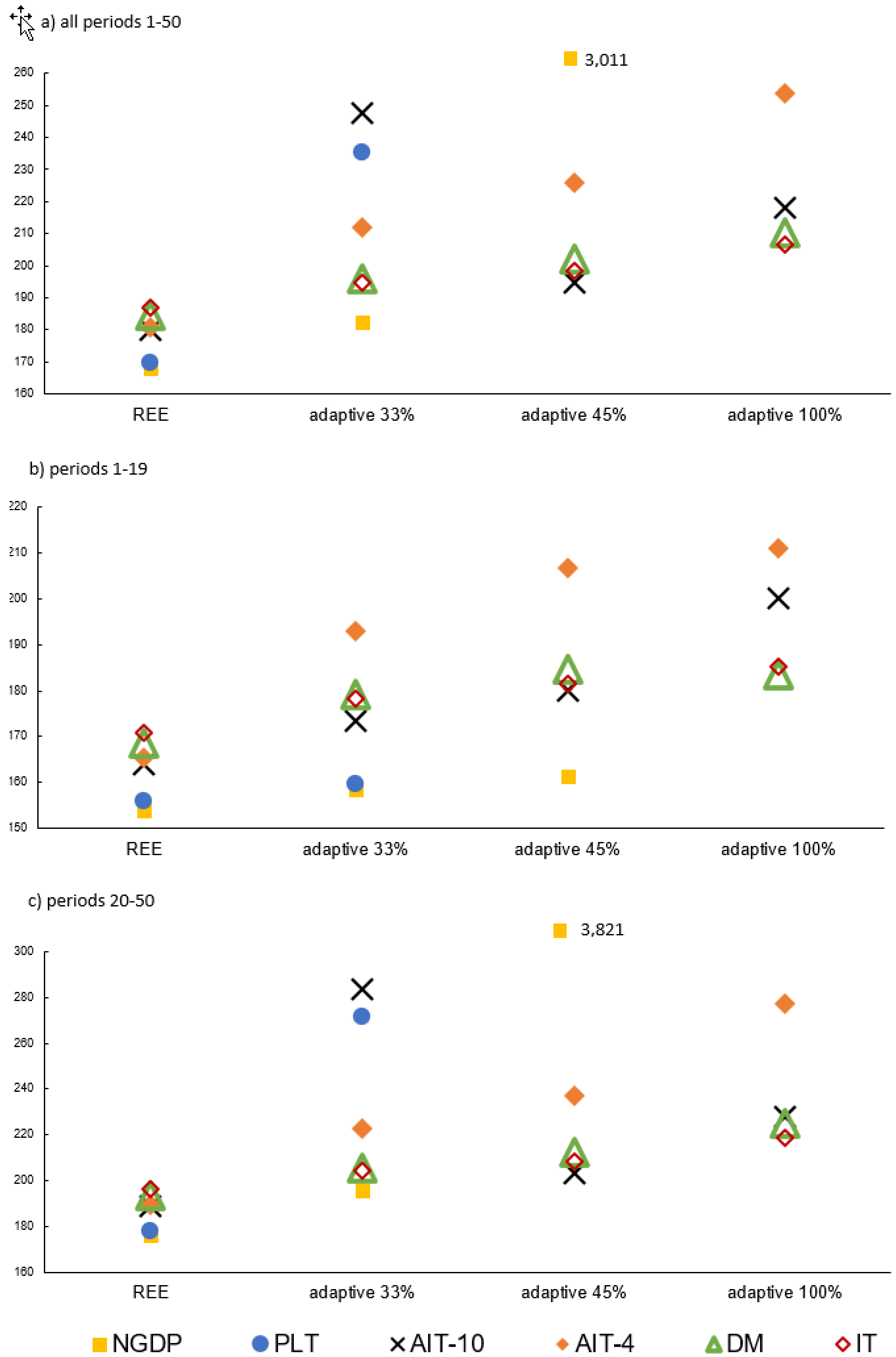


Figure 2: Summary of losses from simulations



This figure presents simulation results from the New Keynesian model under rational expectations (REE) and with varying shares of naïve agents. The proportion of naïve agents ranges from 0% (REE) to 33% (threshold in PLT), 45% (threshold in NGDP) and 100%.

Figure 3: Aggregate dynamics of inflation, output and interest rate in dual mandate (DM) and inflation targeting (IT) treatments

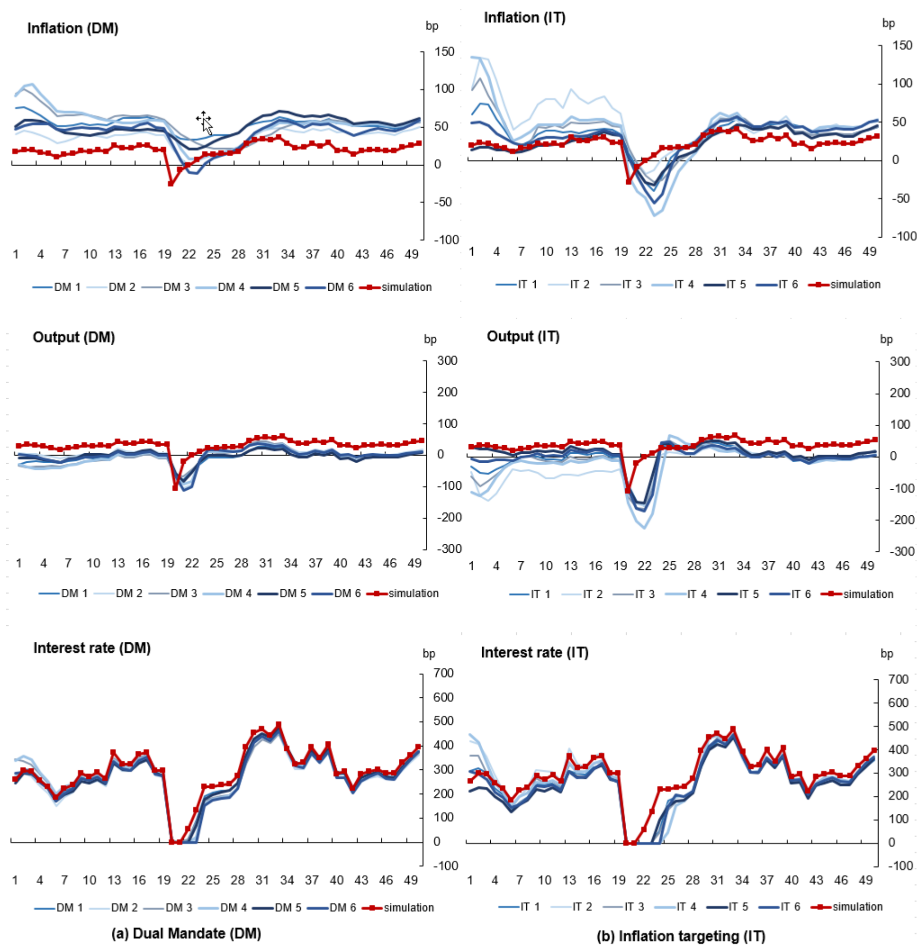


Figure 4: Aggregate dynamics of inflation, output, and interest rate in average inflation targeting (AIT) treatment

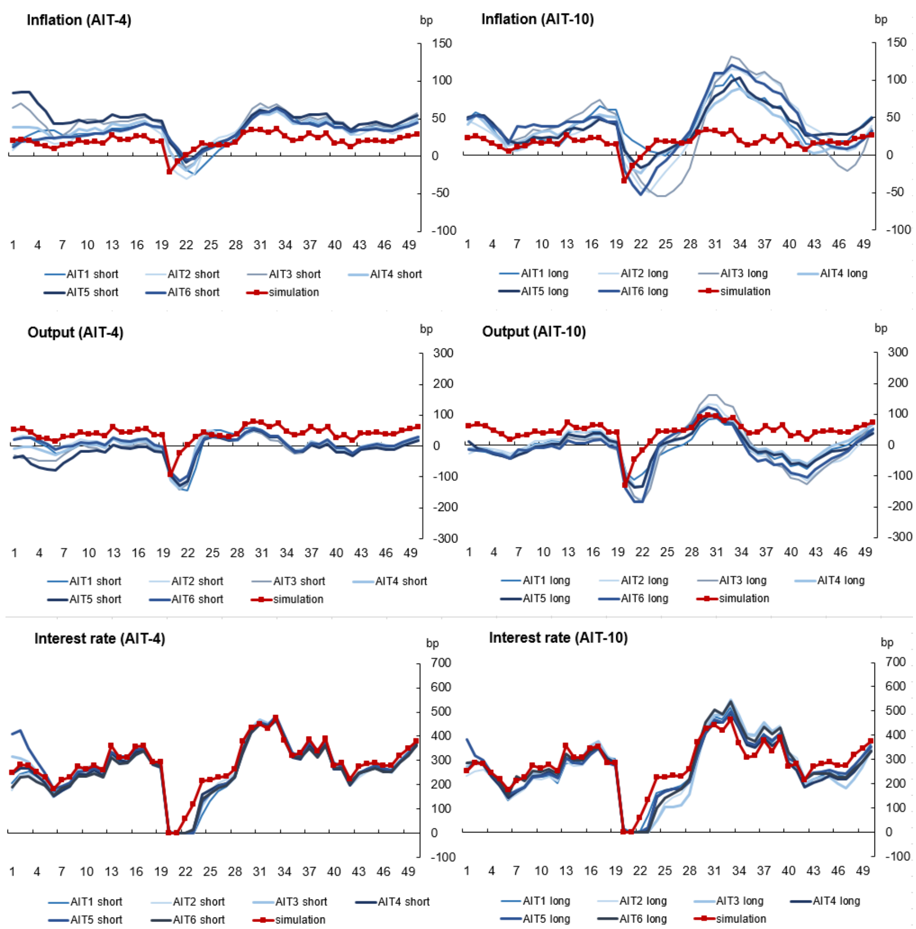


Figure 5: Aggregate dynamics of inflation, output, and interest rate in price-level targeting (PLT) treatment

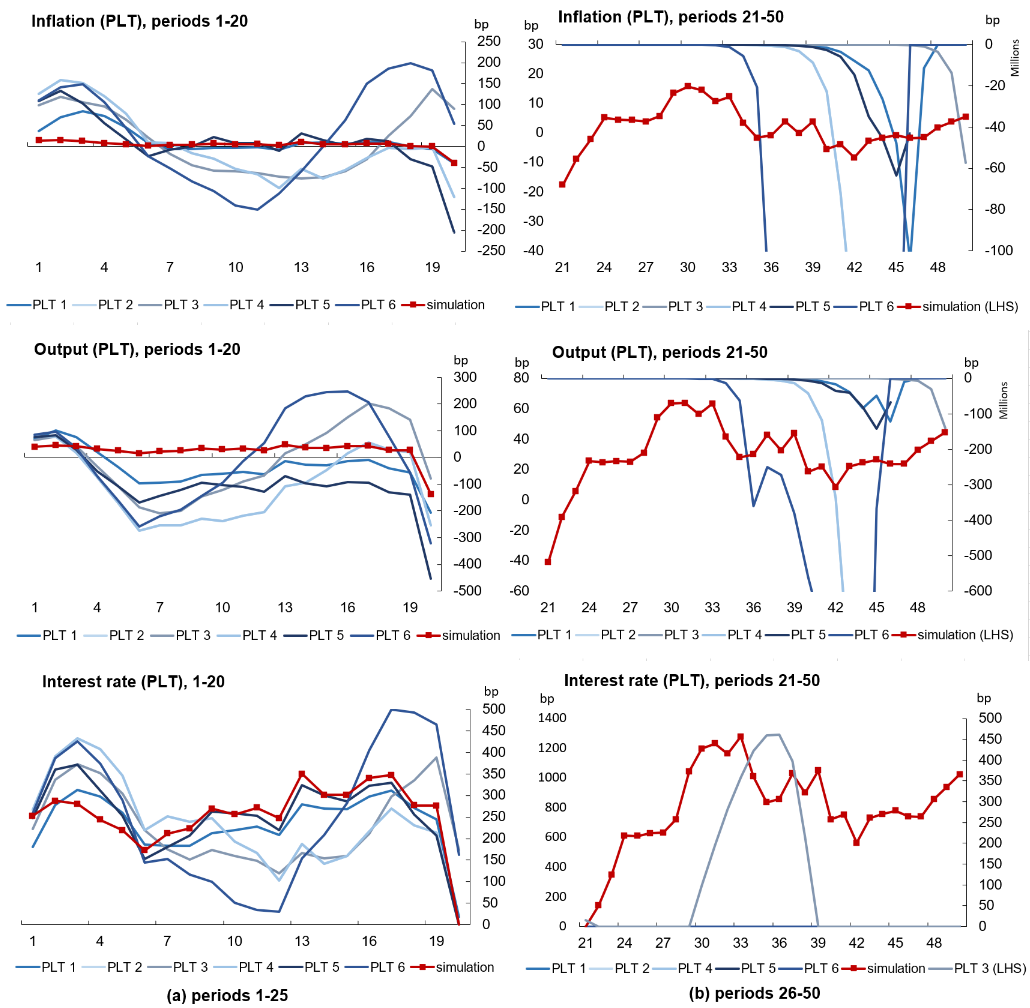
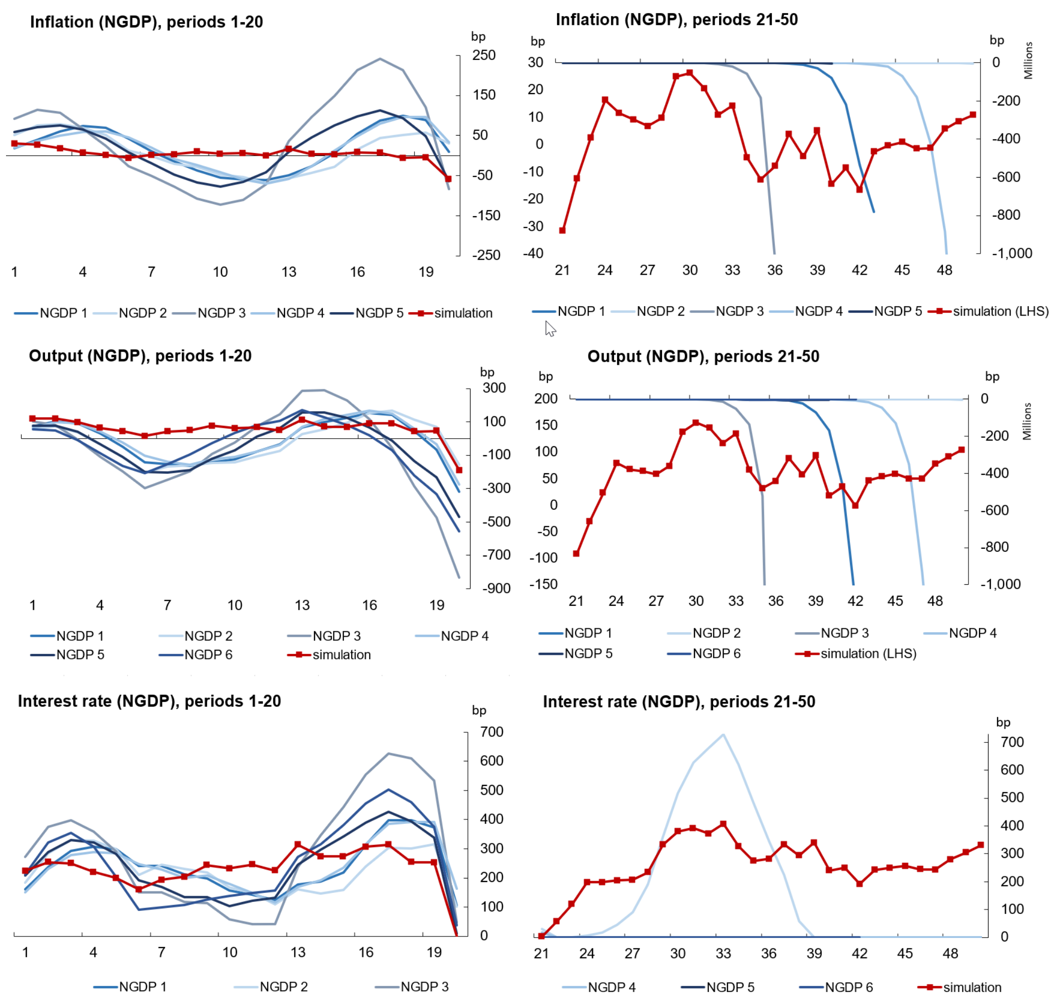


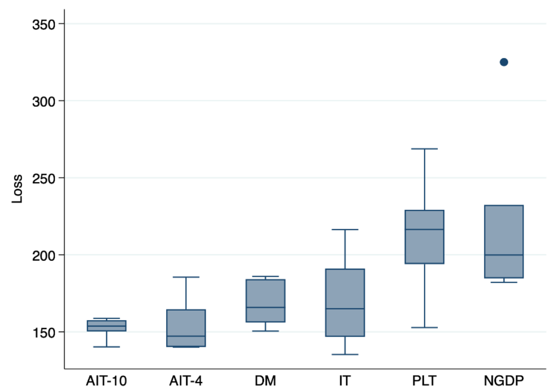
Figure 6: Aggregate dynamics of inflation, output, and interest rate in NGDP level targeting (NGDP) treatment



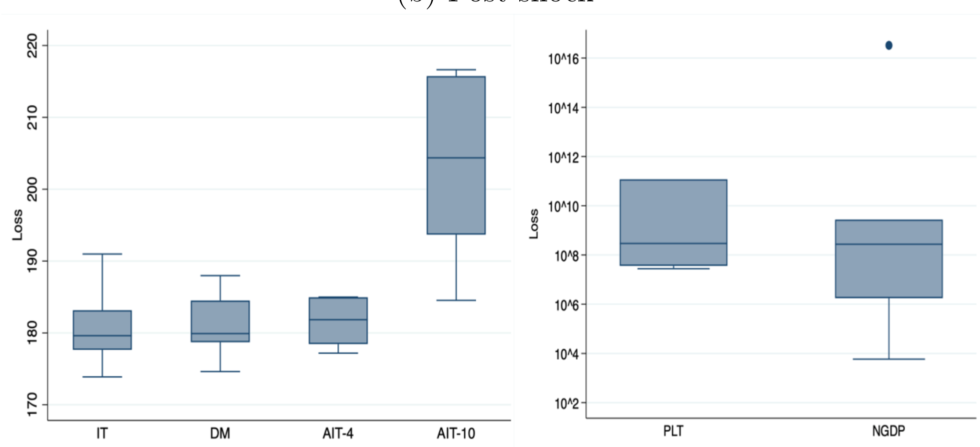
(a) periods 1-20 (b) periods 21-50

Figure 7: Distribution of session losses, by phase

(a) Pre-shock

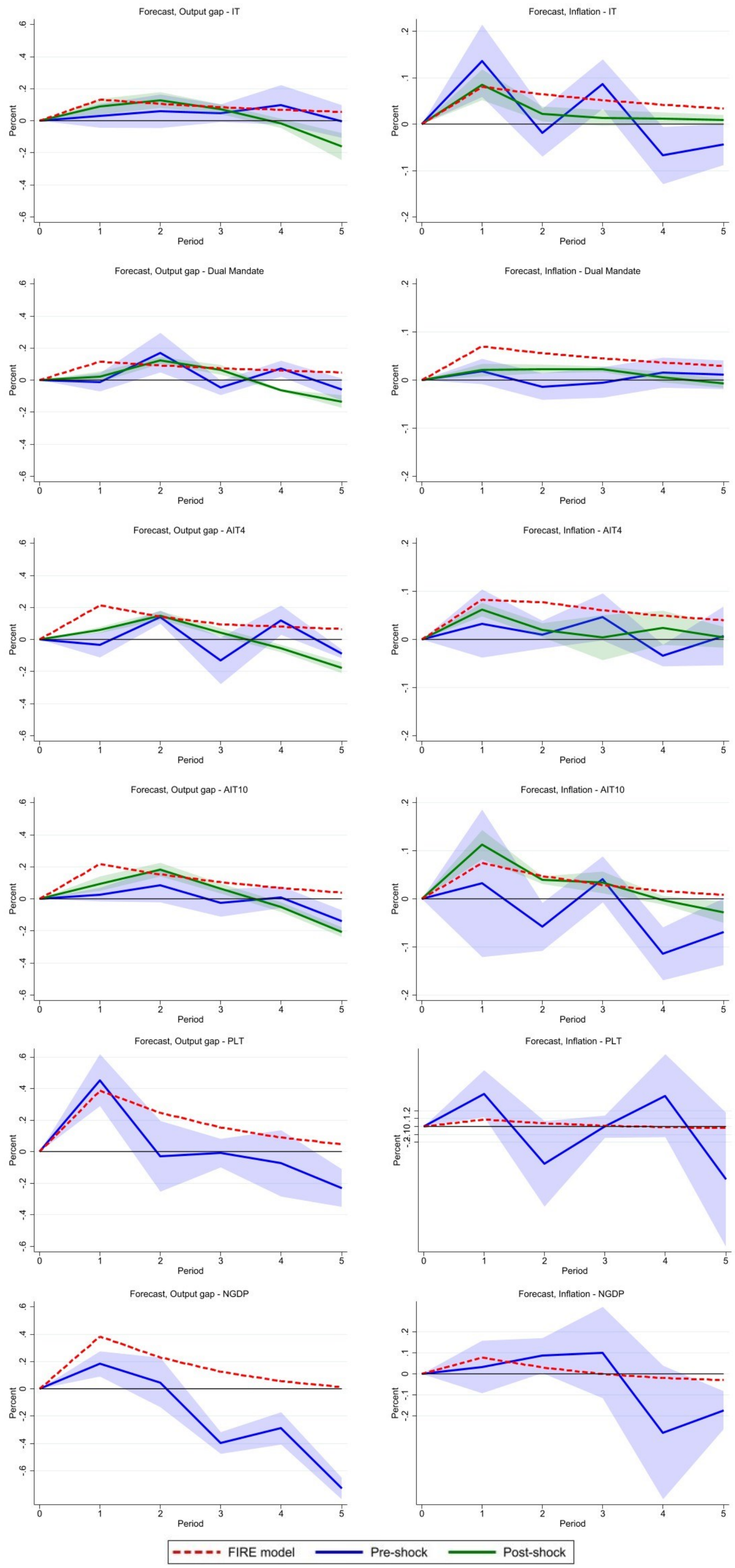


(b) Post-shock



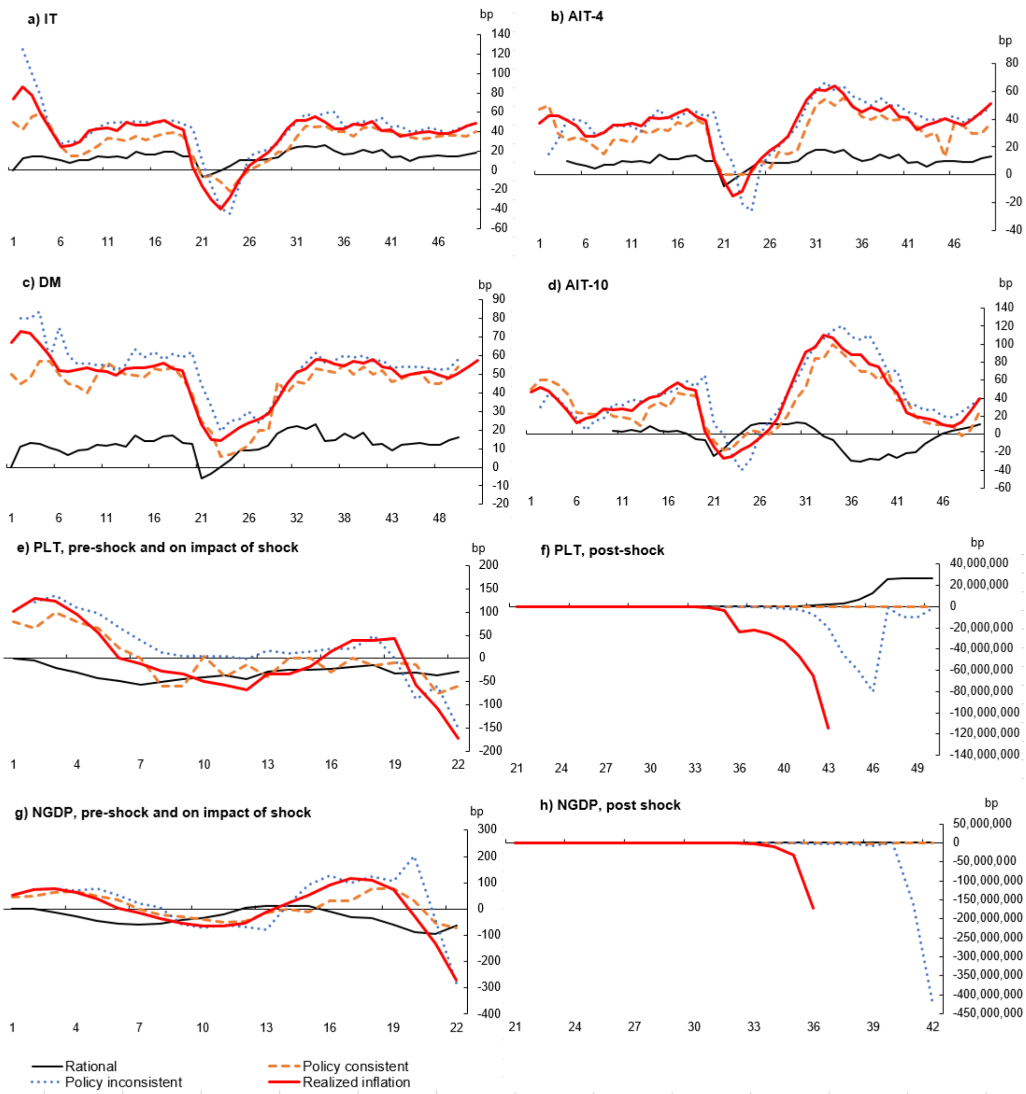
Dots in the figures represent values outside upper adjacent value (upper quartile ± 3/2 times the interquartile range). The y-axis in Panel (b) for PLT and NGDP is a logarithmic scale.

Figure 8: Response of forecasts to a +1% demand impulse.



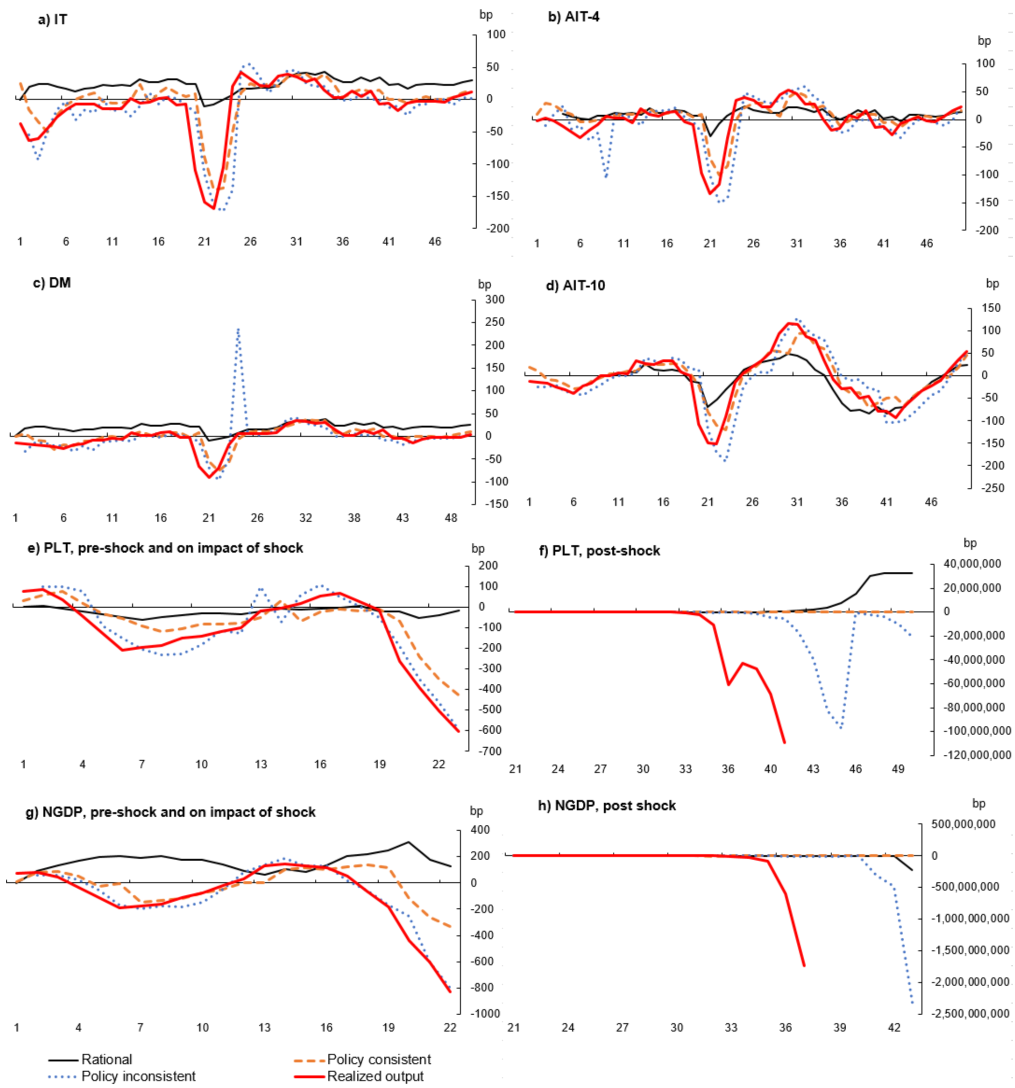
Notes: This figure presents IRFs estimated by local projections using Equation 11. Shaded areas denote the 90% confidence interval. IRFs in the FIRE model are the simulated responses to a one-standard deviation shock to aggregate demand.

Figure 9: Policy consistent and policy inconsistent inflation forecasts



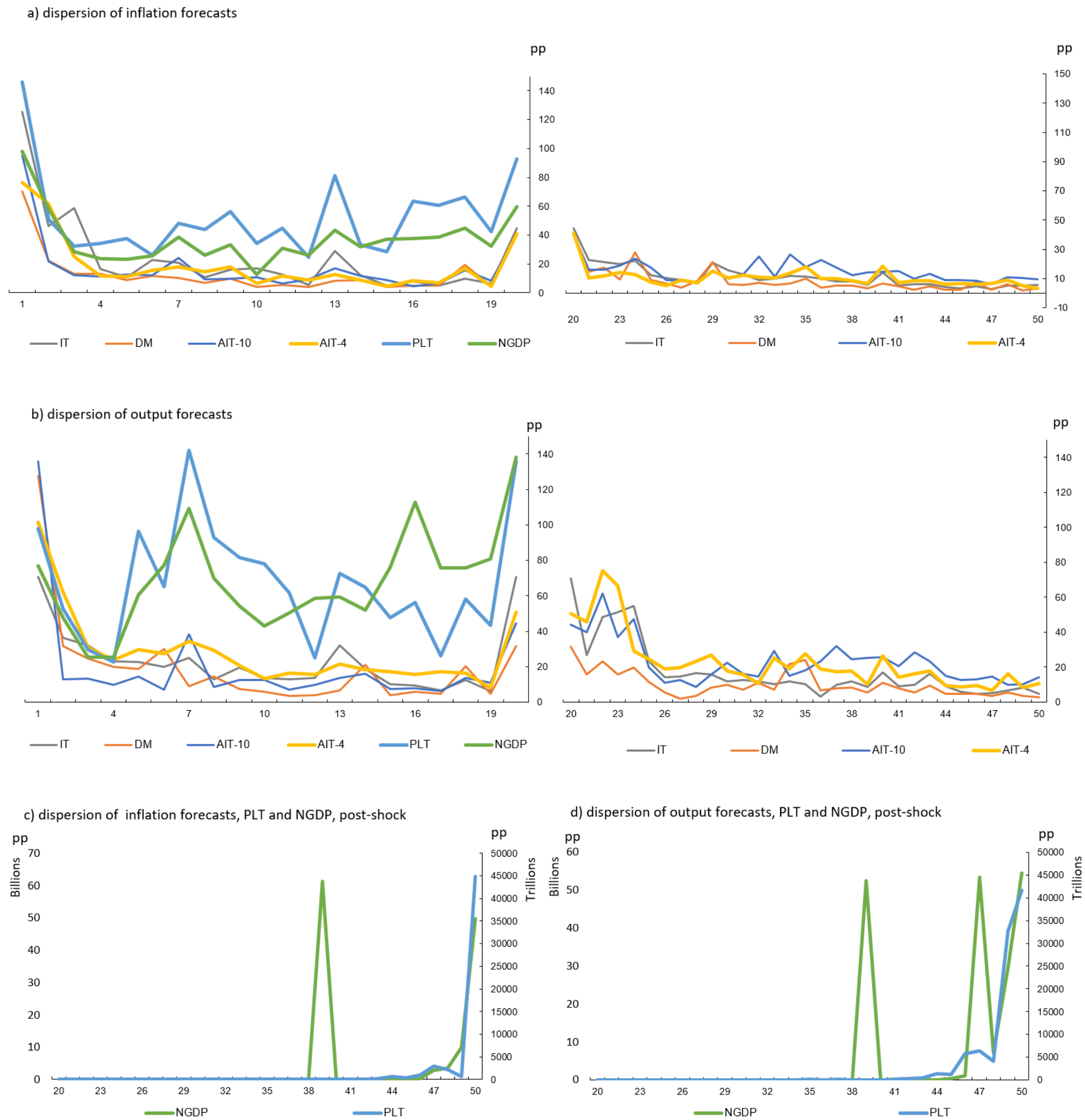
This figure presents the median policy consistent and policy inconsistent forecasts, averaged across all sessions of each treatment.

Figure 10: Policy consistent and policy inconsistent output forecasts



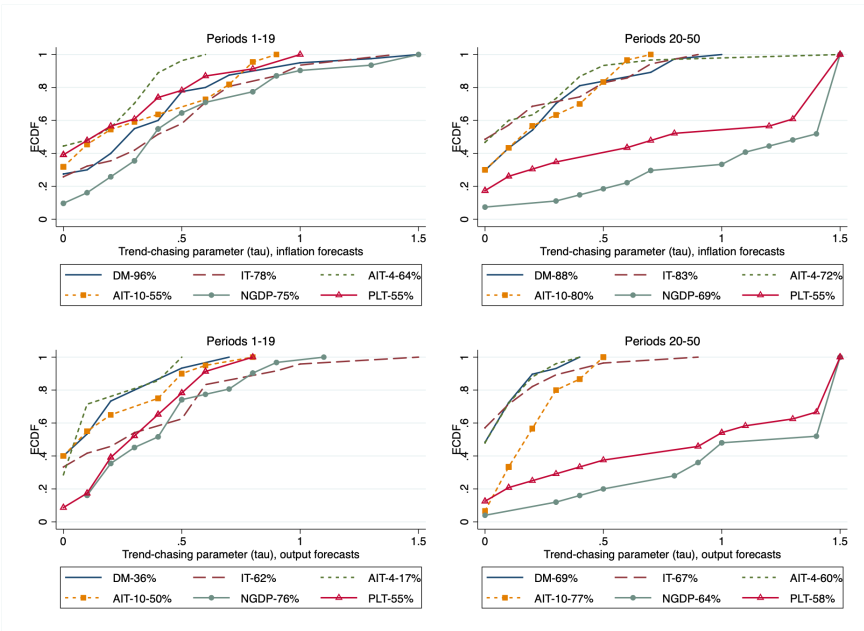
This figure presents the median policy consistent and policy inconsistent forecasts, averaged across all sessions of each treatment.

Figure 11: 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.

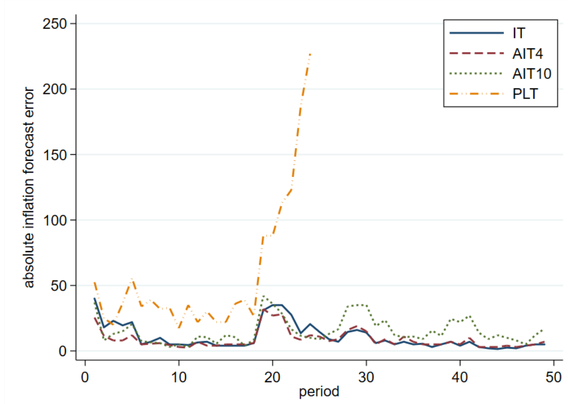
Figure 12: Distribution of trend-chasing parameter τ in inflation and output forecasts by phase



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 13: Median absolute forecast errors, by treatment

(a) Inflation



(b) Output gap

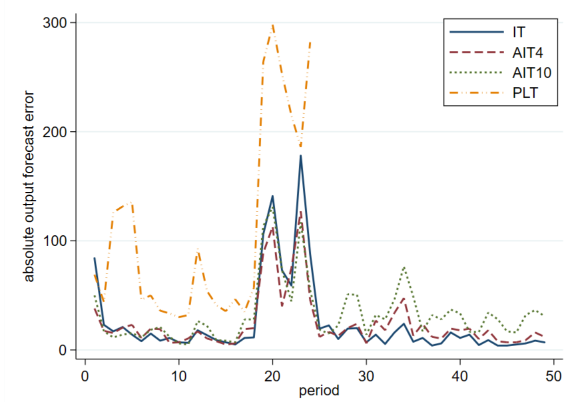


Figure 14: Aggregate dynamics of inflation, output, and interest rate in PLT Comm treatment

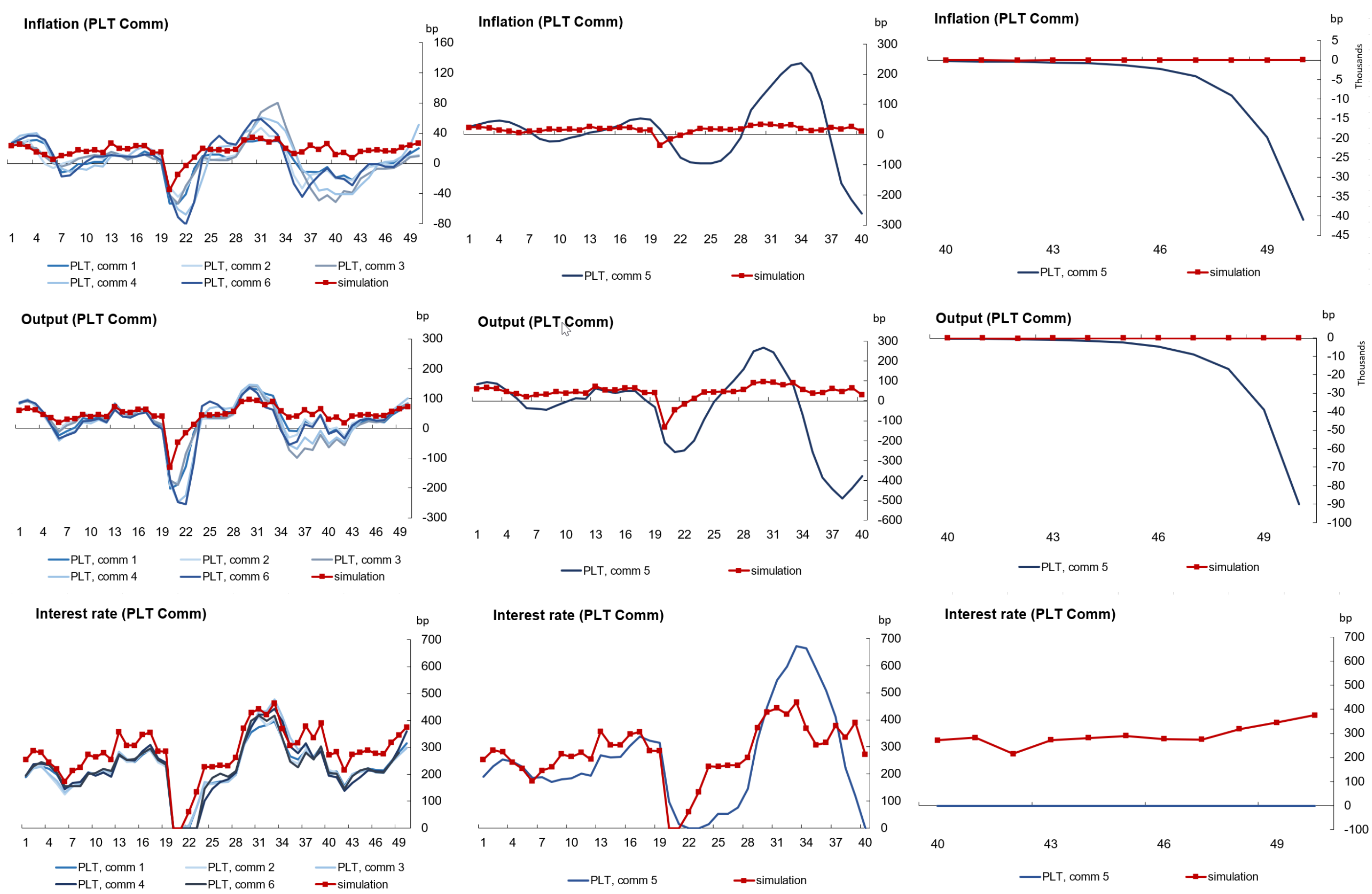


Table 1: Losses from simulations with rational and naive expectations

rational	(1)	naive	(2)
periods 1-50			
NGDP	168.2	AIT-10	194.7
PLT	169.8	IT	198.4
AIT-10	179.7	DM	202.2
AIT-4	180.8	AIT-4	226.0
DM	184.4	PLT	235.3
IT	186.9	NGDP	3010.6
periods 1-19			
NGDP	153.8	PLT	159.6
PLT	155.9	NGDP	161.4
AIT-10	164.0	AIT-10	180.1
AIT-4	165.3	IT	181.4
DM	168.5	DM	184.5
IT	170.8	AIT-4	206.8
periods 20-50			
NGDP	176.4	AIT-10	203.1
PLT	177.7	IT	208.2
AIT-10	188.7	DM	212.3
AIT-4	189.7	AIT-4	237.0
DM	193.5	NGDP	271.4
IT	196.2	PLT	3821.4

Notes: Column (1) reports losses from the simulations with rational expectations. Column(2) reports losses from the simulations with a combination of rational expectations and naïve expectations. Shares of naïve expectations: 33% in PLT and 45% in all other regimes. Losses are expressed in basis points.

Table 2: Losses in laboratory experiments

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

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

Table 3: Wilcoxon rank order test, statistical significance

	NGDP	PLT	DM	IT	AIT-4
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	0.49	0.26	0.85	0.22	0.12	0.29	0.29	0.18
PLT	0.49	0.64	0.29	0.71	0.58	0.49	0.36	0.38	0.26
DM	0.47	0.49	0.25	0.63	0.56	0.27	0.36	0.57	0.16
IT	0.54	0.54	0.33	0.70	0.49	0.26	0.46	0.54	0.21
AIT-4	0.48	0.62	0.32	0.58	0.52	0.20	0.43	0.56	0.25
AIT-10	0.48	0.57	0.33	0.63	0.80	0.51	0.51	0.65	0.36
PLT Comm	0.71	0.74	0.62	0.92	0.85	0.82	0.66	0.62	0.50

Notes: This table presents the share of forecasts in the direction of the rational expectations equilibrium solution.

Table 5: Evolution of policy-consistent expectations

Dep. Var. Pre-shock: Periods 1-19 $\mathbb{1}^{PolicyConsistent}_{i,t}$	IT	DM	AIT-4	AIT-10	PLT	NGDP	PLT Comm	
$\mathbb{1}^{PolicyConsistent}_{i,t-1}$	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)		
$\lvert\pi_{t-1} - \pi^{Proj}_{t-2,t-1}\rvert$								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 $\mathbb{1}^{PolicyConsistent}_{i,t}$	IT	DM	AIT-4	AIT-10	PLT	NGDP	PLT Comm	
$\mathbb{1}^{PolicyConsistent}_{i,t-1}$	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)		
$\lvert\pi_{t-1} - \pi^{Proj}_{t-2,t-1}\rvert$								-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: $\mathbb{1}^{PolicyConsistent}_{i,t} = \alpha + \beta_1 \mathbb{1}^{PolicyConsistent}_{i,t-1} + \beta_2 AbsDevFromTarget_{t-1} + \beta_4 \mu_i + \epsilon_{i,t}$ where $\mathbb{1}^{PolicyConsistent}_{i,t}$ 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: $\lvert\pi_{t-1} - \pi^*\rvert$ is deviation of last-period inflation from target for IT and DM; $\lvert\frac{1}{4}\sum_{t-4}^{t-1}\pi_j - \pi^*\rvert$ is the deviation of last 4-period average inflation from target for AIT-4; $\lvert\frac{1}{10}\sum_{t-4}^{t-1}\pi_j - \pi^*\rvert$ is the deviation of last 10-period average inflation from target in AIT-10; $\lvert P_{t-1} - P^*\rvert$ is the deviation of price level from its target; $\lvert NGDP_{t-1} - NGDP^*\rvert$ 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

Abs.FE	Preshock (Periods 1-19) Inflation	Preshock (Periods 1-19) Output	Shock (Periods 20-21) Inflation	Shock (Periods 20-21) Output	Postshock (Periods 22-50) Inflation	Postshock (Periods 22-50) 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

ln(Abs.FE)	Preshock (Periods 1-19) Inflation	Preshock (Periods 1-19) Output	Shock (Periods 20-21) Inflation	Shock (Periods 20-21) Output	Postshock (Periods 22-50) Inflation	Postshock (Periods 22-50) 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

ln(Abs.FE)	Preshock (Periods 1-19) Inflation	Preshock (Periods 1-19) Output	Shock (Periods 20-21) Inflation	Shock (Periods 20-21) 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.