TI 2023-028/VI
Tinbergen Institute Discussion Paper

HISTORY-DEPENDENT MONETARY REGIMES: A LAB EXPERIMENT AND A HENK MODEL

Jasmina Arifovic[1]
Isabelle Salle[2]
Hung Truong[1]

1 Simon Fraser University, BC, Canada

2 University of Ottawa; University of Amsterdam and Tinbergen Institute

History-Dependent Monetary Regimes: A Lab Experiment And A HENK Model*

Jasmina Arifovic† Isabelle Salle‡ Hung Truong§

May 11, 2023

Abstract

Price-level targeting (PLT) is optimal under the fully-informed rational expectations (FIRE) benchmark but lacks empirical support. Given the hurdles to the implementation of macroeconomic field experiments, we utilize a laboratory group experiment – where expectations are elicited from human subjects – to collect data on expectations, inflation and output dynamics under a traditional inflation targeting (IT) framework and a PLT regime with both deflationary and cost-push shocks. We then emulate the subjects' expectations with a micro-founded heterogeneous-expectation New Keynesian (HENK) model and reproduce the macroeconomic dynamics observed in the lab. Both in the lab and in the HENK model, the benefits of PLT over an IT regime obtained under the FIRE assumption are not observed: both human subjects and HENK agents are unable to learn the underlying implications of PLT, which results in excess macroeconomic volatility. However, once augmented with an inflation guidance from the CB consistent with closing the price gap, the stabilizing benefits of PLT materialize both in the lab and in the model.

Keywords: *heterogeneous expectations, learning, central bank communication, lab experiments*

JEL Classification: *E7; E52; E42; C92.*

*The study has received ethical approval from the SFU Office of Research Ethics numbered 30000994 and is registered under No. AEARCTR-0011128 on the AEA RCT Registry. The present work has benefited from fruitful discussions with Luba Petersen. We also thank the participants of the 56th Canadian Economics Association Conference in Ottawa on June 2022, the 1st PhD Conference on Expectations in Macroeconomics, held virtually by the Barcelona School of Economics on October 2022, the 8th Atlantic Canada Economics Association in Halifax on October 2022. None of the above are responsible for potential errors in the paper.

†Department of Economics, Simon Fraser University. Jasmina sadly passed away before the completion of the present work. However, given her extended intellectual contribution and leadership, we opted for keeping her as a co-author but the views ultimately expressed in this paper may only be attributed to the two remaining authors.

‡Department of Economics, University of Ottawa; Amsterdam School of Economics, University of Amsterdam and Tinbergen Institute, isalle@uottawa.ca.

§Corresponding author: Department of Economics, Simon Fraser University, hungt@sfu.ca.

1 Introduction

The effective-lower bound (ELB) episode that accompanied the Great Recession followed by the steep and persistent rise in inflation in the aftermath of the COVID-19 pandemic have called forth a rethinking of the central bank (CB) policies. In particular, the risk of unanchored expectations, regardless of direction, have historically proven to be a challenge because it may impair the expectation transmission channel of monetary policy. In this context, a remedy could be a shift from standard inflation-targeting (IT) regimes to a form of history-dependent policy, of which price-level targeting (PLT) is a notable representative. Under PLT, nominal rates are adjusted so that the price level gravitates around the path that is consistent with a stable and low inflation target while standard IT is only concerned with adjusting nominal rates to keep *current* inflation around this target.

In this context, the Bank of Canada has considered PLT in its monetary policy framework reviews but fell short of implementing it [Wilkins 2018]. More recently, the Fed (and to a lesser extent the ECB) have announced an adjustment to their monetary policy framework to now emphasize history-dependent inflation targets.[1] We review below the academic literature on the matter, which is limited and discrepant.

This paper brings a twofold contribution to this literature. First, we provide empirical evidence on the stabilizing properties of a history-dependent monetary regime using a laboratory experiment with human subjects that studies the interplay between subjects' forecast dynamics and macroeconomic developments. We then introduce a micro-founded heterogeneous expectation New Keynesian (HENK) model (see Grimaud et al. [2023]) to rationalize the experimental results and shed light on expectation dynamics in the different regimes.

[1] In August 2020, the Fed announced a change to its monetary policy framework and now refers to an average inflation targeting (AIT) regime without specifying explicitly the number of past lags relevant for the setting of the interest rate; see Powell et al. [2020]. Since June 2021, the ECB has been communicating on an inflation objective over a medium term.

Field experiments in macroeconomic settings are impractical and gauging policy-dependent expectations in survey data is difficult. The laboratory can fill this data gap and provides a valuable environment to gain insights into the expectation channel under alternative monetary policy frameworks. Focusing on the expectation channel is particularly relevant in the context of the assessment of history-dependent rules. There is no historical experience with PLT[2] and the theoretical merits of this regime over IT have been established under the standard FIRE (fully-information rational expectations) assumption, which presumes that the private sector understands the underlying implications of history-dependence and aligns their expectations with the PLT monetary policy rule. However, the FIRE assumption falls short in the empirical literature [Coibion et al. 2018, D'Acunto et al. 2022] and models that relax this assumption generally fail to reproduce the benefits of history-dependent rules [Honkapohja & Mitra 2020]. In this respect, CB communication – which is fully redundant under the FIRE assumption – may help nudge the public towards forming policy-consistent expectations and help the properties of PLT under RE materialize with non-FIRE agents. Communication may find a particular rationale once one acknowledges that expectations are also heterogeneous,[3] and the CB therefore also needs to worry about cross-sectional dispersion and coordinate individual beliefs on the target. This is precisely what we test in the experiment.

We utilize a forecasting experiment where inflation and output gap expectations are elicited from a group of human subjects and represent the only degree of freedom in an otherwise standard macroeconomic model. In contrast to the usual implementation of this class of experiment (reviewed in Section 2 below), our experimental design allows for peer imitation, which we provide evidence of in our data. We then introduce a HENK model with social interactions to account for the subjects' behavior. This social behavior in expectation formation echoes the recent

[2] One exception, albeit hardly informative for modern times, is the short Swedish episode between 1931 and 1933.

[3] Earlier survey evidence of heterogeneity in macroeconomic expectations can be found in Branch [2004]; see D'Acunto et al. [2022] and the references herein for the most recent account

empirical findings among forecasters [Coibion et al. 2021, Carroll & Wang 2023]. Our experimental setting is also the first to contrast the same policy framework under both deflationary and inflationary environments, in an effort to reproduce the recent course of inflation. We also elicit long-term forecasts to assess expectation anchorage beyond the short-run. Furthermore, our behavioral model is designed so that agents can learn to utilize an external signal, i.e. the CB guidance. Agents can then learn to either remain backward-looking or shift to a forward-looking behavior within the same model, which is uncommon in other learning models. Our framework is particularly parsimonious and yet fits remarkably well the experimental data.

Our main findings are as follows. First, both in the lab and in the HENK model, the benefits of PLT over an IT regime obtained under the FIRE assumption fails to materialize. On the contrary, inflation and output under PLT exhibit excess volatility compared to a standard IT regime, both in an inflationary and in a deflationary environment. These wild oscillations result from the commitment to correcting past price gaps in a learning environment, where expectations become unanchored and destabilize further output and inflation. However, the picture is reversed if the CB communicates in each period the time-varying transitory inflation target that is consistent with closing the price gap. In this case, expectations are better coordinated around the CB target under PLT than under IT, in the lab as well as in the HENK model. As a result, both deflationary and inflationary shocks result in smaller inflation gaps and shorter ELB episodes under PLT than in the two alternatives. In particular, the commitment to correct the deviations of the price level brings inflation toward the transitory target which, in turn, confirms the relevance of the CB's guidance to forecast inflation and initiates a credibility-stability loop among the subjects and the HENK agents. We also find that long-run expectations remain closer to the target under PLG with guidance despite the short-run fluctuations in inflation generated by the exogenous shocks than under IT and PLT without additional communication. Both the experiment and the HENK model also produce persistent inflation gaps under IT consistent with the recent inflation course. Furthermore, our study contributes to explaining the discrepancy in the literature about the properties of PLT because we find empirical arguments that describe both a PLT regime implying less or more volatility in inflation and output than a traditional IT regime.

The rest of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 describes the experimental design and data. Section 4 introduces the behavioral HENK model and sheds light on the experimental data within this model. Section 5 concludes.

2 Related literature

The theoretical literature on PLT – and history-dependent rules in general – is somewhat limited and rather discrepant. Among earlier contributions, Woodford [2000] argued that optimal monetary policy functions must involve some degree of history-dependence, which is absent from standard forward-looking IT rules; see also Svensson [1999] and Mitra [2003]. In general, under RE, PLT is known to be superior to IT from the point of view of inflation and output gap stabilization. For instance, Cateau et al. [2009] use the ToTEM model[4] to show that switching from IT to PLT can deliver welfare gain as long as the CB is committed to sticking to the new framework. By contrast, once the RE assumption is relaxed and agents form backward-looking expectations (e.g. under econometric learning), whether PLT remains superior to IT depends crucially on whether the agents form expectations that are model consistent. If their forecasting models are misspecified, convergence back to the target after large pessimistic shocks becomes *less* likely under PLT than under IT [Honkapohja & Mitra 2020]. Misspecification could arise from two sources that are not accounted for under RE: a lack of understanding or a lack of credibility of the new policy rule. Bodenstein et al. [2022] find that agents can-

[4]ToTEM is a standard large-scale open-economy DSGE model of the Canadian economy developed and used at the Bank of Canada.

not learn the new interest-rate rule if the transition from IT to PLT coincides with a demand-driven recession where the ELB binds because the two rules become indistinguishable. In short, model-consistent expectations are a key condition for PLT to outperform a standard IT regime but whether expectations would fulfill this property under PLT remains an empirical question.

In this context, forecasting group experiments where expectations are elicited from human subjects constitute a convenient environment to analyze the expectation channel of various monetary policies and CB communication; see Hommes [2021] for a survey of this class of experiments. Within the particular question of the relative merits of history-dependent monetary policy rules, the seminal paper of Arifovic & Petersen [2017] has the flavor of PLT. They set up such an experiment where the CB communicates a time-varying inflation target in each period. They find that such targets lose credibility – i.e., expectations of the subjects become disconnected from the value of this target – as soon as exogeneous shocks push inflation sufficiently down. In other words, they show that agents first 'need to see it to believe it.' Relatedly, Cornand & M'baye [2018] find that communicating the value of the inflation target only brings minor additional stabilization benefits within the context of a standard IT rule. While we also report a credibility loss for such communication at the onset of a recession, our PLT-CB does eventually regain enough credibility to steer expectations. In this sense, our results bear less dire implications than the ones of Arifovic & Petersen [2017]. In Rholes & Petersen [2021], the CB communicates a precise point projection of inflation and the authors find that such projection is more effective at coordinating subjects' forecasts around the target than projecting an inflation density. Within a non-linear NK model, Hommes & Makarewicz [2021] find that a PLT-CB with a sufficiently strong price gap reaction coefficient can induce convergence back to the target once the ELB binds more often than under IT. Salle [2021] also finds experimental evidence of PLT outperforming IT in stabilizing inflation but much of the stability observed under PLT may come from the frictionless model used. In an experiment horse race between different monetary policy regimes, Kostyshyna et al. [2022] find that PLT is outperformed by IT in stabilizing business cycles. In fact, their experimental economy under PLT is drastically dragged into a deflationary spiral after an ELB shock, whereas a similar phenomenon is not observed under other regime treatments. In an additional treatment, these authors also find that the publication of CB projections may help stabilize the economy under PLT.

It is crucial to note that all these contributions analyze the consequences of alternative monetary frameworks in the context of deflationary pressures. By contrast, we implement inflationary shocks in our learning economies and provide the first empirical evidence of the benefits of a history-dependent regime with communication in such a context. Beyond the lab, we further develop a HENK model that matches our experimental data.

More broadly, our contribution relates to the studies of forward guidance because the CB guidance that we implement under PLT may be interpreted as a rule-based forward guidance about the future course of inflation. In this respect, Del Negro et al. [2023] point out that forward guidance has been less effective in reality than the standard DSGEs model would predict (the so-called "forward guidance puzzle"). In our model, forward guidance is plagued by (endogenous) credibility issues and information frictions caused by agents' heterogeneity in expectations but the puzzle is ameliorated.

Finally, we contribute to the growing literature on heterogeneous-expectation models.[5] Earlier contributions incorporate so-called heuristic-switching models into NK frameworks to model endogenous switches between different predictors (see, e.g., Branch & McGough [2010], Massaro [2013], Gasteiger [2014]. Busetti et al. [2017] and Ozden [2021] extend this framework to consider the ELB constraint. Arifovic et al. [2013] and Arifovic et al. [2018] explore the asymptotic stability of the targeted steady state and ZLB steady state in an NK model with a learning mechanism that has the same flavor as the one we use in this paper. Jia & Wu

[5] Our work also builds on the literature that uses evolutionary algorithms to model heterogeneity; see Arifovic [2000] for an early survey and, *inter alia*, Anufriev & Hommes [2012], Anufriev et al. [2019] for more recent applications.

[2021] use social learning to model beliefs under an intentionally ambiguous AIT regime where the number of lags in the policy rule is unknown; see also Hachem & Wu [2017]. Finally, our treatment of endogenous credibility and communication is in line with a series of papers that assess convergence to the target under various learning mechanisms and a traditional IT regime, including adaptive learning [Orphanides & Williams 2004], heuristics-switching model [Hommes & Lustenhouwer 2019] or social learning [Arifovic et al. 2023].

3 Experimental data

We first present the experimental design and then establish the main stylized facts from the resulting experimental data.

3.1 The experimental design

3.1.1 The experimental environment

The experiment is a group experiment with a between-subject design. Each experimental session involves a fixed group of six subjects tasked with making point forecasts of output and inflation gaps in the next period of the experimental economy. The economy unfolds for 75 periods, split into two separate subsessions, one of 40 and one of 35 successive periods. In both subsessions, inflation and output gaps evolve according to a baseline three-equation NK model (see, e.g., Woodford [2003]). The only difference from the standard textbook implementation is that expectations are provided by the subjects and need not align with RE or any *ad-hoc* learning mechanism commonly used in the theoretical literature.

Precisely, in each period, we aggregate[6] the inflation and output gap expectations

[6]The heterogeneous expectation process and the aggregate procedure are fully micro-founded in Grimaud et al. [2023], where a general solution method for this class of HENK models is also

of the six subjects $j = 1, ..., 6$ denoted by, respectively, $\pi^e_{j,t+1}$ and $y^e_{j,t+1}$ using the arithmetic mean, that we denote by $\tilde{E}_t(\pi_{t+1})$ and $\tilde{E}_t(y_{t+1})$ and insert them into the standard IS and Phillips curves given by:

$$\hat{y}_t = \tilde{E}_t \hat{y}_{t+1} - \sigma^{-1} \left(\hat{\imath}_t - \tilde{E}_t \hat{\pi}_{t+1} - \bar{r} \right) + \hat{g}_t, \tag{1}$$

$$\hat{\pi}_t = \beta \tilde{E}_t \hat{\pi}_{t+1} + \kappa \hat{y}_t + \hat{u}_t, \tag{2}$$

where $\hat{y}_t$ is the output gap, $\hat{\pi}_t$ the inflation gap, that is the difference between inflation and the CB target π^T, $\hat{g}_t$ and $\hat{u}_t$ exogenous real and cost-push shocks to be specified below, β and σ households' preference parameters, κ the slope of the Phillips curve and $\hat{\imath}_t$ the nominal interest rate set by the CB either with an IT or a PLT rule (depending on the experimental treatments, see hereafter in Section 3.1.2). The rate $\hat{\imath}_t$ is expressed in deviation from its steady-state level $\bar{r} \equiv \pi^T + \rho$. In all experimental sessions, we use the following parameter values $\beta = 0.994$, $\sigma = 1$, $\kappa = 0.0625$, $\pi^T = 5\%$.

The two subsessions only differ by the nature of the shocks. The first subsession undergoes real shocks and the second cost-push shocks. In all subsessions, we impose an initial learning phase for the subjects consisting of small white noise shocks. A severe and persistent deflationary real shock then disturbs the economy in period 10 of Subsession 1 and a cost-push shock hits the economy in period 8 of Subsession 2. The shock sequences are the same for all sessions and are displayed in Appendix C. In other words, we engineer two recessions in each experiment session. The recession of Subsession 1 is coupled with low inflation and does not imply a trade-off for monetary policy in terms of stabilization of inflation versus output but the severity of the initial shock induces an ELB episode in the model under RE independently of the monetary policy rule; see Section 3.1.2 and Fig. 1 below. By contrast, the recession in Subsession 2 is associated with inflation pressures which may imply such a trade-off. Our experiment thus allows us to collect

introduced. The HENK model that we utilize in Section 4 to account for the experimental data is based on this contribution.

empirical data on the properties of IT and PLT in distinct economic environments. Additionally, to measure long-run inflation expectation anchoring in each of these environments, at period 26 of each subsession, subjects are tasked to submit a one-time inflation forecast for the end of the subsession that is, respectively, 15 and 10 periods ahead.

The role of the subjects in the experiment is framed as professional forecasters for a statistical bureau. The instructions explain the positive and negative signs in the relationships between the main variables in the economy and provide the value of the inflation target. All information across treatments is identical except for the one pertaining to the monetary policy rule and the additional provision of the inflation guidance under the treatment that involves PLT and communication (we specify the guidance in Section 3.1.2 below). At the beginning of every period, subjects observe the current shock, all previously realized endogenous variables, and the value of the inflation target. All these pieces of information are displayed in a table and in charts on the graphical user interface (GUI, see App. E.2 for examples).

Importantly, and in contrast to the existing literature, in our experiment, subjects observe the most recent forecasts of all other forecasters in their group along with their payoff. In particular, an explicit ranking from the best to the worst forecasters in which each given subject is included (tagged by 'YOU') is updated in each period on a table on the GUI.[7] In other words, subjects observe their peers' forecasting performances and can assess their relative performances. This unique feature of our experiment allows for the possibility of peer imitation and the information set in the experiment matches the one of the HENK agents in Section 4.2.

Finally, subjects accumulate points that reward their forecast accuracy throughout the session. In each period, a forecast brings the $\frac{100}{1+\text{absolute forecast errors in p.p.}}$ points, so that a perfect forecast brings 100 points. To rule out strategical hedging across the

[7] In line with standard ethical rules, the identity of each individual subject remained, of course, anonymous and subjects where labeled by their number in the experiment.

inflation and output forecasts, subjects accumulate a separate amount of forecasting points for each variable and only one of the two is randomly drawn at the end of the experiment to be converted into money at a rate of CAD $0.2 per 100 points. To reinforce the salience of the occasional long-run inflation forecasting task, the points corresponding to these two forecasts were converted at a rate of CAD $3 per 100 points.

3.1.2 Experimental treatments and theoretical predictions

We implement three experimental treatments, denoted by `IT`, `PLT`, and `PLT-G`. Tr. `IT` differs from Trs. `PLT` and `PLT-G` only by the monetary rule that is used to determine the nominal interest rate in Eq. (1). Under `IT`, the CB implements a standard contemporaneous Taylor rule subject to the ELB given by:

$$\textbf{IT}\ \hat{\imath}_t = \max\left\{0, \bar{r} + \phi^{\pi}\hat{\pi}_t + \phi^{y}_{IT}\hat{y}_t\right\}, \tag{3}$$

where the ELB is set to 0% and we use $\pi^T = 5\%$ and $\phi^{\pi}_{IT} = 1.5, \phi^y = 1$, which ensures determinacy under RE and stability under backward-looking expectations, see App. A.

In Trs. `PLT` and `PLT-G`, the CB uses a PLT monetary policy rule. To do so, let us define the price gap $\hat{P}_t \equiv \frac{P_t - \bar{P}_t}{\bar{P}_t}$, which measures the gap between the price level P from the predetermined targeted price path $\bar{P}$ that grows at a rate π^T consistent with the CB inflation target. In the log-linearized model, the price gap evolves as follows:

$$\hat{p}_t = \hat{p}_{t-1} + \hat{\pi}_t. \tag{4}$$

In Trs. `PLT` and `PLT-G`, the CB sets the interest rate using the following monetary policy rule:

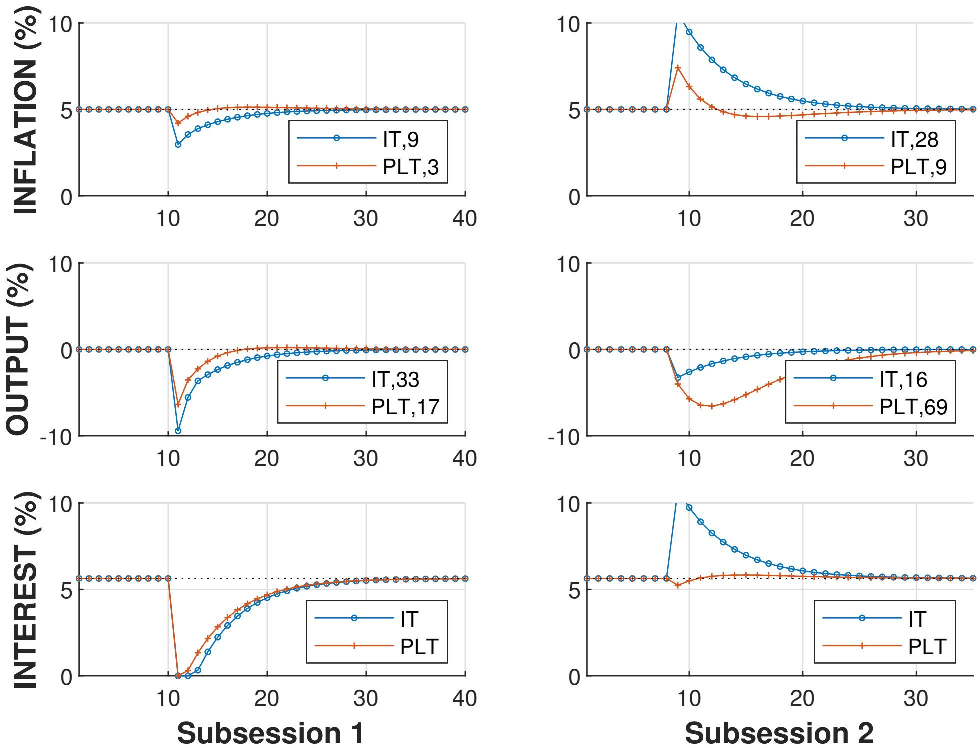
$$\textbf{PLT}\ \hat{\imath}_t = \max\left\{0, \bar{r} + \phi^{P}\hat{p}_t + \phi^{y}\hat{y}_t\right\}, \tag{5}$$

where we use $\phi^P = 1.5$, which rules out any variation in the initial policy response to a shock across treatments (holding all else equal). and ensures determinacy under RE and stability under learning of the targeted steady state; see App. A.

Figure 1 provides the benchmark of the two experimental economies under RE given our calibration and Appendix B.1 and B.2 provide the detail of the derivations of the two model economies under RE. It is clear that PLT is superior to IT for a CB primarily concerned with inflation stabilization. PLT better stabilizes inflation and output gaps than IT in the absence of policy trade-off (namely in face of deflationary pressures, see left panel), even though the ELB binds under both regimes (bottom left panel). When facing inflationary pressures (right panel), PLT better limits the rise in inflation but at the price of a deeper recession than under IT (middle right panel). Comparing inflation and output gap dynamics in Trs. `IT` and `PLT` allow us to assess whether this ranking is robust to laboratory expectations.

Finally, Tr. `PLT-G` differs from Tr. `PLT` only by the provision to the subjects of an inflation guidance. In Tr. `PLT-G`, in each period, subjects receive an inflation guidance that corresponds to the implicit time-varying inflation level that would close the price gap in any period t (see App. A for the detail of the model). We denote such guidance by π_t^{CB}. This guidance is model-consistent and carries the underlying structure of the PLT regime while keeping the frame of the CB communication in terms of inflation rather than price level, which we conjecture to be an easier-to-grasp and more practical communication regime for the public. Under RE, CB communication is redundant because agents already form model-consistent expectations. Hence, under RE, there is no difference between Trs `PLT` and `PLT-G`. Under real-world expectations however, the guidance may help subjects grasp the functioning of PLT and form expectations in line with RE so that the merits of PLT over IT materialize. In each period, subjects receive the guidance in a bold and highlighted manner on the GUI. Furthermore, all previous guidance announcements are displayed on the inflation graph; see, again, App. E.2.

Figure 1: Simulations under RE



Notes: the values in the boxes report the total absolute deviation (in p.p.) for IT (blue circles) and PLT (rec crosses). We use the `occbin` library in `Dynare` to account for the occasionally binding constraint; see [Guerrieri & Iacoviello 2015]. The shocks in the simulations are an AR(1) process with an autoregressive coefficient of 0.8 which fits the series of three negative shocks in the experiment.

3.1.3 Experimental implementation

The experiment sessions were conducted in person at the CRABE laboratory at Simon Fraser University between August and October 2022. The experimental software was programmed using `oTree` [Chen et al. 2016]. We ran six independent sessions composed of groups of six subjects for each treatment, for a total of 108 subjects. The duration of each session was around two-and-a-half hours. Participants earned an average of 24.8 CA$ (from a minimum of 15 to a maximum of 38 with a standard deviation of 6.75), including a $CA 7 participation fee.

Upon entering the lab, subjects were provided with an instruction booklet.[8] Subjects were then asked to carefully read the instructions and complete a pre-experiment quiz in order for the experiment to start. The forecasting game started with Subsession 1. All predictions were submitted on the GUI. At the end of the first subsession and before starting Subsession 2, subjects were provided with the updated instruction booklet that highlights the change in the nature of the shock, from real to cost-push shocks. We introduced information about the shocks in such a sequential manner to prime subjects towards the particular economic environment implemented while avoiding the cognitive overload that could result from providing simultaneous information about various types of shocks. Finally, at the end of the second subsession and before leaving the lab, the subjects were asked to complete an exit survey consisting of demographic questions, central banking literacy, and the strategy(ies) used in the experiment, if any.

We now describe our experimental results and highlight the main stylized facts from our data.

3.2 Experimental results

We first describe the macroeconomic dynamics and then dig into the individual forecast times series to study cross-sectional dispersion and show evidence of peer imitation.

3.2.1 Macroeconomic dynamics in the experimental economies

We first discuss cross-treatment comparisons after a visual inspection of the data in Fig. 2-4 and then confirm these qualitative descriptions with statistical tests in

[8] Standard laboratory procedures apply, e.g. identification, consent acknowledgment, etc. We provide the instructions booklet in Appendix E.

Table 3 hereafter.

Figs. 2-4 report the dynamics of inflation, output gap and interest rate in all experimental economies (dashed blue lines) in, respectively, Trs. `IT`, `PLT`, and `PLT-G`, along with the median realization over the six sessions in each treatment (red thick lines). In Tr. `PLT-G`, we are also interested in whether subjects use the CB guidance to form their inflation forecasts. Since we may not directly observe this behavior, we compute a "credibility index" in each period (see, e.g., Cecchetti et al. [2002]). The index is aggregated over all subjects in each period in a given session s and corresponds to the distance between a subject's inflation forecast and the CB inflation guidance in any given period computed as:

$$Cred_{s,t}^{CB} = \frac{1}{6}\sum_{j=1}^{6} exp(-(\pi_{j,t+1}^{e} - \pi_{t}^{CB})^2), \tag{6}$$

where a higher value indicates a higher CB credibility. The bottom panel of Fig. 4 displays the dynamics of this credibility index in Tr. `PLT-G`.

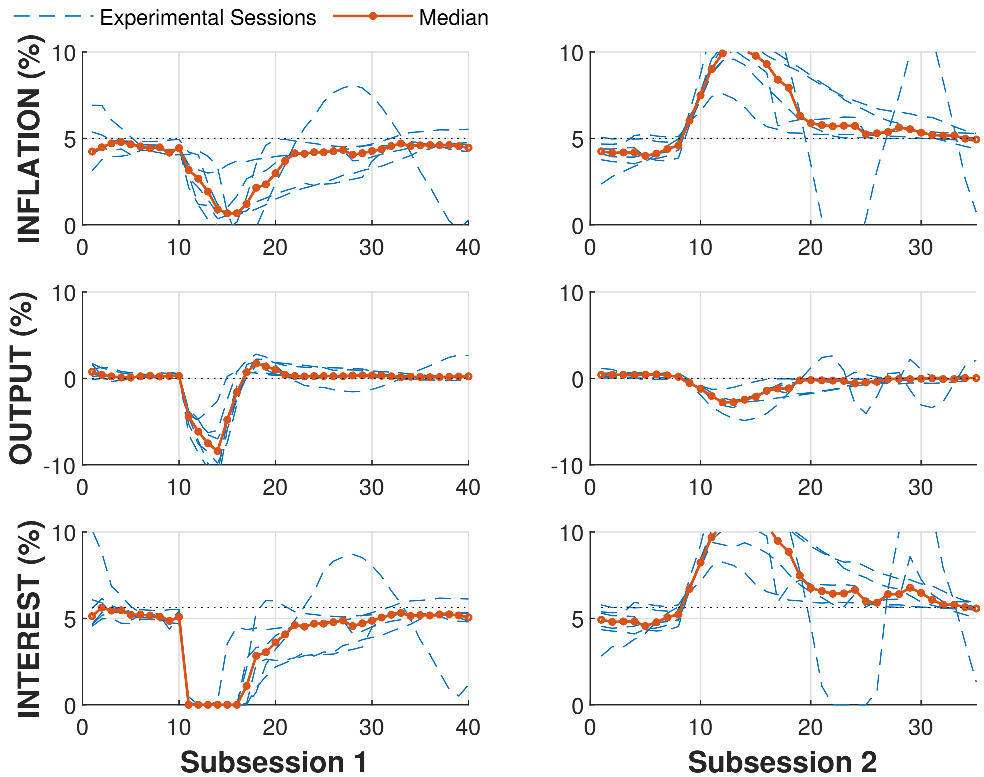
Table 1 further reports cross-treatment comparisons regarding inflation and output gap stabilization. To do so, we compute the per-period *a*verage of the *a*bsolute *d*eviation (AAD) of inflation and output gap from their targets. For a given session $s = 1, .., 6$ and subsession $ss = 1, 2$ and a given time frame of $T = 40, 35$ periods, the AAD for each variable reads as:

$$AAD_{s,ss}^{\pi} = \frac{1}{T}\sum_{t \in T} |\pi_{s,ss,t} - \pi^{T}|, \tag{7}$$

$$AAD_{s,ss}^{y} = \frac{1}{T}\sum_{t \in T} |y_{s,ss,t}|, \tag{8}$$

where lower values indicate a better stabilization of inflation and output around their target. In Table 1, we disentangle pre-shock (top panel) from post-shock periods, which include the shock periods themselves (bottom panel).

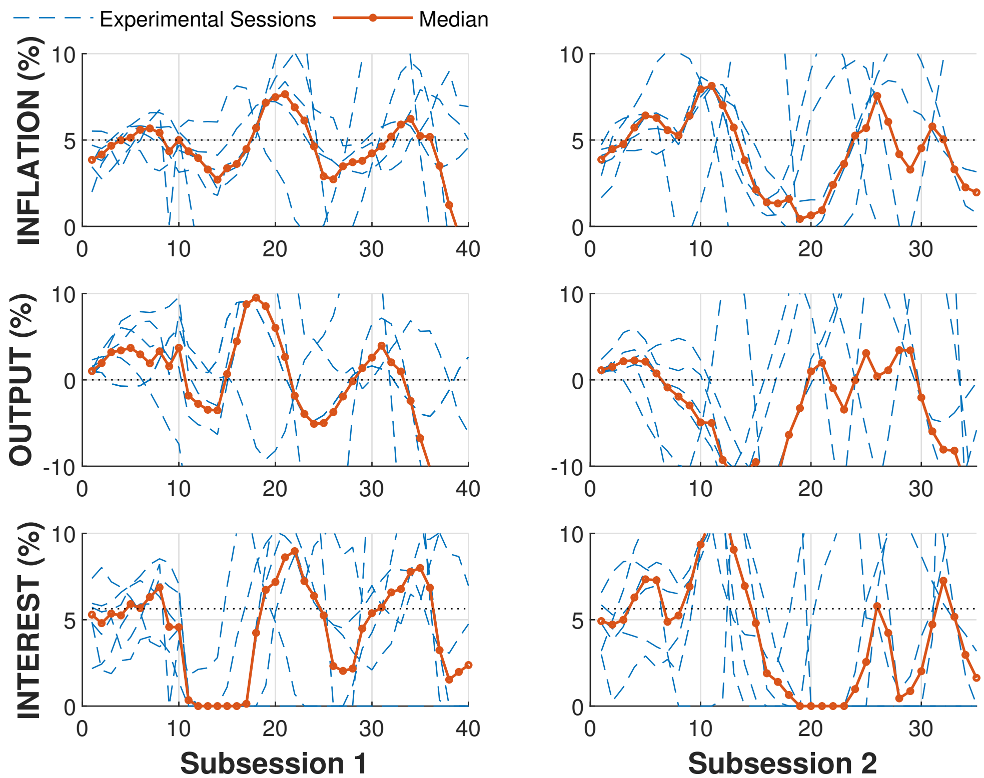
Figure 2: Inflation, output and interest rate dynamics in Tr. `IT`



Overall, the merits of PLT under RE do not seem to extend to the lab environment: experimental economies under PLT are more unstable than economies under IT. However, once augmented with CB guidance in Tr. `PLT-G`, the stabilization benefits of a PLT regime obtained under RE materialize in the lab and this treatment appears to deliver the smallest inflation and output gaps, in particular regarding the inflation gap.

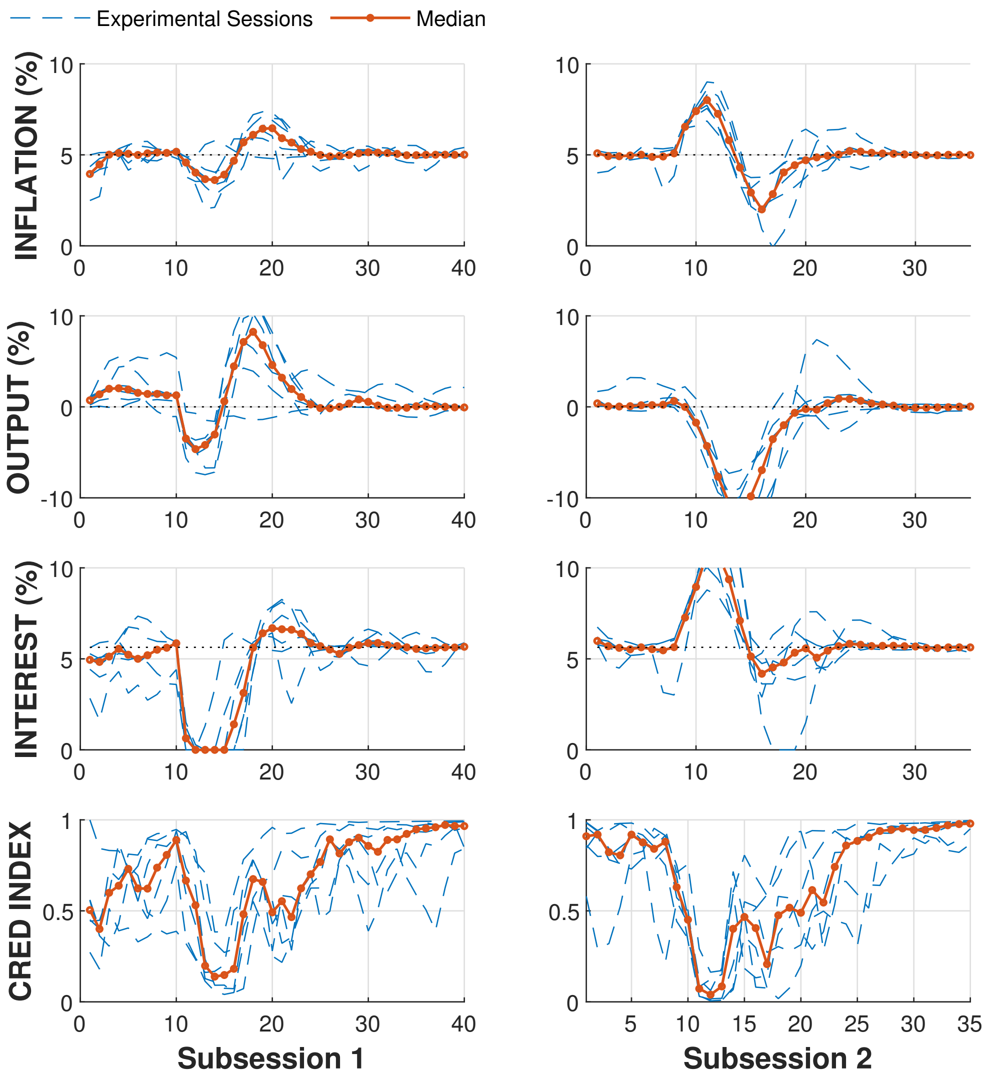
In detail, in the first periods of each subsession – i.e. before the shocks, inflation, and output appear similarly stabilized around the CB's target under IT and PLT with guidance while displaying some wilder oscillations under PLT. This visual impression is confirmed by the cross-treatment comparisons of the AADs

Figure 3: Inflation, output and interest rate dynamics in Tr. `PLT`



(top panel of Table 1): Overall, the ADDs of inflation and output are significantly higher in Tr. `PLT` than in Trs. `IT` and `PLT-G`: the PLT environment without guidance seems to hinder subjects' learning about the experimental environment. In the aftermath of the large shocks, in both subsessions, the cross-treatment differences become more salient: while economies in Trs. `IT` and `PLT-G` eventually converge back to the targeted steady-state, economies in Tr. `PLT` exhibit wilder and non-dampening oscillations. The bottom panel of Table 1 reports that the AADs of both inflation and output are considerably and significantly higher under PLT than under IT or PLT with guidance. The standard deviations of the AADs under PLT are also strikingly larger than in the two other treatments, which indicates more variation across independent sessions.

Figure 4: Inflation, output and interest rate dynamics in Tr. `PLT-G`



Even if inflation returns towards the target in both Trs. `IT` and `PLT-G`, the inflation gap under IT exhibits more persistence in the aftermath of the recessions. Inflation remains persistently below target in the deflationary-shock subsession (left panel) and persistently above target in the cost-push-shock subsession (right panel). These observations are consistent with the persistently low inflation experience post-Great Recession and the persistently high inflation in the post-COVID-19 pandemic. These circumstances may result in an unanchoring of long-run inflation expectation, which is confirmed in Section 3.2.2 hereafter.

By contrast, in Tr. `PLT-G`, inflation returns back on target in a quicker time frame as it veers around the CB's target with negligible fluctuations as shown in Fig. 4. This inflation stabilization in Tr. `PLT-G` is observed under both types of shocks but is most salient after the deflationary shock. In particular, in the wake of the deflationary recession, the ELB binds for a median time of only four periods under Tr.`PLT-G`, which is still more than under RE (see again Fig. 1) but considerably less than the six-or-seven-period-long ELB episodes under IT. We may also look at the credibility index in Tr. `PLT-G` (bottom panel of Fig. 4): while credibility substantially drops at the onset of the recessions, it increases back toward one (full credibility) along the recoveries. Finally, in line with the RE prediction, a better inflation stabilization in Tr. `PLT-G` than in Tr. `IT` comes at the cost of a wider output gap, in particular after a cost-push shock. This is confirmed in the bottom panel of Table 1: inflation and output are significantly better stabilized in Tr. `PLT-G` than in Trs. `IT` but the output gap in Subsession 2 is significantly wider in Tr. `PLT-G` than under Tr. `IT`.

We now turn to the analysis of individual forecasts.

3.2.2 Heterogeneity in individual forecasts

Fig. 5 reports the median short-run forecast dispersion across all six sessions of each treatment measured by the standard deviation across the six subjects' forecasts in each period. Table 2 (top panel) complements the graphical analysis with the corresponding cross-treatment comparative statistics.

Forecast dispersion varies both within treatments, i.e. is time-varying, and be-

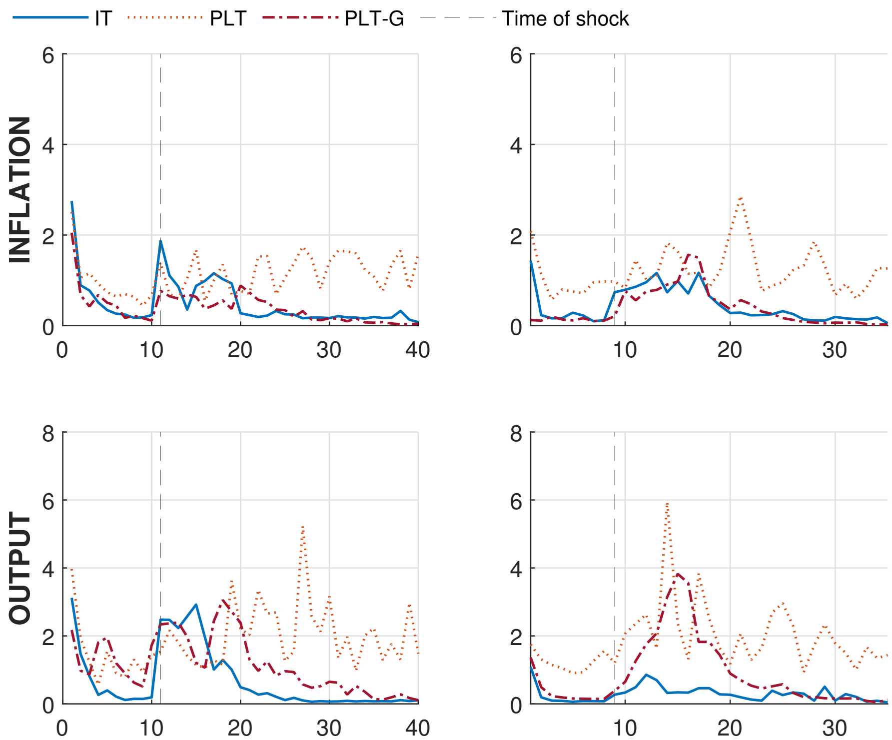
Table 1: Cross-treatment comparative statistics of the experiment data

		Numerical values			*Associated p-values*		
		IT	PLT	PLT-G	IT *vs* PLT	IT *vs* PLT-G	PLT *vs* PLT-G
Top panel: ***Pre-shock AAD*** *(p.p.)*							
π	**Sub 1**	0.62	0.90	0.38	0.18	0.18	0.02
		(0.30)	(0.31)	(0.18)			
	Sub 2	0.78	1.35	0.23	0.31	0.06	0.01
		(0.52)	(0.90)	(0.21)			
y	**Sub 1**	0.40	3.41	1.90	0.00	0.02	0.18
		(0.27)	(1.91)	(1.52)			
	Sub 2	0.37	3.22	0.65	0.00	1.00	0.01
		(0.23)	(1.58)	(0.89)			
Bottom panel: ***Post-shock AAD*** *(p.p.)*							
π	**Sub 1**	1.77	6.54	0.56	0.39	0.00	0.00
		(0.72)	(9.77)	(0.31)			
	Sub 2	2.27	9.51	0.92	0.04	0.06	0.00
		(1.49)	(8.82)	(0.37)			
y	**Sub 1**	1.64	13.86	2.20	0.00	0.18	0.00
		(0.57)	(16.03)	(0.95)			
	Sub 2	1.10	19.54	2.81	0.00	0.01	0.00
		(0.73)	(14.91)	(0.94)			

Note: The average of the absolute deviation (AAD) of inflation π and output y from their target displayed in the first three columns are computed as per, respectively, Eq. (7) and (8), and are averaged across sessions within a treatment. Standard deviations are reported between brackets. The p-values displayed in the last three columns correspond to the non-parametric Wilcoxon rank-sum test for comparison between each pair of treatments.

tween treatments. From Fig. 5, we see that the pre-recession dispersion in forecasts drops to near zero for all sessions under Trs. IT and PLT-G, which suggests that subjects learn to coordinate on the target in the absence of large shocks in these two treatments. Disagreement in Tr. PLT also decreases but does not drop towards zero before the shocks and towards the end of the subsessions as in the other two treatments. This observation shows that more heterogeneity persists in the short-run forecasts of both inflation and output gap even in the absence of large disturbances in the PLT treatment. Interestingly, at the beginning of Subses-

Figure 5: Between-subject disagreement in short-run forecasts



Note: Median over time (x-axis) across all sessions in each treatment in p.p. (y-axis).

sion 2, inflation forecasts are particularly homogeneous in `PLT-G`, which reflects the credibility of the guidance and the resulting coordination on the steady-state inherited from the first subsession.

Post-shock, the forecast dispersion rises in all treatments. In Trs. `IT` and `PLT-G`, in line with the better stabilization of the endogenous variables with respect to Tr. `PLT` discussed in earlier section 3.2.1, forecast dispersion tends to revert back towards its pre-shock level, while it is not the case in Tr. `PLT`. Also in line with the more volatile output gap in Tr. `PLT-G` than in Tr. `IT` after the inflation shock in Subsession 2, output gap forecast dispersion is temporarily larger in Tr. `PLT-G` than in Tr. `IT`. This observation is confirmed by the statistical tests in the top

Table 2: Cross-treatment comparative statistics of the experiment data

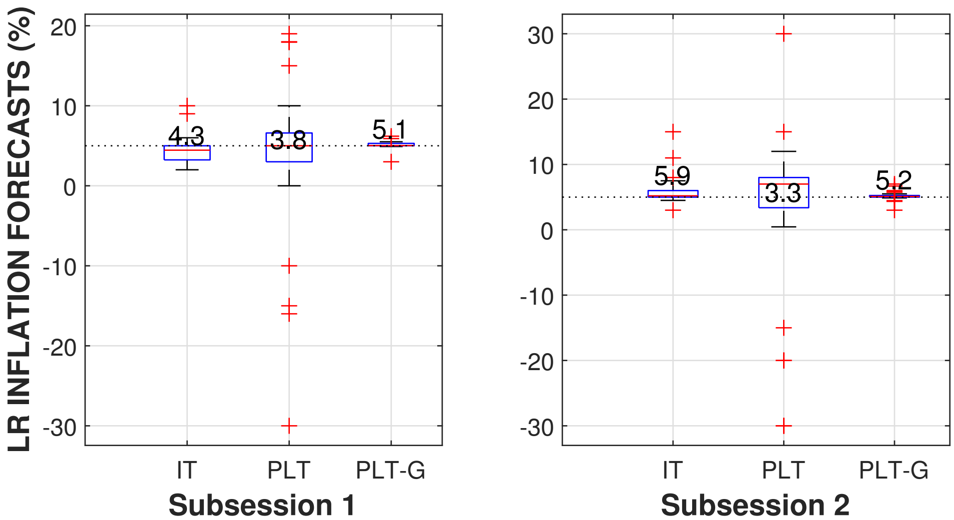
		Numerical values			*Associated p-values*		
		IT	PLT	PLT−G	IT *vs* PLT	IT *vs* PLT−G	PLT *vs* PLT−G
Top panel: ***Short-run forecasts disagreement*** *(p.p.)*							
π	**Sub 1**	0.68	1.51	0.63	0.06	0.82	0.06
		(0.26)	(0.69)	(0.41)			
	Sub 2	0.76	1.72	0.47	0.04	0.59	0.02
		(0.68)	(0.99)	(0.25)			
y	**Sub 1**	0.82	2.94	1.39	0.03	0.18	0.09
		(0.34)	(1.64)	(0.75)			
	Sub 2	0.57	2.82	1.03	0.01	0.09	0.03
		(0.39)	(2.19)	(0.33)			
Bottom panel: ***Long-run inflation forecasts*** *(p.p.)*							
π	**Sub 1**	4.33**	3.83	5.12	0.10	0.00	0.67
		(1.64)	(9.25)	(0.47)			
	Sub 2	5.86**	3.29	5.17	0.23	0.20	0.04
		(2.06)	(11.91)	(0.63)			

Note: Forecast disagreement (first three columns) is computed as the standard deviation across the six forecasts in each period and is averaged over six sessions in a treatment. Standard deviations are reported between brackets. The p-values displayed in the last three columns correspond to the non-parametric Wilcoxon rank-sum test for comparison between each pair of treatments. The superscripts of the long-run inflation forecast values correspond to the t-test of differences with respect to the 5% target at 1% (***), 5% (**) and 10% (*) significance level.

panel of Table 2: disagreement between subjects is significantly lower in Trs. IT and PLT−G than in Tr. PLT, both concerning inflation and output gap and in both shock environments.

The bottom panel of Table 2 reports on the long-run inflation forecasts (for the final period of each subsession). Fig. 6 plots the cross-treatment distributions of these expectations. Tr. PLT−G delivers the best anchoring of the long-run inflation forecasts, which is also evident from Fig. 6: they are not significantly different from the inflation target and the variance among subjects is relatively small compared to the one observed in the other treatments, which is indicative of a lower disagreement about long-run inflation in Tr. PLT−G. By contrast, in Tr. IT, long-

Figure 6: Long-run inflation forecasts of all subjects in all sessions by treatment



Note: the distributions of the long-run inflation forecasts are represented by "boxplots" where the red crosses denote outliers and the numerical values displayed are the mean per treatment.

run expectations are significantly below target in the deflationary environment (Subsession 1) and significantly above target in the inflationary environment (Subsession 2). In Tr. `PLT`, while the long-run forecasts are not significantly different from the target, their variance is highest and fairly large, which indicates a large dispersion in long-run inflation expectations.

To conclude, we highlight the two main experimental results:

Result 1 (IT>PLT) *Contrary to the RE predictions, inflation and output gaps are better stabilized around their target under IT than under PLT, no matter the type of shocks. Short-run expectations are also more homogeneous and long-run expectations are better anchored at the target under IT than PLT.*

Result 2 (PLT-G>IT) *Once the CB communicates in each period the inflation rate that is consistent with closing the price gap, the superior stabilization properties of PLT under*

RE materialize in the lab: shocks have less persistent effects, the ELB binds for fewer periods, and expectations are better coordinated and anchored at the target than under IT.

Before turning to the behavioral model of heterogeneous expectations that we introduce to account for these results, we further dig into individual forecast time series. Doing so allows us to identify which pieces of information subjects use to form their forecasts and find evidence of peer imitation, which is useful to motivate the use of this class of models and specify its assumptions.

3.2.3 Expectation formation and evidence of peer imitation

To give a first impression of the subjects' behavior, we use panel models at the subject-period level to identify which pieces of the available information significantly relate to their inflation and output forecasts in each treatment. The details of the models are deferred to Table 4 Appendix D. The main takeaway is that subjects' forecasts react to the most recent inflation and output gap values, the shocks, and the best inflation forecasts in their group, and, additionally in Tr. `PLT-G`, the CB guidance, but never relate to the past price gap, even under PLT. This indicates that model-consistent expectations are not realized in the history-dependent regime but subjects may imitate their peers and rely on the recent past and the current shocks in all treatments, while also considering the CB guidance in Tr. `PLT-G`.[9]

To refine our description of the forecasts formation and find further evidence of peer imitation in the lab, we estimate heuristics on each subject's forecast time series. We modify the set of first-order heuristics that usually approximate well

[9]The fit is substantially poorer in Tr. `PLT` (Cols. III and IV) than in Tr. `IT` (Cols. I and II) and Tr. `PLT-G` (Cols V and VI), which reflects the greater confusion and the poorer forecasting performances of the subjects in the context of greater macroeconomic volatility in Tr. `PLT` than in the two other treatments. To see that, Table 5 in Appendix D reports the cross-treatment comparisons of the payoff over the entire experimental sessions: subjects' payoff are about one-third to one-half smaller in Tr. `PLT` than in the two other treatments.

the subjects' forecasting behaviors in this class of experiments [10] to include the forecasts of the most accurate forecaster of the group in any given period. We denote this additional piece of information by $\pi^e_{BF,t}$. For each subject, we estimate each forecasting model from the following set:

$$\pi^e_{t+1} = f(\pi^e_{BF,t}) + \epsilon_t, \tag{9}$$

$$\pi^e_{t+1} = f(\pi^e_{BF,t}, \pi^e_t) + \epsilon_t, \tag{10}$$

$$\pi^e_{t+1} = f(\pi_{t-1}, \pi_{t-2}) + \epsilon_t, \tag{11}$$

$$\pi^e_{t+1} = f(\pi_{t-1}, \pi_{t-2}, \pi^e_{BF,t}) + \epsilon_t, \tag{12}$$

$$\pi^e_{t+1} = f(\pi_{t-1}, \pi_{t-2}, \pi^e_{BF,t}, \pi^e_t, F_t) + \epsilon_t, \tag{13}$$

$$\pi^e_{t+1} = f(\pi^e_t, \pi_{t-1}) + \epsilon_t, \tag{14}$$

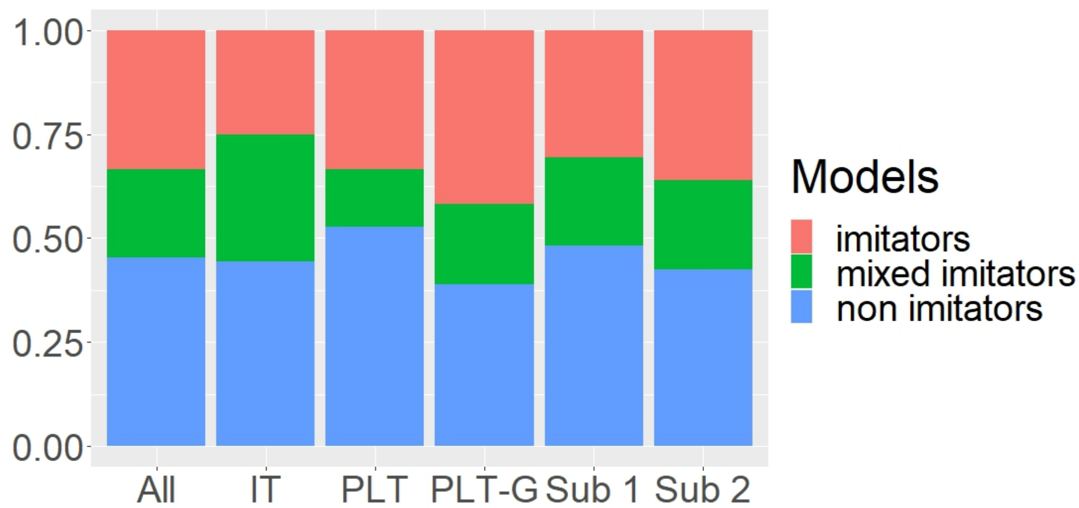
$$\pi^e_{t+1} = f(F_t) + \epsilon_t, \tag{15}$$

$$\pi^e_{t+1} = f(\pi_{t-1}, \pi_{t-2}, F_t) + \epsilon_t, \tag{16}$$

where $f(\cdot)$ is a linear function in its arguments. Models (9), (10) , (12) and (13) make use of the best forecasts and correspond to "imitation models." Model (9) is a "copy-cat" strategy that consists in adopting the best forecast. Model (10) also relies on the individual previous forecasts and may therefore be defined as an "adaptive imitator." Model (12) corresponds to an "anchoring-and-adjustment" rule where the anchor is the best forecast and the adjustments take place in the direction of the most recent inflation trend. Model (13) is a mixed model where all these pieces of information are used to make predictions. Models (11), (14) (15) and (16) deliver, respectively, trend-following expectations, adaptive expectations (including the limiting case of naive expectations), 'fundamentalist' expectations (i.e. anchored at the target in this model) and a mixture of these models, which are the four benchmark in the related literature. We include these standard

[10] See, *inter alia*, Anufriev & Hommes [2012], Assenza et al. [2021], Mokhtarzadeh & Petersen [2021]. To the best of our knowledge, in neither of the existing learning-to-forecast experiments do subjects observe the forecasts of others or their performances. Therefore, forecasting heuristics involving up to a couple of lags, such as Eqs. (11), (14), (15) or (16) are used to describe the subjects' forecasts.

Figure 7: Proportions of imitation, mixed and non-imitation heuristics among all subjects

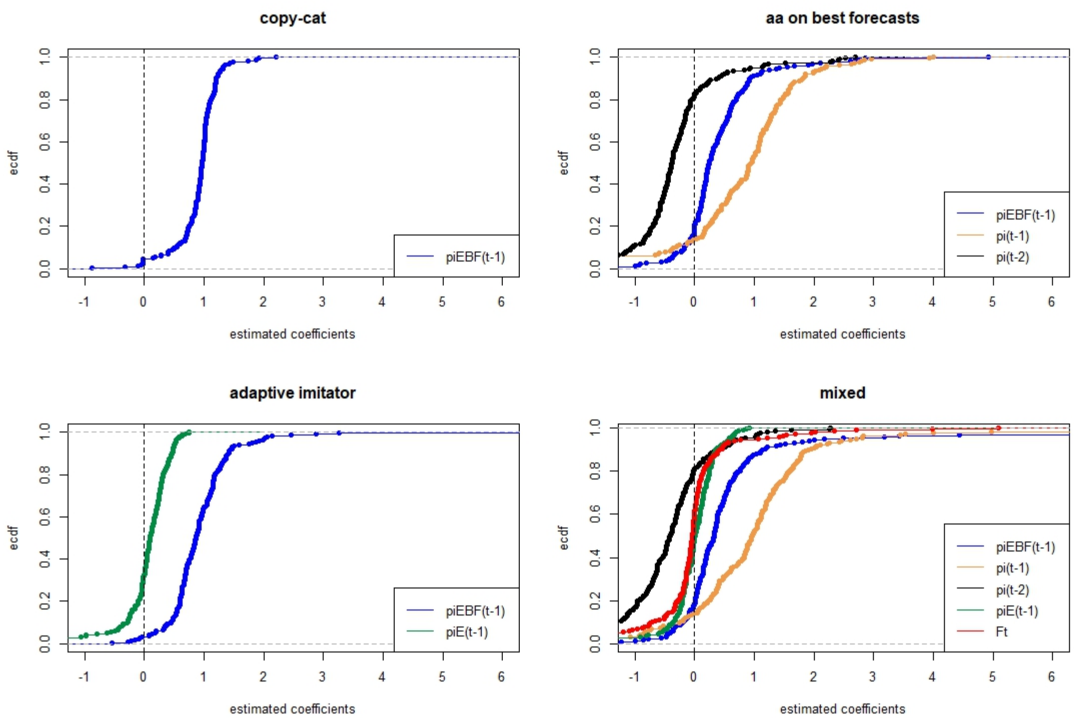


'non-imitation' models to allow for the possibility that subjects ignore the other forecasts in their group and, therefore, do not significantly imitate.

We use the Akaike Information Criterion (AIC) to classify each subject across one of these forecasting heuristics. The results from this exercise are presented in Fig. 7 where we show the fraction of 'imitators' (defined as subjects assigned to Models (9), (10) or (12)), 'mixed imitators' (assigned to Model (13)) and non-imitators (assigned to any other models) over all sessions, per treatment and per subsession.

Overall, we find evidence of peer imitation. A considerable share of subjects uses the best forecasts of their group to form their own forecasts in all treatments and subsessions. A significantly higher fraction of subjects relies, at least partly, on the best forecasts of their group in Tr. PLT-G than in the other treatments (the p-value of the Chi-squared proportion test is 0.03 with respect to Tr. PLT and 0.05 with respect to Tr. IT). This observation is consistent with the guidance of the CB acting as a self-fulfilling focal point so that forecasts are more homogeneous in Tr. PLT-G than in the other treatments. Fig. 8 further plots the distributions

Figure 8: Estimated coefficients of the imitation and mixed models (models 9, 10, 12, and 13) in all sessions



of the estimated coefficients of the imitation models. The coefficient values are consistent with the interpretation of each rule. In the "copycat" model (9), the estimated coefficients associated with the best forecasts are distributed around 1 (see the top-left panel of Fig. 8). In the "anchoring-and-adjustment" rule (12), the coefficients associated with the past two inflation lags are of the same magnitude on the unit interval but of opposite signs, which is consistent with an adjustment in line with the latest inflation trend (see top-right panel). As for the "adaptive imitator" model (10), the weights on the best forecasts and on the last inflation rate also falls on the unit interval, with a higher weight on the best forecasts, in line with a weighted average between the two variables (see bottom-left panel). The estimated coefficients associated with these pieces of information remain of the same order of magnitude in the mixed model (see bottom-right panel).

We now introduce a HENK model with peer imitation through social interactions and show how the resulting behavioral model may account for the main results of the experiment.

4 A behavioral model of the experiment

4.1 A HENK model with peer imitation

The behavioral model builds on the seminal contribution of Arifovic [1996] and relies on the micro-foundations and the general solution method for this class of models derived in Arifovic et al. [2023, Appendix A] and Grimaud et al. [2023], to which we refer for a comprehensive presentation.

4.1.1 Main ingredients of the model

The economy is populated by an infinitely-lived household composed of a finite number N of members indexed by $j = 1, ..., N$. Each member decides about their consumption, savings and labor plans to maximize their expected discounted utility flow with non-separable preferences under their intertemporal budget constraint, which is binding in each period. The production side of the economy is composed of an intermediary-good sector and retailers. Each member of the family operates a firm that produces an intermediate good under monopolistic competition, constant return to scale and price rigidity *à la* Rotemberg.[11] Retailers

[11] These assumptions are not restrictive, nor are they necessary but they simplify the aggregation of heterogeneous behaviors. We refer to Grimaud et al. [2023] for a detailed discussion. For instance, if we consider decreasing returns to scale in the production function, the wages of households may differ because their labor supply may differ as a result of distinct expectations but a mechanism such as a labor packer could equalize labor quantities. To map these homogeneous labor supplies with possibly heterogeneous labor demands, household members are further assumed to evenly split their working hours across all intermediate-sector firms at no cost. Similarly, an intra-household insurance can equalize post-consumption individual wealth patterns; see, e.g., Andrade et al. [2019]. Arifovic et al. [2023, Appendix A.1.7] derive the restrictions on the

aggregate and price the non-perfectly substitutable intermediary goods. Monetary policy is set either under a traditional IT Taylor rule or a PLT rule, depending on the experimental treatment that we aim to match, while fiscal policy is passive.

Household members' plans are conditional upon each member's inflation and output gap expectations which are heterogeneous. We follow the strand of the adaptive learning literature that studies so-called "steady-state learning".[12] The general formulation of SL expectations in the non-linear model then evolves as:

$$E_{j,t}^{SL}\left(X_{t+1}\right) \equiv E_t^{SL}\left(X_{t+1} \mid \xi_{j,t}^x\right) = \bar{X}\exp(\xi_{j,t}^x), \text{ with } X = \Pi, Y, \tag{17}$$

where $E_{j,t}^{SL}(X_{t+1})$ is the one-step-ahead expectation under SL of agent j of variable X conditional on their information set available in period t; this information set includes all past realizations of the endogenous variables X up until period $t-1$ (where Π is gross inflation and Y denotes output) as well as $\exp(\xi_{j,t}^x)$, which is a stochastic process of unconditional mean 1 that represents the private information of agent j about the future realization of X. $\bar{X}$ refers to the steady state value of variable X in the non-stochastic model. The private information of each agent j is subject to idiosyncratic "news" shocks and affected by the interactions under SL (see Section 4.1.2).

Log-linearizing the SL expectations (17) around the targeted steady state results in:

$$E^{SL}(x_{j,t+1}) = \xi_{j,t}^x \text{ with } x = \pi, y, \tag{18}$$

where π and y are now expressed in gaps, i.e. in deviations from their steady-

utility parameters that ensure the correspondence between the model under Rotemberg pricing and the reduced form utilized in the experiment and in the textbook three-equation NK model with separable preferences and Calvo pricing.

[12]This assumption is common in non-linear models; see, *inter alia*, Evans et al. [2008, 2022]. Our HENK model is non-linear not only due to the ELB but most importantly, because of the social dynamics in expectations. In a history-dependent regime, steady-state learning implies restricted perceptions because the perceived law of motion of the agents does not include the lagged variables and is therefore under-parametrized; see, e.g., Branch [2006]. Within the context of our model, subjects did not utilize all the information relevant to form model-consistent expectations as discussed in Section 3.2.3 and, hence, this specification is well-suited.

state values. In the model, each agent j is then uniquely defined by a pair of $(\xi^{\pi}_{j,t}, \xi^{y}_{j,t})$ which stands for their idiosyncratic information about, respectively future inflation and output gaps. Steady-state learning allows us to interpret these expectations as beliefs about the long-run or steady-state values of these variables, which conveniently conveys the concept of the (un)-anchoring of expectations.

4.1.2 Evolution of beliefs under SL

Expectation formation under SL involves two steps. First, agents receive a private signal, or news, about the variables that they need to forecast. Second, they interact with each other, exchange their idiosyncratic information and form their expectations using the information that has proven most relevant to forecast the endogenous variables in the recent past.[13]

Formally, at the beginning of each period t, each agent j first receives a news about each endogenous variable $x = \{y, \pi\}$, denoted by $\iota^{x}_{j,t}$, and updates their belief $\xi^{x}_{j,t-1}$ as follows:

$$m^{x}_{j,t} = \xi^{x}_{j,t-1} + \iota^{x}_{j,t}, \tag{19}$$

where $m^{x}_{j,t}$ refers to their "*m*utated" or updated forecast. The news ι^{x}, $x = \pi, y$ are random draws with unconditional mean zero. The exact distributions of the news are the only processes that need to be specified to close the model – besides the exogenous disturbances g and u – which keeps the behavioral model parsimonious. In what follows, we use $\iota^{x}_{j,t} \sim \mathcal{N}(\theta^{x}(g_t, u_t), \sigma^2_x)$, i.e. the news are idiosyncratic Gaussian random draws, where the 'average news' depends on the latest observable shocks g_t and u_t and, hence, the private signals convey fundamental information about the endogenous variables.[14] As usual in the learning

[13] The SL algorithm is inspired by the use of genetic algorithms to solve global optimization problems, where two operators are mobilized: i) a 'mutation' operator, that is the exploration mechanism of the state space, and ii) an 'imitation' operator, that is an exploitation mechanism of the information collected during the exploration phase which selects the values that optimize the objective function.

[14] Note that this assumption does not conflict with Eq. (17) because the shocks have an uncondi-

literature, agents do not know the true parameter values of the model but in the same way as the lab subjects are informed about these qualitative effects in the instructions of the experimental game, they have a qualitative understanding of the effects of the shocks on the economy. In this respect, we impose $\frac{\partial\theta^{\pi}}{\partial g}, \frac{\partial\theta^{\pi}}{\partial u}, \frac{\partial\theta^{y}}{\partial g} \geq 0$ and $\frac{\partial\theta^{y}}{\partial u} \leq 0$, so that the negative real shock g in Subsession 1 tends to bring, on average, lower inflation and output gap outlooks and a positive cost-push shock in Subsession 2 tends to bring inflationary and recessionary prospects. Furthermore, a higher noise σ_x^2 translates into more heterogeneity, that is more dispersion, among agents' news.

After receiving their news, agents assess the relative value of their private information through social interactions and imitate their peer who have the most relevant news, in the spirit of the "survival-of-the-fittest"-metaphor underlying evolutionary algorithms used for optimization purposes. In line with the payoff in the experiment that is based on forecast errors, the relevance of any individual forecast is assessed using the 'Foregone' forecast error, denoted by $F_{k,t}^{x}$, that is the error that would have resulted would the updated forecast $m_{j,t}^{x}$ have been used in the last period, namely: $F_{k,t}^{x} = (\widehat{x}_{t-1} - m_{j,t}^{x})^2$, for $x = \pi, y$ (lower values denote a *more* relevant news).

Social interactions are modeled *via* a behavioral model of peer imitation called 'tournament.' In each period, each agent is randomly paired with another agent,[15] against who they compare the relative relevance of their news. The agent with the highest foregone forecast errors discards their news and instead adopts the forecast of the other agent in the pair. There may be two distinct tournaments for inflation and output gaps. Formally, in each period, for each pair of agents

tional 0 mean.

[15]Without loss of generality, the number of agents is chosen to be even. Pairing is performed through random draws with equal probability and without replacement. It is also possible to model the diffusion of news in a population using probabilistic choice models, see e.g. Brock & Hommes [1997], Anufriev & Hommes [2012]. Our results are robust to this alternative.

$k \neq l \in J$, their resulting beliefs are given by:

$$(\xi^y_{k,t}, \xi^y_{l,t}) = \mathbb{1}_{F^x_{k,t} > F^x_{l,t}}(m^x_{l,t}, m^x_{l,t}) + \mathbb{1}_{F^x_{k,t} \leq F^x_{l,t}}(m^x_{k,t}, m^x_{k,t}), x = \pi, y. \tag{20}$$

Eqs. (19) and (20) determine the distribution of heterogeneous expectations $\{\xi^\pi_{j,t}, \xi^y_{j,t}\}$ in Eq. (18). The rest of the model is given by the same set of equations as in the experiment, namely Eqs. (1), (2) and (3) in Tr. IT or (1), (2), (4) and (5) in Tr. PLT. Hence, the only difference between the experimental environment of Section 3 and the behavioral model is the origin of the expectations: while these expectations are elicited from human subjects in the experiment, they are produced by agents under SL in the HENK model.

Before turning to the comparison between the SL simulation results and the experimental results, the CB guidance needs to be integrated into the SL algorithm to account for Tr. PLT-G.

4.1.3 CB guidance in the SL beliefs

In Tr. PLT-G, the CB provides the time-varying inflation targets π^{CB}_t consistent with the closure of the price gap. Therefore, we need to modify how inflation expectations are formed by the SL agents to account for this public signal. We add a third component to the individual pair of beliefs that corresponds to the probability for a given agent j of using the public signal rather than their private signal as their inflation forecast. We denote this probability by $c_{j,t} \in [0, 1]$ (for "*c*redibility" of the CB guidance, where credibility is defined as the probability of anchoring one's inflation forecast to the announcement of the CB).[16] The complete set of beliefs for each agent in each period t becomes $\{\xi^y_{j,t}, (\xi^\pi_{j,t}, c_{j,t})\}$. Probabilities $c_{j,t}$ also evolve as a result of news and peer imitation during social interactions so that agents learn to follow (when $c_{j,t}$ increases) or ignore (when $c_{j,t}$ decreases) the CB guidance. In particular, if the guidance grows disconnected from the recent

[16] The credibility $\bar{c}_t$ of the CB in the economy may be measured by the average of the individual probabilities: $\bar{c}_t = \frac{1}{N}\sum_j^N c_{j,t}$.

inflation developments, agents tend to discard it and follow their private news instead (and the other way around if the guidance turns relevant to the recent inflation data).

Formally, in the model under Tr. `PLT-G`, only two modifications to the framework developed in Section 4.1.2 are necessary to account for the CB guidance. First, the updated inflation forecast of agent j is given by:

$$m_{j,t}^{\pi} = c_{j,t}\pi_t^{CB} + (1 - c_{j,t})\,\xi_{j,t}^{\pi} \tag{21}$$

while the output gap forecast is formed as described in Section 4.1.2.

Second, the relevance of any updated information pertaining to inflation $\left(m_{j,t}^{\pi}, c_{j,t}\right)$ becomes a weighted average of the foregone forecast errors that would have resulted from using the private signal $m_{j,t}^{\pi}$ or the CB guidance π_t^{CB}, i.e. $F_{j,t}^{\pi} = c_{j,t}\left(\hat{\pi}_{t-1} - \pi_t^{CB}\right)^2 + (1 - c_{j,t})\left(\hat{\pi}_{t-1} - m_{j,t}^{\pi}\right)^2$.

In the next section, we compare the model outcomes in the three treatments to the experimental results.

4.2 Model outcomes

4.2.1 Cross-treatment comparisons

Figs. 9, 10 and 11 plot the model counterparts of Figs. 2, 3 and 4 in respectively, Trs. `IT`, `PLT` and `PLT-G`. At first glance, it is striking to see that the behavioral model can account for the main stylized facts of the experiment, namely the following five observations.

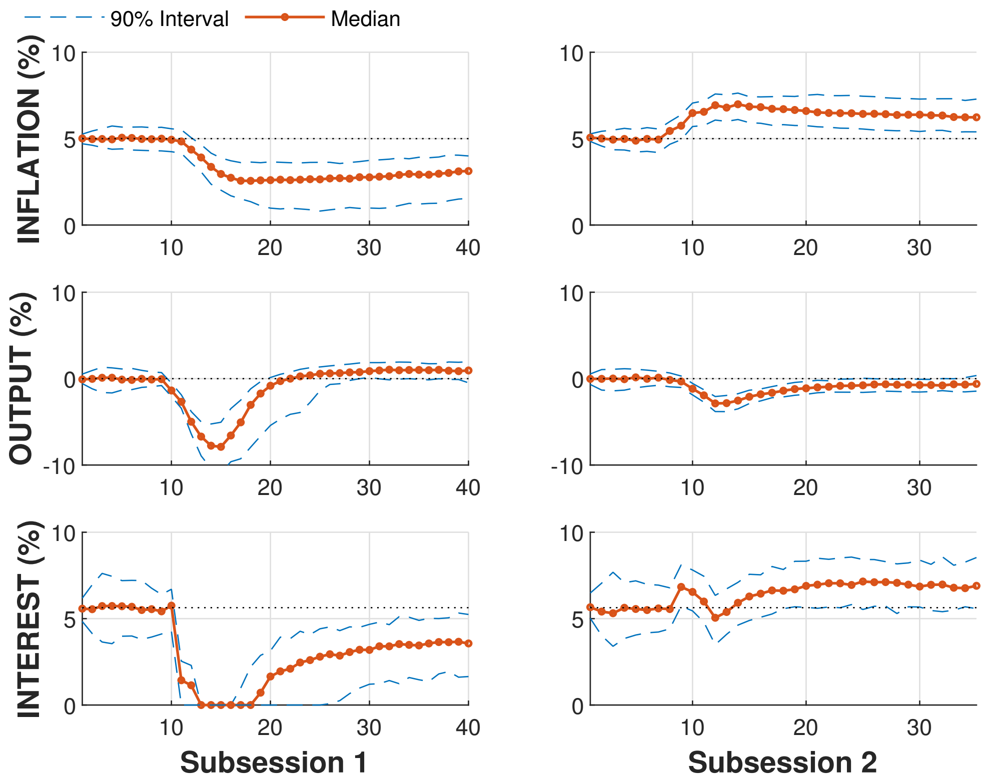
First, inflation and output exhibit wilder and non-dampening oscillations in Tr. `PLT` compared to the stabilization observed towards the end of the subsessions in Trs. `IT` and `PLT-G`. Second, the inflation gap in the aftermath of the shocks is more persistent under `IT` than under `PLT-G`, where inflation quickly settles back on target along mild and symmetric oscillations around the target. By contrast, under IT, the inflation gap remains negative in the first subsession and positive in the second one, in line with the experimental and empirical evidence. Third, this persistent inflation gap translates into a fairly long episode of low nominal interest rates in Subsession 1 under IT, which echoes the 'low-for-long' narrative before 2022. By contrast, under PLT with guidance, the nominal rate converges back to its steady state level after a few periods. Fourth, the model accounts well for the larger recession resulting from the cost-push shock in Subsession 2 under PLT with guidance than under IT. Fifth, under Tr. `PLT-G`, SL agents learn to follow the guidance in the first periods of the model (to see that, look at the bottom panel of Fig. 11 where the average strategy c across agents is reported). While CB credibility suffers a temporary deep at the onset of the recessions, it recovers towards a strong level (close to one) by the end of the subsessions.

These cross-treatment differences are confirmed by statistical tests in Table 3.[17] At least qualitatively, Table 3 replicates its counterpart Table 1. Let us first focus on the top and the middle panels. Before the shocks (see top panel), Trs. `IT` and `PLT-G` exhibit similar AAD-levels but the picture changes after the shocks (middle panel). `PLT-G` has the lowest inflation AAD and `IT` has the lowest output AAD, implying a quicker closure of the inflation gap under PLT with guidance, but at a higher stabilization cost in terms of output than under IT. Without guidance, PLT exhibits the largest AADs. The differences for inflation are not as striking as for output because the large oscillations in Tr. `PLT` in the experiment takes more time to materialize in the simulations than in the experiment. When extrapolating the model for a few more periods, it becomes clear that the oscillations grow wilder, in line with the dynamics observed in Tr. `PLT` in the experiment (see Fig. 13 in appendix D).

Finally, the HENK model is able to capture a sixth stylized fact of the experiment

[17] Of course, with many more independent observations in the model than in the experiment, we obtained much smaller p-values and more statistically significant treatment differences than in Table 1 pertaining to the experimental data.

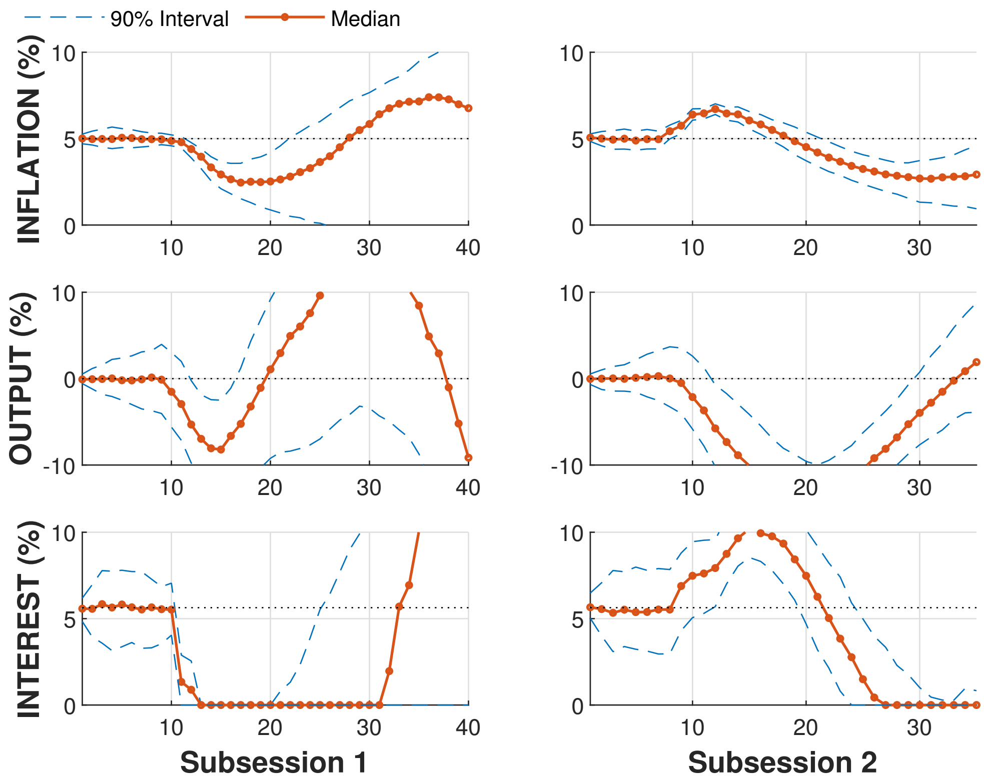
Figure 9: Inflation, output and interest rate dynamics in `IT` treatment in NK-SL



Notes: 1,000 Monte Carlo simulations of the HENK model, $\sigma_y^2 = 0.1$, $\sigma_\pi^2 = 0.02$, $\sigma_c^2 = 0.15$.

pertaining to the cross-treatment dispersion in individual forecasts. Comparing the bottom panel of Table 3 with its experimental counterpart Table 2, one can see that the best coordination among inflation forecasts (i.e. the lowest cross-sectional dispersion) is obtained in Tr. `PLT-G`, which reflects the coordination of expectations on the CB's guidance in this treatment. Furthermore, the best coordination among output gap forecasts is observed under IT, due to the stronger fluctuations in output under a PLT rule than under a standard IT rule, in particular in face of cost-push shocks (in Subsession 2). For both inflation and output, individual forecasts are the most heterogeneous under Tr. `PLT`, in the experiment as well as in the model.

Figure 10: Inflation, output and interest rate dynamics in `PLT` treatment in NK-SL



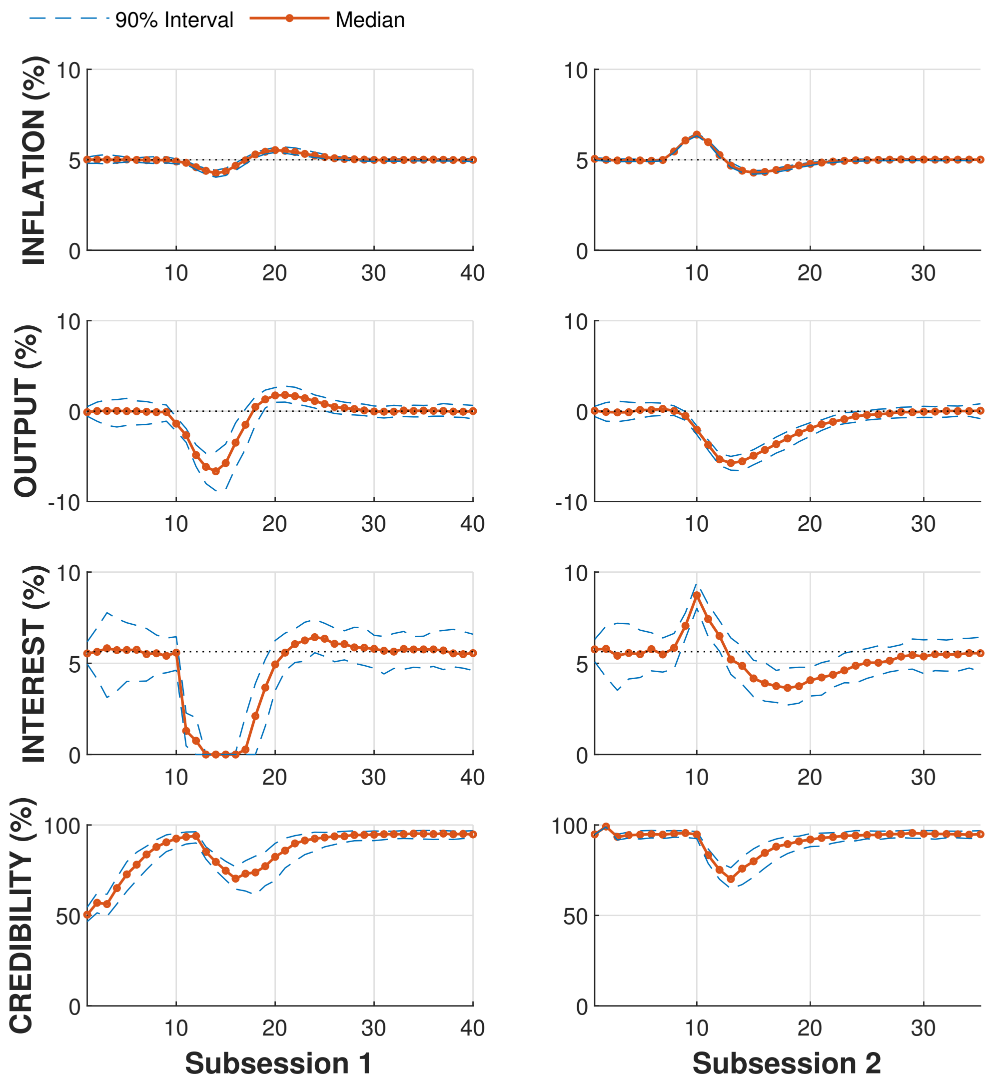
Notes: see Fig. 9.

Given that our behavioral model explains well the key empirical findings obtained in the lab, we now discuss further the behavioral reasoning behind the observed dynamics.

4.2.2 Discussion

Unraveling the expectation dynamics under SL may help shed light on the mechanisms at play in the experiment. In particular, persistence is a common feature of this learning mechanism [Bullard 2023]. Under IT, the persistent inflation gap stems from long-lasting pessimistic deflationary (in Subsession 1) and inflationary

Figure 11: Inflation, output and interest rate dynamics in `PLT-G` treatment in NK-SL



Notes: see Fig. 9. The agents' beliefs $\{c_{j,0}\}$ are initialized based on the experimental data in Figure 4. In particular, credibility in Subsession 2 is inherited from Subsession 1.

(in Subsession 2) outlooks: at the onset of the recession, SL agents adopt off-target forecasts which, in combination with the severity and persistence of the recessionary shocks, quickly become contagious because inflation expectations are almost self-fulfilling per the NKPC Eq. (2). Hence, agents learn to live in a low (or high) inflation environment. This mechanism generates lasting coordination on pessimistic expectations, both in the experiment and in the model.

With stable yet off-target inflation forecasts, inflation can be alternatively brought back on target by pressure from aggregate demand. Due to the nominal rigidities of the NK economy (i.e. a small slope κ in Eq. (2)), it requires a sufficiently large change in output to turn around inflation. Yet, this takes time to unfold because the IT-CB starts adjusting the interest rate to movements in output, which slows down the lifting of inflation. Post-recession, the economy enters a "new normal", where both inflation and interest rate settle on an off-target path. Consequently, both SL agents and the participants in the experiment will not coordinate on a different inflation expectation path unless actual inflation leads the way. Put differently, the lack of history-dependence in the monetary policy rule limits the adjustment of aggregate demand that would be necessary to move inflation and, consequently, expectations toward the target.

Table 3: Cross-treatment comparative statistics of the simulation data

		Numerical values			*Associated p-values*		
		IT	PLT	PLT−G	IT *vs* PLT	IT *vs* PLT−G	PLT *vs* PLT−G
Top panel: ***Pre-shock AAD*** *(p.p.)*							
π	**Sub 1**	0.38	0.27	0.11	0.00	0.00	0.00
		(0.23)	(0.14)	(0.05)			
	Sub 2	0.39	0.33	0.1	0.06	0.00	0.00
		(0.21)	(0.15)	(0.01)			
y	**Sub 1**	0.75	1.66	0.78	0.00	0.47	0.00
		(0.34)	(0.90)	(0.35)			
	Sub 2	0.64	1.20	0.53	0.00	0.01	0.00
		(0.31)	(0.62)	(0.25)			
Middle panel: ***Post-shock AAD*** *(p.p.)*							
π	**Sub 1**	2.29	2.44	0.23	0.32	0.00	0.00
		(1.23)	(1.34)	(0.05)			
	Sub 2	1.50	1.53	0.31	0.79	0.00	0.00
		(0.63)	(0.47)	(0.02)			
y	**Sub 1**	2.85	10.71	1.65	0.00	0.00	0.00
		(2.48)	(3.00)	(0.38)			
	Sub 2	1.23	7.93	1.93	0.00	0.00	0.00
		(0.32)	(1.90)	(0.22)			
Bottom panel: ***Short-run forecasts disagreement*** *(p.p.)*							
π	**Sub 1**	1.64	1.65	0.74	0.01	0.00	0.00
		(0.04)	(0.05)	(0.06)			
	Sub 2	1.63	1.64	0.73	0.01	0.00	0.00
		(0.04)	(0.05)	(0.06)			
y	**Sub 1**	8.17	8.32	8.23	0.00	0.10	0.01
		(0.24)	(0.25)	(0.24)			
	Sub 2	8.12	8.27	8.18	0.00	0.07	0.01
		(0.23)	(0.24)	(0.23)			

Note: The average absolute deviation (AAD) of inflation π (first two rows of the top and middle panels, see Eq. (7)) and output y (second two rows of the top and middle panels, see Eq. (8)) and the forecast disagreement (bottom panel), computed as the standard deviation across all SL agents' forecasts in each period, are displayed in the first three columns and are averaged across all simulations of each treatment. Standard deviations are reported between brackets. The p-values displayed in the last three columns correspond to the non-parametric Wilcoxon rank-sum test for comparison between each pair of treatments.

By contrast, the PLT-CB reacts in a more accommodating (in Subsession 1) or restrictive (in Subsession 2) way to the shocks than under IT to close the price gap. This stronger reaction pressurizes prices and output up (or down) to escape the persistent off-target region. This reaction prompts agents to quickly discard off-target inflation forecasts and edge closer to the inflation guidance of the CB. This more aggressive interest-rate response is what causes the deeper recession in Subsession 2 in Tr. `PLT-G` than in Tr. `IT`. Per the self-fulfilling property of inflation in Eq. (2), the CB announcement becomes self-validating and inflation expectations turn consistent with the monetary policy rule and its objective of price-gap closure. Moreover, if credibility is sufficiently high at the onset of the shocks, it tends to carry on because of the self-fulfilling property of inflation expectations and the contagion under social interactions. As a result, a credible guidance limits the movements in inflation expectations caused by the shocks because part of the agents form model-consistent expectations. This process ensures a reinforcing loop between credibility and macroeconomic stability.

Without this loop, namely under Tr. `PLT`, expectations (whether under SL or in the lab) follow the recent inflation developments instead of reacting to the past price gap that is a key component of model-consistent expectations under a PLT rule. As a result, inflation plunges deeper following the deflationary shock in Subsession 1 and increases higher following the cost-push shocks in Subsession 2 than under `PLT-G`. These larger deviations accumulate in the price gap so that it remains well below or above zero even when output has recovered and the PLT-CB reacts by extending its recession responses rather than normalizing its policy rate, as does the IT-CB. Hence, in Subsession 1, the ELB episode is longer than in the other treatments, in an attempt to boost inflation to close the price gap but at the price of a large overheating of the economy. As a result, higher inflation and output forecasts become more relevant under SL and further feed the boom up until the point that this upward course of both inflation and output becomes problematically enduring (especially for output). The PLT-CB then reacts by lifting the interest rate to fight off the positive output gaps (but with a long delay because of the time required to close the persistent negative price gap inherited from the shock). By the time the PLT-CB finally abandons the ELB policy in Subsession 1 (or the high-rate policy in Subsession 2), the output gap is particularly large (positive in Subsession 1 and negative in Subsession 2), which necessitates a particularly brutal adjustment of the interest rate. These brutal adjustments, amplified by the reaction of unanchored and contagious expectations, generate the large unsettling oscillations observed in both the experiment and the model.

5 Conclusion

The contributions of this paper are twofold. First, we empirically find supporting evidence for the merits of the history-dependent PLT policy – a monetary framework that has lacked empirical evidence so far – in stabilizing inflation and output gaps in face of both real and cost-push shocks. We do so by utilizing a forecasting laboratory experiment where we vary the monetary policy rule in a between-subject design and expectations are elicited from human subjects. Second, we provide a micro-founded HENK model based on the diffusion of private news in the market through peer imitation. This parsimonious and intuitive model can account remarkably well for the observed cross-treatment differences in the lab and offer a behavioral description of these differences.

Our findings are important to CB affairs for at least two reasons. First, our findings stress the importance of a carefully designed CB communication strategy in case of the adoption of a history-dependent policy rule. By carefully designed CB communication, we mean communication that is framed in terms of inflation and conveys the relationship between the future inflation gap and the past price gap. This relationship is key for the theoretical benefits of history-dependent regimes to materialize under real-world expectations but unfortunately, it seems particularly hard to grasp by human subjects. In this context, CB communication, while redundant under RE, can convey this relationship to them, at least when it is framed in terms of inflation forecasts. Our finding furthermore suggests that this strategy is robust to the type of shocks. A history-dependent regime with a sufficiently credible communication can lead agents to form model-consistent expectations and, by doing so, steer inflation expectations to the target, under both deflationary or inflationary pressures. Second, and more broadly, we have presented the added value of a method, based on the combination of lab evidence with a HENK model to analyze monetary policy rules. This framework is an interesting complement to the CB's policy analysis toolkit because it may easily be deployed to evaluate alternative communication contents and policy rules.

References

Andrade, P., Gaballo, G., Mengus, E. & Mojon, B. [2019], 'Forward guidance and heterogeneous beliefs', *American Economic Journal: Macroeconomics* **11**(3), 1–29.

Anufriev, M. & Hommes, C. [2012], 'Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments', *American Economic Journal: Microeconomics* **4**(4), 35–64.

Anufriev, M., Hommes, C. & Makarewicz, T. [2019], 'Simple forecasting heuristics that make us smart: Evidence from different market experiments', *Journal of the European Economic Association* **17**(5), 1538–1584.

Arifovic, J. [1996], 'The behavior of the exchange rate in the genetic algorithm and experimental economies', *Journal of Political economy* **104**(3), 510–541.

Arifovic, J. [2000], 'Evolutionary algorithms in macroeconomic models', *Macroeconomic Dynamics* **4**(3), 373–414.

Arifovic, J., Bullard, J. & Kostyshyna, O. [2013], 'Social learning and monetary policy rules', *The Economic Journal* **123**(567), 38–76.

Arifovic, J., Grimaud, A., Salle, I. & Vermandel, G. [2023], Social learning and monetary policy at the effective lower bound. Available at SSRN, `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3728108`, revised 2023.

Arifovic, J. & Petersen, L. [2017], 'Stabilizing expectations at the zero lower bound: Experimental evidence', *Journal of Economic Dynamics and Control* **82**, 21–43.

Arifovic, J., Schmitt-Grohé, S. & Uribe, M. [2018], 'Learning to live in a liquidity trap', *Journal of Economic Dynamics and Control* **89**, 120–136.

Assenza, T., Heemeijer, P., Hommes, C. H. & Massaro, D. [2021], 'Managing self-organization of expectations through monetary policy: A macro experiment', *Journal of Monetary Economics* **117**, 170–186.

Bodenstein, M., Hebden, J. & Winkler, F. [2022], 'Learning and misperception of makeup strategies', *Journal of Economic Dynamics and Control* **139**, 104417.

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

Branch, W. A. [2006], 'Restricted perceptions equilibria and learning in macroeconomics', *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model. Cambridge University Press, New York* pp. 135–160.

Branch, W. A. & McGough, B. [2010], 'Dynamic predictor selection in a new keynesian model with heterogeneous expectations', *Journal of Economic Dynamics and Control* **34**(8), 1492–1508.

Brock, W. A. & Hommes, C. H. [1997], 'A Rational Route to Randomness', *Econometrica* **65**(5), 1059–1095.

Bullard, J. [2023], 'Social Learning for the Masses'. Federal Reserve Bank of St. Louis, Speech at the Computational & Experimental Economics Workshop, Simon Fraser University, Vancouver, British Columbia.

Bullard, J. & Mitra, K. [2002], 'Learning about monetary policy rules', *Journal of monetary economics* **49**(6), 1105–1129.

Busetti, F., Delle Monache, D., Gerali, A. & Locarno, A. [2017], Trust, but verify. De-anchoring of inflation expectations under learning and heterogeneity, Technical Report 1994, ECB working paper.

Carroll, C. & Wang, T. [2023], Epidemiological expectations, *in* 'Handbook of Economic Expectations', Elsevier, pp. 779–806.

Cateau, G., Kryvtsov, O., Shukayev, M. & Ueberfeldt, A. [2009], Adopting Price-Level Targeting under Imperfect Credibility in ToTEM, Staff Working Papers 09-17, Bank of Canada.

Cecchetti, S. G., Krause, S. et al. [2002], 'Central bank structure, policy efficiency, and macroeconomic performance: exploring empirical relationships', *Review-Federal Reserve Bank of Saint Louis* **84**(4), 47–60.

Chen, D. L., Schonger, M. & Wickens, C. [2016], 'otreean open-source platform for laboratory, online, and field experiments', *Journal of Behavioral and Experimental Finance* **9**, 88–97.

Coibion, O., Gorodnichenko, Y. & Kamdar, R. [2018], 'The formation of expectations, inflation, and the Phillips curve', *Journal of Economic Literature* **56**(4), 1447–91.

Coibion, O., Gorodnichenko, Y., Kumar, S. & Ryngaert, J. [2021], 'Do You Know that I Know that You Know? Higher-Order Beliefs in Survey Data', *The Quarterly Journal of Economics* **136**(3), 1387–1446.

Cornand, C. & M'baye, C. K. [2018], 'Does inflation targeting matter? an experimental investigation', *Macroeconomic Dynamics* **22**(2), 362–401.

D'Acunto, F., Malmendier, U. & Weber, M. [2022], What do the data tell us about inflation expectations?, Technical Report 29825, National Bureau of Economic Research.

Del Negro, M., Giannoni, M. P. & Patterson, C. [2023], 'The forward guidance puzzle', *Journal of Political Economy Macroeconomics* **1**(1), 000–000.

Evans, G., Guse, E. & Honkapohja, S. [2008], 'Liquidity traps, learning and stagnation', *European Economic Review* **52**(8), 1438–1463.

Evans, G., Honkapohja, S. & Mitra, K. [2022], 'EXPECTATIONS, STAGNATION, AND FISCAL POLICY: A NONLINEAR ANALYSIS', *International Economic Review* **63**(3), 1397–1425.

Gasteiger, E. [2014], 'Heterogeneous Expectations, Optimal Monetary Policy, and the Merit of Policy Inertia', *Journal of Money, Credit and Banking* **46**(7), 1535–1554.

Grimaud, A., Salle, I. & Vermandel, G. [2023], 'Social learning expectations: microfoundations and a Dynare toolbox', Available at SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4437177.

Guerrieri, L. & Iacoviello, M. [2015], 'Occbin: A toolkit for solving dynamic models with occasionally binding constraints easily', *Journal of Monetary Economics* **70**, 22–38.

Hachem, K. & Wu, J. C. [2017], 'Inflation announcements and social dynamics', *Journal of Money, Credit and Banking* **49**(8), 1673–1713.

Hommes, C. [2021], 'Behavioral and experimental macroeconomics and policy analysis: A complex systems approach', *Journal of Economic Literature* **59**(1), 149–219.

Hommes, C. & Lustenhouwer, J. [2019], 'Inflation targeting and liquidity traps under endogenous credibility', *Journal of Monetary Economics* **107**, 48–62.

Hommes, C. & Makarewicz, T. [2021], 'Price level versus inflation targeting under heterogeneous expectations: a laboratory experiment', *Journal of Economic Behavior & Organization* **182**, 39–82.

Honkapohja, S. & Mitra, K. [2020], 'Price level targeting with evolving credibility', *Journal of Monetary Economics* **116**, 88–103.

Jia, C. & Wu, J. C. [2021], Average Inflation Targeting: Time Inconsistency And Intentional Ambiguity, Working Papers 21-19R, Federal Reserve Bank of Cleveland.

Kostyshyna, O., Petersen, L. & Yang, J. [2022], A horse race of monetary policy regimes: An experimental investigation, Technical report, National Bureau of Economic Research.

Massaro, D. [2013], 'Heterogeneous expectations in monetary dsge models', *Journal of Economic Dynamics and Control* **37**(3), 680–692.

Mitra, K. [2003], 'Desirability of nominal gdp targeting under adaptive learning', *Journal of Money, Credit and Banking* pp. 197–220.

Mokhtarzadeh, F. & Petersen, L. [2021], 'Coordinating expectations through central bank projections', *Experimental economics* **24**, 883–918.

Orphanides, A. & Williams, J. [2004], Imperfect knowledge, inflation expectations, and monetary policy, *in* 'The inflation-targeting debate', University of Chicago Press, pp. 201–246.

Ozden, T. [2021], Heterogeneous Expectations and the Business Cycle at the Effective Lower Bound, Technical Report 714, De Nederlandsche Bank Working Paper.

Powell, J. H. et al. [2020], 'New Economic Challenges and the Fed's Monetary Policy Review: a speech at "Navigating the Decade Ahead: Implications for Monetary Policy," an economic policy symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 27, 2020'.

Rholes, R. & Petersen, L. [2021], 'Should central banks communicate uncertainty in their projections?', *Journal of Economic Behavior & Organization* **183**, 320–341.

Salle, I. [2021], 'What to target? insights from a lab experiment', *Tinbergen Institute Discussion Paper 2021-081/II* .

Svensson, L. E. O. [1999], 'Price-Level Targeting versus Inflation Targeting: A Free Lunch?', *Journal of Money, Credit and Banking* **31**(3), 277–295.

Wilkins, C. A. [2018], 'Choosing the best monetary policy framework for Canada'. Remarks at the Conference "Toward 2021: Reviewing the monetary policy framework," Max Bell School of Public Policy, McGill University, Montreal.

Woodford, M. [2000], 'Pitfalls of forward-looking monetary policy', *American Economic Review* **90**(2), 100–104.

Woodford, M. [2003], *Interest and prices: Foundations of a theory of monetary policy*, Princeton University Press.

A Theoretical predictions and guidance specification

The IT equations (1), (2), (3) can be reduced to a two-variable matrix form:

$$\mathbf{z}_t = \mathbf{A} + \mathbf{B}E_t\mathbf{z}_{t+1} + \chi\epsilon_t, \tag{22}$$

and the PLT equations (1), (2), (4) (5) to a three-variable matrix form:

$$\tilde{\mathbf{z}}_t = \mathbf{A} + \mathbf{B}E_t\tilde{\mathbf{z}}_{t+1} + \mathbf{C}\tilde{\mathbf{z}}_{t-1} + \chi\epsilon_t, \tag{23}$$

where $\mathbf{z} = (\hat{y}, \hat{\pi})'$, $\tilde{\mathbf{z}} = (\hat{y}, \hat{\pi}, \hat{p})'$ and $\epsilon = (g, u)'$. $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and χ are matrices of parameters that depend on whether the ZLB is binding or not. Details of these matrices can be found hereafter. The data-generating process in the experiment is obtained from Systems (22) or (23).

Remark 1 *The necessary and sufficient conditions for the non-ZLB rational expectation equilibrium under IT to be locally determinate is $\kappa(\phi^\pi - 1) + (1-\beta)\phi^y > 0$ (see Bullard & Mitra [2002]).*

Remark 2 *The necessary and sufficient conditions for the non-ZLB rational expectation equilibrium under PLT to be locally determinate is $\phi^P > 0$ and $\phi^y \geq 0$ (Proof: see Appendix B.3 below).*

The rational expectation equilibrium (REE) of both economies are solved hereafter.

At the targeted steady state, the policy parameters, as mentioned in the main text $\phi^y_{IT} = \phi^y_{PLT} = 1$, $\phi^\pi = \phi^P = 1.5$, satisfy both remarks. To construct the guidance specification under `PLT-G`, we first solve for the rational expectation equilibrium in the PLT model using the method of undetermined coefficients, i.e. MSV-REE (see hereafter). The MSV-REE solution under non-ZLB binding PLT takes the form:

$$\tilde{\mathbf{z}}_t = \mathbf{\Theta}\tilde{\mathbf{z}}_{t-1} + \mathbf{X}\epsilon_t, \tag{24}$$

where $\mathbf{\Theta}$ and $\mathbf{X}$ are the MSV-REE matrices. The equilibria include the lagged variable $\tilde{\mathbf{z}}_{t-1}$. Therefore, the equilibria under PLT are path-dependent as the price gap contains all previous inflation gaps (per Eq. (4)).

It is worthwhile to shortly digress from the guidance specification and discuss the property of $\mathbf{\Theta}$, which carries the distinctive feature of the PLT regime, which is in stark contrast to IT, where the MSV-REE solution under non-ZLB binding reduces to $\mathbf{z}_t = \mathbf{X}\epsilon_t$, which depends only on the shocks and not on the history of the endogenous variables. Given our parameterization, $\mathbf{\Theta}$ takes these approximate numerical values:

$$\begin{bmatrix} 0 & 0 & -0.78 \\ 0 & 0 & -0.18 \\ 0 & 0 & 0.56 \end{bmatrix}.$$

Abstracting from the shock structure, the MSV-REE reduces to $\hat{y}_t = -0.78\hat{p}_{t-1}$, $\hat{\pi}_t = -0.18\hat{p}_{t-1}$ and $\hat{p}_t = 0.56\hat{p}_{t-1}$. The two negative entries indicate that any deviations of inflation from the target that cause the price level to be off-track are corrected by opposite and hawkish movements in output and inflation until the price gap is closed.

B MSV-REE derivations

B.1 PLT

We use the method of undetermined coefficients to solve for the MSV-REE under PLT. Abstracting again from the shock structure, under PLT, there are three state variables: the output gap, the inflation gap, and the price gap. We define $\widetilde{\mathbf{z}}_t = \begin{pmatrix}\hat{y}_t & \hat{\pi}_t & \hat{p}_t\end{pmatrix}'$ and rewrite the system of equations (1), (2), (4), and (5) as:

$$\tilde{\mathbf{z}}_t = \alpha + \mathbf{B}E_t\tilde{\mathbf{z}}_{t+1} + \mathbf{C}\tilde{\mathbf{z}}_{t-1} + \chi^g\hat{g}_t + \chi^u\widehat{a}_t, \tag{25}$$

where $\alpha_{3\times3}$, $\mathbf{B}_{3\times3}$, $\mathbf{C}_{3\times3}$, $\chi^g_{3\times1}$ and $\chi^u_{3\times1}$ are parameter matrices that depend on whether the ELB binds or not. The general form of the MSV-REE solution then takes the expression:

$$\tilde{\mathbf{z}}_t = \mathbf{a} + \mathbf{\Theta}\tilde{\mathbf{z}}_{t-1} + \mathbf{c}\hat{g}_t + \mathbf{d}\hat{u}_t, \tag{26}$$

where the coefficient matrices $\mathbf{a}_{3\times1}$, $\mathbf{\Theta}_{3\times3}$,$\mathbf{c}_{3\times1}$ and $\mathbf{d}_{3\times1}$ also depend on whether the ELB binds or not. Applying the unconditional expectation operator to (26), rational expectations look like:

$$E_t\tilde{\mathbf{z}}_{t+1} = \mathbf{a} + \mathbf{\Theta}\tilde{\mathbf{z}}_t + \mathbf{c}\rho^g\hat{g}_t + \mathbf{d}\rho^u\hat{u}_t. \tag{27}$$

Inserting these into (25) yields the actual law of motion of the economy:

$$\tilde{\mathbf{z}}_t = (\mathbf{I} - \mathbf{B}\boldsymbol{\Theta})^{-1} \left[(\alpha + \mathbf{B}\mathbf{a}) + \mathbf{C}\tilde{\mathbf{z}}_{t-1} + (\mathbf{B}\mathbf{c}\rho^g + \chi^g)\,\hat{g}_t + (\mathbf{B}\mathbf{d}\rho^u + \chi^u)\,\hat{u}_t \right]. \tag{28}$$

Equalizing (26) with (28) yields:

$$\boldsymbol{\Theta} = (\mathbf{I} - \mathbf{B}\boldsymbol{\Theta})^{-1}\mathbf{C}, \mathbf{a} = (\mathbf{I} - \mathbf{B}\boldsymbol{\Theta} - \mathbf{B})^{-1}\alpha, \mathbf{c} = (\mathbf{I} - \mathbf{B}\boldsymbol{\Theta} - \mathbf{B}\rho^g)^{-1}\chi^g, \mathbf{d} = (\mathbf{I} - \mathbf{B}\boldsymbol{\Theta} - \mathbf{B}\rho^u)^{-1}\chi^u. \tag{29}$$

At the targeted REE (denoted by a superscript 'T'), the ELB does not bind and the MSV-REE and inserting the interest-rate rule into the IS curve (1) allows one to obtain α^T, $\mathbf{B}^T$, $\mathbf{C}^T$, χ^g and χ^u in (25). We find the MSV parameters at the target:

$$\boldsymbol{\Theta}^T = \left(\mathbf{I} - \mathbf{B}^T\boldsymbol{\Theta}^T\right)^{-1}\mathbf{C}^T, \quad \mathbf{a}^T = \left(\mathbf{I} - \mathbf{B}^T\boldsymbol{\Theta}^T - \mathbf{B}^T\right)^{-1}\alpha^T, \quad \mathbf{c}^T = \left(\mathbf{I} - \mathbf{B}^T\boldsymbol{\Theta}^T - \mathbf{B}^T\rho^g\right)^{-1}\chi^{g^T}, \mathbf{d}^T = \left(\mathbf{I} - \mathbf{B}^T\boldsymbol{\Theta}^T - \mathbf{B}^T\rho^u\right)^{-1}\chi^{u^T}. \tag{30}$$

Notice that $\boldsymbol{\Theta} = (\mathbf{I} - \mathbf{B}\mathbf{e})^{-1}\mathbf{C}$ is a quadratic matrix equation. We solve for the stable solution $\boldsymbol{\Theta}^T$ and obtain unique expressions for $\mathbf{a}^T$, $\boldsymbol{\Theta}^T$, $\mathbf{c}^T$ and $\mathbf{d}^T$.

Let us now consider the case when the ELB binds, i.e. $\hat{\imath}_t = -\bar{r}$. Insert this interest rate into the IS curve (1) results in α, $\mathbf{B}$, $\mathbf{C}$, χ^g and χ^u in Equation (29) at the ZLB. We find the MSV solution at the ZLB $\mathbf{a}^{elb}$, $\boldsymbol{\Theta}^{elb}$, $\mathbf{c}^{elb}$ and $\mathbf{d}^{elb}$:

$$\boldsymbol{\Theta}^{elb} = \left(\mathbf{I} - \mathbf{B}^{elb}\boldsymbol{\Theta}^{elb}\right)^{-1}\mathbf{C}^{elb}, \quad \mathbf{a}^{elb} = \left(\mathbf{I} - \mathbf{B}^{elb}\boldsymbol{\Theta}^{elb} - \mathbf{B}^{elb}\right)^{-1}\alpha^{elb}, \quad \mathbf{c}^{elb} = \left(\mathbf{I} - \mathbf{B}^{elb}\boldsymbol{\Theta}^{elb} - \mathbf{B}^{elb}\rho^g\right)^{-1}\chi^{g^{elb}}, \mathbf{d}^{elb} = \left(\mathbf{I} - \mathbf{B}^{elb}\boldsymbol{\Theta}^{elb} - \mathbf{B}^{elb}\rho^u\right)^{-1}\chi^{u^{elb}}. \tag{31}$$

Taken together, the MSV-REE under PLT is a piece-wise solution:

$$\tilde{\mathbf{z}}_t = \begin{pmatrix} \hat{y}_t & \hat{\pi}_t & \hat{p}_t \end{pmatrix}' = \begin{cases} \mathbf{a}^T + \boldsymbol{\Theta}^T\tilde{\mathbf{z}}_{t-1} + \mathbf{c}^T\hat{g}_t + \mathbf{d}^T\hat{u}_t & \text{, if } \hat{\imath}_t > -\bar{r}, \\ \mathbf{a}^{elb} + \mathbf{c}^{elb}\hat{g}_t + \mathbf{d}^{elb}\hat{u}_t & \text{, if } \hat{\imath}_t = -\bar{r}, \end{cases} \tag{32}$$

where we have: $\mathbf{A} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} \frac{1}{\Omega} & \frac{\sigma^{-1} - \beta\sigma^{-1}\phi^P}{\Omega} & 0 \\ \frac{\kappa}{\Omega} & \frac{\kappa\sigma^{-1} + \beta(1+\sigma^{-1}\phi^y)}{\Omega} & 0 \\ \frac{\kappa}{\Omega} & \frac{\kappa\sigma^{-1} + \beta(1+\sigma^{-1}\phi^y)}{\Omega} & 0 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 0 & 0 & \frac{\sigma^{-1}\phi^P}{\Omega} \\ 0 & 0 & \frac{-\kappa\sigma^{-1}\phi^P}{\Omega} \\ 0 & 0 & \frac{1+\sigma^{-1}\phi^y}{\Omega} \end{bmatrix}$,
$\chi = \begin{bmatrix} \frac{1}{\Omega} & \frac{-1}{\Omega} \\ \frac{\kappa}{\Omega} & \frac{1+\sigma^{-1}\phi^y}{\Omega} \\ \frac{\kappa}{\Omega} & \frac{1+\sigma^{-1}\phi^y}{\Omega} \end{bmatrix}$, $\Omega = 1 + \sigma^{-1}\phi^y + \kappa\sigma^{-1}\phi^{\pi,P}$;

and: $\mathbf{A}^{ZLB} = \begin{bmatrix} \sigma^{-1}\bar{r} \\ \kappa\sigma^{-1}\bar{r} \\ \kappa\sigma^{-1}\bar{r} \end{bmatrix}$, $\mathbf{B}^{ZLB} = \begin{bmatrix} 1 & \sigma^{-1} & 0 \\ \kappa & \beta+\kappa\sigma^{-1} & 0 \\ \kappa & \beta+\kappa\sigma^{-1} & 0 \end{bmatrix}$, $\mathbf{C}^{ZLB} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\chi^{ZLB} = \begin{bmatrix} 1 & 0 \\ \kappa & 1 \\ \kappa & 1 \end{bmatrix}$.

B.2 IT

Using a similar approach, we obtain the REE under IT. The reduced form of the three NK equations reads as:

$$\mathbf{z}_t = \begin{pmatrix} \hat{y}_t & \hat{\pi}_t \end{pmatrix}' = \alpha + \mathbf{B}E_t\mathbf{z}_{t+1} + \chi^g\hat{g}_t + \chi^u\hat{u}_t. \tag{33}$$

The MSV-REE under IT is:

$$\mathbf{z}_t = \begin{pmatrix} \hat{y}_t & \hat{\pi}_t \end{pmatrix}' = \begin{cases} \mathbf{a}^T + \mathbf{c}^T\hat{g}_t + \mathbf{d}^T\hat{u}_t & \text{, if } \hat{\imath}_t > 0, \\ \mathbf{a}^{elb} + \mathbf{c}^{elb}\hat{g}_t + \mathbf{d}^{elb}\hat{u}_t & \text{, if } \hat{\imath}_t = 0, \end{cases} \tag{34}$$

where the MSV matrices have the following representation:

$$\mathbf{a} = (\mathbf{I}-\mathbf{B})^{-1}\alpha, \mathbf{c} = (\mathbf{I}-\mathbf{B}\rho^g)^{-1}\chi^g, \text{ and } \mathbf{d} = (\mathbf{I}-\mathbf{B}\rho^u)^{-1}\chi^u, \tag{35}$$

that depends on the corresponding steady state.

At non-ZLB binding for IT, we have: $\mathbf{A} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} \frac{1}{\Omega} & \frac{\sigma^{-1}-\beta\sigma^{-1}\phi^\pi}{\Omega} \\ \frac{\kappa}{\Omega} & \frac{\kappa\sigma^{-1}+\beta(1+\sigma^{-1}\phi^y)}{\Omega} \end{bmatrix}$, $\chi = \begin{bmatrix} \frac{1}{\Omega} & \frac{-1}{\Omega} \\ \frac{\kappa}{\Omega} & \frac{1+\sigma^{-1}\phi^y}{\Omega} \end{bmatrix}$.

At binding-ZLB for IT, we have: $\mathbf{A}^{ZLB} = \begin{bmatrix} \sigma^{-1}\bar{r} \\ \kappa\sigma^{-1}\bar{r} \end{bmatrix}$, $\mathbf{B}^{ZLB} = \begin{bmatrix} 1 & \sigma^{-1} \\ \kappa & \beta+\kappa\sigma^{-1} \end{bmatrix}$, $\chi^{ZLB} = \begin{bmatrix} 1 & 0 \\ \kappa & 1 \end{bmatrix}$.

B.3 Proof of remark 2

We first show the determinacy of the non-ZLB equilibrium under PLT. The system of four equations of the PLT-economy is as follows:

$$\begin{aligned}
\hat{\pi}_t &= \beta \tilde{E}_t \hat{\pi}_{t+1} + \kappa \hat{y}_t + \hat{u}_t, \\
\hat{y}_t &= \tilde{E}_t \hat{y}_{t+1} - \sigma^{-1} \left(\hat{\imath}_t - \tilde{E}_t \hat{\pi}_{t+1} - \bar{r} \right) + \hat{g}_t, \\
\hat{p}_t &= \hat{p}_{t-1} + \hat{\pi}_t, \\
\hat{\imath}_t &= \bar{r} + \phi^P \hat{p}_t + \phi^y \hat{y}_t.
\end{aligned}$$

Abstracting from the shocks, removing the hat and expectation operator for brevity, and letting $x_t \equiv p_{t-1}$ results in:

$$\begin{bmatrix} 1 & -\kappa & 0 & 0 \\ 0 & 1+\sigma^{-1}\phi^y & \sigma^{-1}\phi^P & 0 \\ 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \\ p_t \\ x_t \end{bmatrix} = \begin{bmatrix} \beta & 0 & 0 & 0 \\ \sigma^{-1} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{t+1} \\ \pi_{t+1} \\ p_{t+1} \\ x_{t+1} \end{bmatrix}, \tag{36}$$

which can be rewritten as:

$$\begin{bmatrix} y_t \\ \pi_t \\ p_t \\ x_t \end{bmatrix} = \dot{\mathbf{B}} \begin{bmatrix} y_{t+1} \\ \pi_{t+1} \\ p_{t+1} \\ x_{t+1} \end{bmatrix}, \tag{37}$$

for $\dot{\mathbf{B}} = \begin{bmatrix} \beta + \frac{\kappa}{\phi^y+\sigma} & \frac{\kappa\sigma}{\phi^y+\sigma} & 0 & \frac{-\kappa\phi^P}{\phi^y+\sigma} \\ \frac{1}{\phi^y+\sigma} & \frac{\sigma}{\phi^y+\sigma} & 0 & \frac{-\phi^P}{\phi^y+\sigma} \\ 0 & 0 & 0 & 1 \\ -\beta - \frac{\kappa}{\phi^y+\sigma} & \frac{-\kappa\sigma}{\phi^y+\sigma} & 0 & \frac{\phi^y+\sigma+\kappa\phi^P}{\phi^y+\sigma} \end{bmatrix}.$

Dropping the definition equation, $\dot{\mathbf{B}}$ becomes as 3×3 matrix. The eigenvalues λ of $\dot{\mathbf{B}}$ are the solutions of $\det(\dot{\mathbf{B}} - \lambda \mathbf{I}) = 0$, where $\mathbf{I}$ is a 3-by-3 identity matrix. The targeted REE is locally determinate if and only if $\dot{\mathbf{B}}$ has two eigenvalues inside the unit circle, which holds if $\phi^P > 0$ and $\phi^y \geq 0$. This is consistent with the counterpart proof under a non-linearized NK model with a PLT rule in Honkapohja & Mitra [2020, Online Appendix C].

Similarly, we examine the stability property of the ELB-REE under PLT. When the ELB binds, the system reduces to two dimensions:

$$\begin{bmatrix} 1 & -\kappa \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \beta & 0 \\ \sigma^{-1} & 1 \end{bmatrix} \begin{bmatrix} y_{t+1} \\ \pi_{t+1} \end{bmatrix}, \tag{38}$$

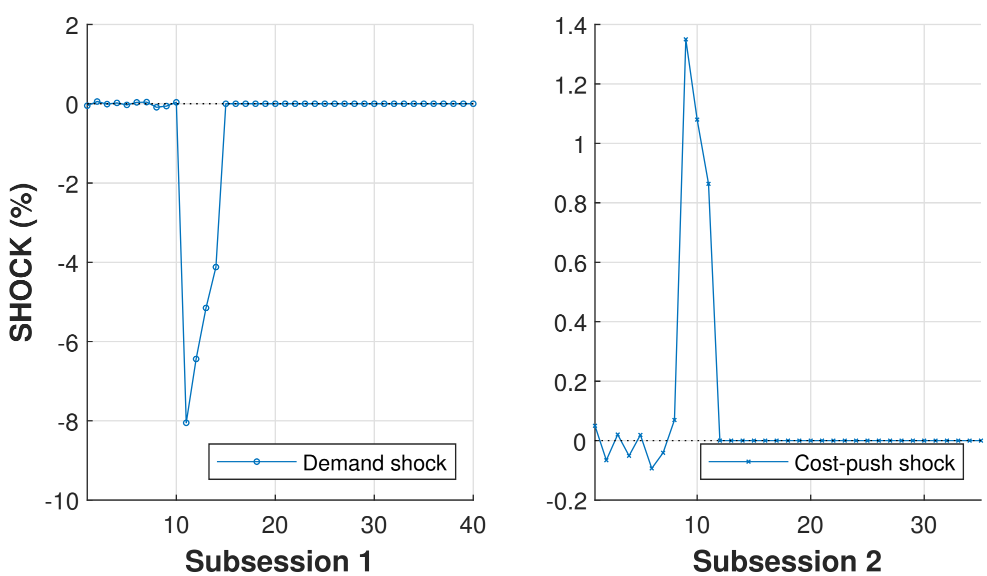
or, equivalently:

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \beta + \sigma^{-1}\kappa & \kappa \\ \sigma^{-1} & 1 \end{bmatrix} \begin{bmatrix} y_{t+1} \\ \pi_{t+1} \end{bmatrix},$$

of which the remaining 2-by-2 matrix has both eigenvalues within the unit circle iif $\beta < 1$ and $\sigma^{-1}\kappa < 0$, which is violated given that $\sigma, \kappa > 0$. Therefore, the ZLB-binding REE is indeterminate.

C The exogenous shock process

Figure 12: Shocks sequence used in all sessions



D Additional results

Table 4: Panel models of the subjects' forecasts in each treatment

	Tr. `IT`		*Tr.* `PLT`		*Tr.* `PLT-G`	
Dependent variable	$\pi^e_{j,t+1}$	$y^e_{j,t+1}$	$\pi^e_{j,t+1}$	$y^e_{j,t+1}$	$\pi^e_{j,t+1}$	$y^e_{j,t+1}$
y^e_{t-1}	0.01	0.29***	−0.17	0.19	0.12***	0.62***
	(0.01)	(0.02)	(0.23)	(0.23)	(0.01)	(0.02)
π^e_{t-1}	0.52***	0.05**	−0.44	−0.64*	0.18***	−0.18***
	(0.02)	(0.02)	(0.34)	(0.34)	(0.02)	(0.03)
y_{t-1}	−0.13***	0.93***	−0.15	0.68	−0.07**	0.65***
	(0.03)	(0.04)	(0.49)	(0.49)	(0.03)	(0.04)
y_{t-2}	0.10***	−0.27***	0.12	0.03	−0.05	−0.23***
	(0.04)	(0.05)	(0.69)	(0.69)	(0.04)	(0.06)
y_{t-3}	−0.01	−0.11***	−0.02	−0.26	−0.01	−0.15***
	(0.02)	(0.03)	(0.46)	(0.46)	(0.02)	(0.03)
π_{t-1}	0.63***	−0.29***	1.94**	0.28	0.70***	0.10
	(0.08)	(0.11)	(0.98)	(0.99)	(0.07)	(0.11)
π_{t-2}	−0.64***	0.68***	−1.59	−0.88	−0.39***	−0.22*
	(0.08)	(0.11)	(1.31)	(1.31)	(0.08)	(0.12)
π_{t-3}	0.02	−0.23***	0.68	0.75	0.02	0.30***
	(0.05)	(0.07)	(0.75)	(0.75)	(0.06)	(0.08)
$\pi^e_{BF,t}$	0.33***	−0.06	0.60*	0.43	0.37***	−0.10
	(0.04)	(0.06)	(0.36)	(0.36)	(0.05)	(0.08)
Shocks	0.08***	0.12***	−0.62	−0.20	0.04**	0.16***
	(0.02)	(0.02)	(0.77)	(0.76)	(0.02)	(0.03)
p_{t-1}	0.0002	−0.0001	−0.01	−0.001	−0.06	−0.09
	(0.0004)	(0.0005)	(0.02)	(0.02)	(0.04)	(0.06)
i_{t-1}	0.04	−0.22***	−0.29	−0.17	−0.03	−0.18***
	(0.03)	(0.05)	(0.40)	(0.40)	(0.03)	(0.05)
CB guidance					−0.30*	−0.74***
					(0.15)	(0.23)
Constant	0.67	−0.05	−1.22	−1.81	3.69	5.08
			(3.41)	(3.40)		(30.15)
Nb. Obs.	2,484	2,484	2,484	2,484	2,484	2,484
R^2	0.87	0.68	0.06	0.07	0.54	0.87

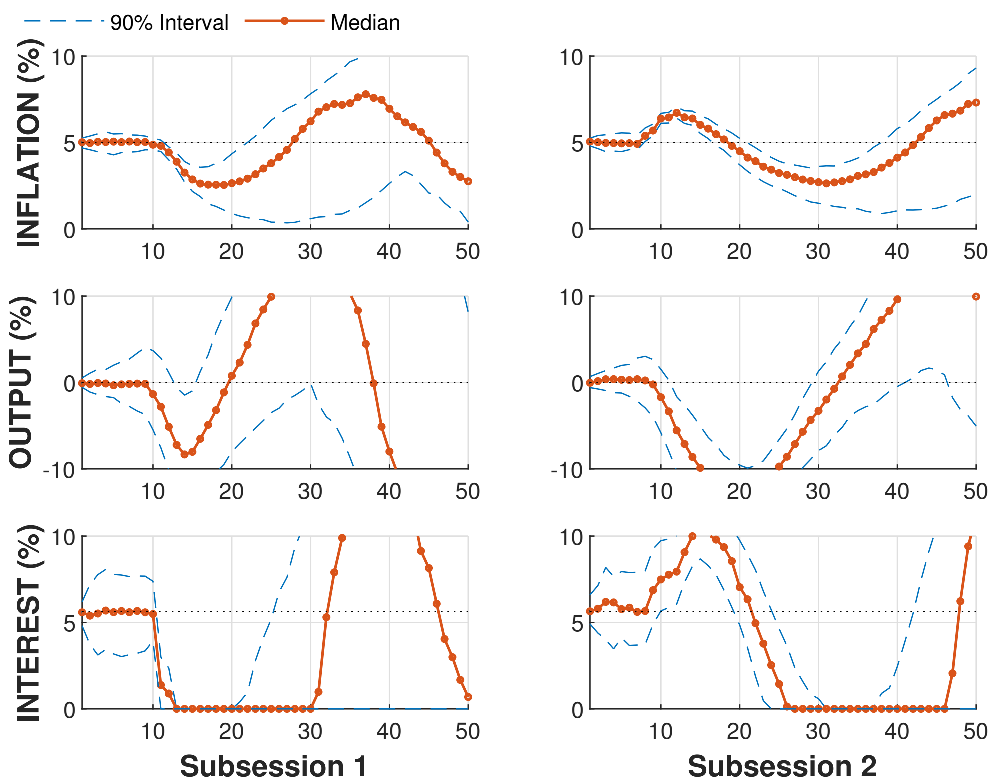
Note: Panel regression models including session and subsession fixed effects ($N = 36$ subjects $\times T = 69$ periods); robust clustered standard errors between brackets. The Wu-Hausman test fails to reject the null hypothesis in all case (the minimum p-value across the six models is 0.19). Varible $\pi^e_{BF,t}$ refers to the best inflation forecast in the group observed at the beginning of period t (from $t-1$). The variable shocks include the demand shocks in Subsession 1 and the cost-push shocks in Subsession 2. *p<0.1; **p<0.05; ***p<0.01.

Table 5: Cross-treatment comparison of forecasting performances in the lab

		Numerical values			*Associated p-values*		
		IT	PLT	PLT-G	IT *vs* PLT	IT *vs* PLT-G	PLT *vs* PLT-G
Total prediction scores							
π	Mean	2537	1555	2672	0.00	0.20	0.00
	SD	(462)	(455)	(397)			
y	Mean	2582	904	2085	0.00	0.00	0.00
	SD	(362)	(314)	(415)			

Note: See Table 3. Payoff in experimental points, higher values indicate a lower forecast errors and higher earnings in the experiment.

Figure 13: Inflation, output gap and interest rate dynamics in Tr. PLT in the HENK model over 50 periods for each subsession



Note: This simulation extends to 50 periods per subsession to illustrate more clearly the persistent fluctuations observed under PLT in the lab.

E Instructions booklet and GUI

E.1 Instructions

E.1.1 IT

Welcome to the experiment! You and others are participating in a simulated economic environment as **a group of professional forecasters at a statistical bureau** who earn money by **providing predictions of output and inflation**.

In this experiment, you do not need to have prior experience in performing such tasks nor having a sound economic background. The instructions will provide a description of your task and sufficient information to help you earn a considerable amount of money that will be paid shortly after the experiment. If you have any questions at any point, we will be glad to clarify them. Each participant is paid $7 for attending. Throughout this experiment, you will also earn points based on the decisions you make.

During the experiment, you are not allowed to use your mobile phone. You are also not allowed to communicate with other participants. If you have a question, please raise your hand and someone will come to your station.

General information about the experimental economy

The experiment will consist of two sub-sessions. This section describes the experimental economy of the first sub-session, which lasts for 40 periods. In each period you will be asked to **predict output and inflation** for the next period. You can think the current period is today and the next period is tomorrow. Every period, the statistical bureaus then take all the predictions made by you and other participants; then average them to get the **bureaus prediction of output and inflation**.

These predictions are useful for households and firms in the economy to guide their decisions.

The economy you are participating in is described by three variables: output, inflation, and interest rate. All variables are expressed in **percentage points**; for instance, 4%, -2.5%, 0.25%, 8% etc.

OUTPUT is the amount of goods and services produced by firms and consumed by consumers. In this simulated economy, we consider a simple measurement for output called output gap. The output gap is simply the percentage difference of GDP from the best level of GDP. For instance, if the output gap is +1%, the GDP level is 1% higher than the best level. If the output gap is -1%, the GDP level is 1% lower than the best level. And if the output gap is 0%, the GDP level is at the best level. The intended output gap in this economy is therefore 0%.

- Todays output gap has a **positive** relationship with the **bureaus prediction of tomorrows output gap**. That is, if the bureaus prediction of tomorrows output gap increases, todays output gap also increases; and vice versa.
- Todays output gap has a **positive** relationship with the **bureaus prediction of tomorrows inflation**.
- Todays output gap has a **negative** relationship with the **interest rate** set by the central bank, and vice versa. That is if the interest rate increases, todays output gap decreases; and vice versa.
- Todays output gap has a **positive** relationship with **random demand shock**. When shocks are **small** (between -0.5% to 0.5%), you can expect shocks in each period to be independent and random (i.e., shock today does not determine shock tomorrow). When a shock is **large**, it will determine the following shocks and you can expect the following shocks to have the same sign but a smaller magnitude (i.e., a large shock will dissipate over a few periods).

$$\text{output today} = \begin{array}{c}(+)\text{ output prediction}\\(+)\text{ inflation prediction}\\(-)\text{ interest rate}\\(+)\text{ shock}\end{array}$$

INFLATION is the percentage change in the average price of goods and services between two consecutive periods. For instance, if inflation is 5% today, it means that the price level **increased** by 5% compared to yesterday. If inflation is -5% today, it means that the price level **decreased** by 5% compared to yesterday.

The **intended level of inflation** is 5% in this economy. This means that the price should grow at a rate of 5% per period.

- Todays inflation has a positive relationship with the **bureaus prediction of tomorrows inflation**.
- Todays inflation has a positive relationship with **todays output gap**.

$$\text{inflation today} = \begin{array}{c} (+)\ \text{output today} \\ (+)\ \text{inflation prediction} \end{array}$$

- Because a rise in the interest rate decreases output today, there is thus a negative relationship between interest rate and inflation: a **rise in interest rate decreases inflation and vice-versa** (this is how the policy maker can act upon inflation by adjusting the interest rate).

INTEREST RATE is the cost of borrowing money (or the benefit of saving money). There is a policymaker called the **central bank**.

- The **objective** of the central bank is to keep the **inflation at 5% and the output gap at 0%** .
- The shocks may cause the output gap and inflation to deviate from their intended level. The CB increases or decreases the interest rate to achieve its objective.
- If the CB wants to increase the output gap or inflation it decreases the interest rate, if it wants to decrease the output gap or inflation it increases the interest rate. Note that the interest rate cannot go below zero.
- Note that due to shocks, the economy will fluctuate and not always at the central banks objective; however, the economy will be more stabilized under the central banks actions.

Your prediction tasks

Your task in each period of the experiment is to predict the output gap and inflation of the next period. When the experiment starts, you enter the start of period 1 and have to predict the output gap and inflation for period 2. Once all participants have submitted their two predictions, the average predictions for period 2 and the shock in period 1 determine the interest rate, inflation, and output gap in period 1. You then enter period 2 and have to submit output gap and inflation predictions for period 3. This process repeats itself for the length of the session.

Additionally, you will be asked one time for a long-run inflation prediction 11 periods before a sub-session ends. That is, for the first sub-session; in period 29 you will be asked: What do you think is the long-run inflation in this economy. You can think the long-run inflation as the inflation at the last period of a sub-session.

Note: your prediction can be positive, negative, or zero, and decimal. For example, if you want to submit a prediction of 0.53% type 0.53; for a prediction of -1.75% type -1.75.

Important remarks

- In every period, the bureau averages your and others predictions to compute **the bureaus predictions**.

- **The predictions of the bureau** have an important **influence on the output gap and inflation today**.
- The central banks objective is to keep the **inflation at 5% and the output gap at 0%**.
- To **help** you with your output gap and inflation predictions, the bureau will reveal at the end of each period, **other forecasters latest output gap and inflation predictions and prediction scores** (see below how the scores are computed). This information will be anonymous but will allow you to see who the most accurate forecasters in the bureau are, on average so far in the experiment and in the last period, and what their predictions are.

Your payment

Your payment will depend on the **accuracy** of your predictions, measured as the **absolute distance** (error) between your **predictions** and the **actual values**. For each period, the prediction errors are calculated as soon as the actual values are known, that is at the end of the next period. For instance, in period 6, you predict output gap and inflation for period 7, so your prediction errors and payment for period 6 will be computed at the end of period 7. You will be paid based on your **individual** performance, not the group performance.

Your prediction score decreases as the prediction error increases. For any error, you make 100/(1+error) points. Hence, in the case of perfect prediction (zero prediction error), you get a maximum of 100 points in each period.

Example: If you predict 6 percent inflation and the actual inflation turns out to be 5 percent, your prediction error is $|6 - 5| = 1$. *If you predict 4.5 percent and the actual inflation turns out 4.9 percent, your error is* $|4.5 - 4.9| = 0.4$.

After the experiment, you will have **two total scores**, one for inflation predictions and one for output predictions. These total scores simply consist of the sum of all prediction scores you got during the experiment, separately for output gap and inflation predictions. You will be paid either for predicting inflation or for output prediction. One of the two scores will be randomly selected with equal probability for payment at a rate of **20 cents per 100 points**. The selected score is the same for all participants and the random draw is independent of how well you did in the two tasks.

For the one-time long-run inflation prediction, the payment rate is **3 dollars per 100 points**, which is 15 times higher than the short-run prediction. The score for this long-run prediction is evaluated by its absolute distance to the **last period**s inflation of the sub-session.

Computer interface and information

The computer interface is mainly self-explanatory. At the beginning of every period, prior to **submitting your predictions**. You are provided with the following information.

- The current shock of today.
- All past inflation and your predictions up until yesterday.
- All past outputs and your predictions up until yesterday.
- All past interest rates up until yesterday.
- All past prices and the price if inflation always grows at 5%.
- Your prediction scores of the last period and cumulative scores so far in the experiment.

Additionally, you also see

- The last period output gap and inflation predictions of your peers, and their corresponding forecasting scores in the last period only and so far in the experiment.

All this information may be relevant to form your predictions, but this is up to you to make use of it or not.

There will be boxes for you to enter your predictions. When submitting your predictions, use a decimal point if necessary (not a comma). For example, if you want to submit a prediction of 0.5% type 0.5; for a prediction of -1.75% type -1.75.

[the following information is provided to subjects after the first subsession concluded]

This page describes the experimental economy of the second sub-session, which lasts for 35 periods.

In this sub-session, there will be a change in the nature of the random shocks. Shocks are no longer affecting the output (therefore removed from the OUTPUT box). Shocks are now affecting inflation as now highlighted in the INFLATION box. The rest will be identical to the first sub-session.

OUTPUT is **no longer** affected by the shocks.

$$\text{output today} = \begin{array}{c} (+)\text{ output prediction} \\ (+)\text{ inflation prediction} \\ (-)\text{ interest rate} \end{array}$$

INFLATION todays inflation now has a **positive** relationship with **random cost-push shock**. When shocks are small (between -0.5% to 0.5%), you can expect shock in each period to be independent and random (i.e., shock today does not determine shock tomorrow). When a shock is large, it will determine the following shocks and you can expect the following shocks to have the same sign but a smaller magnitude (i.e., a large shock will dissipate over a few periods).

$$\text{inflation today} = \begin{array}{c} (+)\text{ output today} \\ (+)\text{ inflation prediction} \\ (+)\text{shock} \end{array}$$

E.1.2 PLT

Welcome to the experiment! You and others are participating in a simulated economic environment as **a group of professional forecasters at a statistical bureau** who earn money by **providing predictions of output and inflation**.

In this experiment, you do not need to have prior experience in performing such tasks nor having a sound economic background. The instructions will provide a description of your task and sufficient information to help you earn a considerable amount of money that will be paid shortly after the experiment. If you have any questions at any point, we will be glad to clarify them. Each participant is paid $7 for attending. Throughout this experiment, you will also earn points based on the decisions you make.

During the experiment, you are not allowed to use your mobile phone. You are also not allowed to communicate with other participants. If you have a question, please raise your hand and someone will come to your station.

General information about the experimental economy

The experiment will consist of two sub-sessions. This section describes the experimental economy of the first sub-session, which lasts for 40 periods. In each period you will be asked to **predict output and inflation** for the next period. You can think the current period is today and the next period is tomorrow. Every period, the statistical bureaus then take all the predictions made by you and other participants; then average them to get the **bureaus prediction of output and inflation**.

These predictions are useful for households and firms in the economy to guide their decisions.

The economy you are participating in is described by four variables: output, inflation, price gap, and interest rate. All variables are expressed in **percentage points**; for instance, 4%, -2.5%, 0.25%, 8% etc.

OUTPUT is the amount of goods and services produced by firms and consumed by consumers. In this simulated economy, we consider a simple measurement for output called output gap. The output gap is simply the percentage difference of GDP from the best level of GDP. For instance, if the output gap is +1%, the GDP level is 1% higher than the best level. If the output gap is -1%, the GDP level is 1% lower than the best level. And if the output gap is 0%, the GDP level is at the best level. The intended output gap in this economy is therefore 0%.

- Todays output gap has a **positive** relationship with the **bureaus prediction of tomorrows output gap**. That is, if the bureaus prediction of tomorrows output gap increases, todays output gap also increases; and vice versa.
- Todays output gap has a **positive** relationship with the **bureaus prediction of tomorrows inflation**.
- Todays output gap has a **negative** relationship with the **interest rate** set by the central bank, and vice versa. That is if the interest rate increases, todays output gap decreases; and vice versa.
- Todays output gap has a **positive** relationship with **random demand shock**. When shocks are **small** (between -0.5% to 0.5%), you can expect shocks in each period to be independent and random (i.e., shock today does not determine shock tomorrow). When a shock is **large**, it will determine the following shocks and you can expect the following shocks to have the same sign but a smaller magnitude (i.e., a large shock will dissipate over a few periods).

$$\text{output today} = \begin{array}{c} (+)\text{ output prediction} \\ (+)\text{ inflation prediction} \\ (-)\text{ interest rate} \\ (+)\text{ shock} \end{array}$$

INFLATION is the percentage change in the average price of goods and services between two consecutive periods. For instance, if inflation is 5% today, it means that the price level **increased** by 5% compared to yesterday. If inflation is -5% today, it means that the price level **decreased** by 5% compared to yesterday.

The intended level of inflation is 5% in this economy. This means that the price should grow at a rate of 5% per period. The initial price is 100. In period 1, the intended price is therefore 100*(1+0.05) = 105. In period 2, the intended price is 105*(1+0.05) = 110.25, then 115.8, 121,

- Todays inflation has a positive relationship with the **bureaus prediction of tomorrows inflation**.
- Todays inflation has a positive relationship with **todays output gap**.

$$\text{inflation today} = \begin{array}{c} (+)\text{ output today} \\ (+)\text{ inflation prediction} \end{array}$$

- Because a rise in the interest rate decreases output today, there is thus a negative relationship between interest rate and inflation: a **rise in interest rate decreases inflation and vice-versa** (this is how the policy maker can act upon inflation by adjusting the interest rate).

PRICE GAP is the difference between the actual price and the intended price. Under intended inflation of 5%, the intended price sequence is 105, 110.25, 115.8, 121,... For example, in period 2, suppose inflation is only 4% and then the price is 109.02, given that the intended price for this period is 110.25, we then have a negative price gap being (109.02-110.25) = -1.23. If inflation is 6% the price is 111.3, and we have a positive price gap being (111.3-110.25) = +1.05. And if inflation is 5% the price is 110.25, we then have a zero price gap, and zero price gap is the target.

INTEREST RATE is the cost of borrowing money (or the benefit of saving money). There is a policymaker called the **central bank**.

- The **objective** of the central bank is to keep the **inflation at 5% and the output gap at 0%** .
- The shocks may cause the output gap and inflation to deviate from their intended level. The CB increases or decreases the interest rate to achieve its objective.
- To keep the price gap at 0%, **the level of target inflation will fluctuate and not always be in line with the 5% intended level**. If the **price gap** is **negative** (**positive**), then inflation needs to be **higher** (**lower**) in the next period.
- If the CB wants to increase the output gap or inflation it decreases the interest rate, if it wants to decrease the output gap or inflation it increases the interest rate. Note that the interest rate cannot go below zero.
- Note that due to shocks, the economy will fluctuate and not always at the central banks objective; however, the economy will be more stabilized under the central banks actions.

Your prediction tasks

Your task in each period of the experiment is to predict the output gap and inflation of the next period. When the experiment starts, you enter the start of period 1 and have to predict the output gap and inflation for period 2. Once all participants have submitted their two predictions, the average predictions for period 2 and the shock in period 1 determine the interest rate, inflation, and output gap in period 1. You then enter period 2 and have to submit output gap and inflation predictions for period 3. This process repeats itself for the length of the session.

Additionally, you will be asked one time for a long-run inflation prediction 11 periods before a sub-session ends. That is, for the first sub-session; in period 29 you will be asked: What do you think is the long-run inflation in this economy. You can think the long-run inflation as the inflation at the last period of a sub-session.

Note: your prediction can be positive, negative, or zero, and decimal. For example, if you want to submit a prediction of 0.53% type 0.53; for a prediction of -1.75% type -1.75.

Important remarks

- In every period, the bureau averages your and others predictions to compute **the bureaus predictions**.
- **The predictions of the bureau** have an important **influence on the output gap and inflation today**.
- The central banks objective is to keep the **price gap at 5% and the output gap at 0%**.
- To keep the **price gap at 0%**, the required level of inflation may be different from the intended level of 5%.
- To **help** you with your output gap and inflation predictions, the bureau will reveal at the end of each period, **other forecasters latest output gap and inflation predictions and prediction scores** (see below how the scores are computed). This information will be anonymous but will allow you to see who the most accurate forecasters in the bureau are, on average so far in the experiment and in the last period, and what their predictions are.

Your payment

Your payment will depend on the **accuracy** of your predictions, measured as the **absolute distance** (error) between your **predictions** and the **actual values**. For each period, the prediction errors are calculated as soon as the actual values are known, that is at the end of the next period. For instance, in period 6, you predict output gap and inflation for period 7, so your prediction errors and payment for period 6 will be computed at the end of period 7. You will be paid based on your **individual** performance, not the group performance.

Your prediction score decreases as the prediction error increases. For any error, you make 100/(1+error) points. Hence, in the case of perfect prediction (zero prediction error), you get a maximum of 100 points in each period.

Example: If you predict 6 percent inflation and the actual inflation turns out to be 5 percent, your prediction error is $|6 - 5| = 1$. *If you predict 4.5 percent and the actual inflation turns out 4.9 percent, your error is* $|4.5 - 4.9| = 0.4$.

After the experiment, you will have **two total scores**, one for inflation predictions and one for output predictions. These total scores simply consist of the sum of all prediction scores you got during the experiment, separately for output gap and inflation predictions. You will be paid either for predicting inflation or for output prediction. One of the two scores will be randomly selected with equal probability for payment at a rate of **20 cents per 100 points**. The selected score is the same for all participants and the random draw is independent of how well you did in the two tasks.

For the one-time long-run inflation prediction, the payment rate is **3 dollars per 100 points**, which is 15 times higher than the short-run prediction. The score for this long-run prediction is evaluated by its absolute distance to the **last period**s inflation of the sub-session.

Computer interface and information

The computer interface is mainly self-explanatory. At the beginning of every period, prior to **submitting your predictions**. You are provided with the following information.

- The current shock of today.
- All past inflation and your predictions up until yesterday.
- All past outputs and your predictions up until yesterday.
- All past interest rates up until yesterday.
- All past prices and the price if inflation always grows at 5%.
- Your prediction scores of the last period and cumulative scores so far in the experiment.

Additionally, you also see

- The last period output gap and inflation predictions of your peers, and their corresponding forecasting scores in the last period only and so far in the experiment.

All this information may be relevant to form your predictions, but this is up to you to make use of it or not.

There will be boxes for you to enter your predictions. When submitting your predictions, use a decimal point if necessary (not a comma). For example, if you want to submit a prediction of 0.5% type 0.5; for a prediction of -1.75% type -1.75.

[the following cost-push shock description provided to subjects after the first subsession concluded is similar across treatments, see above instruction]

E.1.3 `PLT–G`

Welcome to the experiment! You and others are participating in a simulated economic environment as **a group of professional forecasters at a statistical bureau** who earn money by **providing predictions of output and inflation**.

In this experiment, you do not need to have prior experience in performing such tasks nor having a sound economic background. The instructions will provide a description of your task and sufficient information to help you earn a considerable amount of money that will be paid shortly after the experiment. If you have any questions at any point, we will be glad to clarify them. Each participant is paid $7 for attending. Throughout this experiment, you will also earn points based on the decisions you make.

During the experiment, you are not allowed to use your mobile phone. You are also not allowed to communicate with other participants. If you have a question, please raise your hand and someone will come to your station.

General information about the experimental economy

The experiment will consist of two sub-sessions. This section describes the experimental economy of the first sub-session, which lasts for 40 periods. In each period you will be asked to **predict output and inflation** for the next period. You can think the current period is today and the next period is tomorrow. Every period, the statistical bureaus then take all the predictions made by you and other participants; then average them to get the **bureaus prediction of output and inflation**.

These predictions are useful for households and firms in the economy to guide their decisions.

The economy you are participating in is described by four variables: output, inflation, price gap, and interest rate. All variables are expressed in **percentage points**; for instance, 4%, -2.5%, 0.25%, 8% etc.

OUTPUT is the amount of goods and services produced by firms and consumed by consumers. In this simulated economy, we consider a simple measurement for output called output gap. The output gap is simply the percentage difference of GDP from the best level of GDP. For instance, if the output gap is +1%, the GDP level is 1% higher than the best level. If the output gap is -1%, the GDP level is 1% lower than the best level. And if the output gap is 0%, the GDP level is at the best level. The intended output gap in this economy is therefore 0%.

- Todays output gap has a **positive** relationship with the **bureaus prediction of tomorrows output gap**. That is, if the bureaus prediction of tomorrows output gap increases, todays output gap also increases; and vice versa.
- Todays output gap has a **positive** relationship with the **bureaus prediction of tomorrows inflation**.
- Todays output gap has a **negative** relationship with the **interest rate** set by the central bank, and vice versa. That is if the interest rate increases, todays output gap decreases; and vice versa.
- Todays output gap has a **positive** relationship with **random demand shock**. When shocks are **small** (between -0.5% to 0.5%), you can expect shocks in each period to be independent and random (i.e., shock today does not determine shock tomorrow). When a shock is **large**, it will determine the following shocks and you can expect the following shocks to have the same sign but a smaller magnitude (i.e., a large shock will dissipate over a few periods).

$$\text{output today} = \begin{array}{c} (+)\text{ output prediction} \\ (+)\text{ inflation prediction} \\ (-)\text{ interest rate} \\ (+)\text{ shock} \end{array}$$

INFLATION is the percentage change in the average price of goods and services between two consecutive periods. For instance, if inflation is 5% today, it means that the price level **increased** by 5% compared to yesterday. If inflation is -5% today, it means that the price level **decreased** by 5% compared to yesterday.

The intended level of inflation is 5% in this economy. This means that the price should grow at a rate of 5% per period. The initial price is 100. In period 1, the intended price is therefore 100*(1+0.05) = 105. In period 2, the intended price is 105*(1+0.05) = 110.25, then 115.8, 121,

- Todays inflation has a positive relationship with the **bureaus prediction of tomorrows inflation**.
- Todays inflation has a positive relationship with **todays output gap**.

$$\text{inflation today} = \begin{array}{c}(+)\ \text{output today}\\(+)\ \text{inflation prediction}\end{array}$$

- Because a rise in the interest rate decreases output today, there is thus a negative relationship between interest rate and inflation: a **rise in interest rate decreases inflation and vice-versa** (this is how the policy maker can act upon inflation by adjusting the interest rate).
- **To help you with your inflation predictions, at the beginning of every period the central bank will publicly announce the target inflation rate to bring the price to its intended level.**

PRICE GAP is the difference between the actual price and the intended price. Under intended inflation of 5%, the intended price sequence is 105, 110.25, 115.8, 121,... For example, in period 2, suppose inflation is only 4% and then the price is 109.02, given that the intended price for this period is 110.25, we then have a negative price gap being (109.02-110.25) = -1.23. If inflation is 6% the price is 111.3, and we have a positive price gap being (111.3-110.25) = +1.05. And if inflation is 5% the price is 110.25, we then have a zero price gap, and zero price gap is the target.

INTEREST RATE is the cost of borrowing money (or the benefit of saving money). There is a policymaker called the **central bank**.

- The **objective** of the central bank is to keep the **inflation at 5% and the output gap at 0%** .
- The shocks may cause the output gap and inflation to deviate from their intended level. The CB increases or decreases the interest rate to achieve its objective.
- To keep the price gap at 0%, **the level of target inflation will fluctuate and not always be in line with the 5% intended level**. If the **price gap** is **negative** (**positive**), then inflation needs to be **higher** (**lower**) in the next period.
- If the CB wants to increase the output gap or inflation it decreases the interest rate, if it wants to decrease the output gap or inflation it increases the interest rate. Note that the interest rate cannot go below zero.
- **Again, o help you with your inflation predictions, at the beginning of every period the central bank will publicly announce the target inflation rate to bring the price to its intended level.**
- Note that due to shocks, the economy will fluctuate and not always at the central banks objective; however, the economy will be more stabilized under the central banks actions.

Your prediction tasks

Your task in each period of the experiment is to predict the output gap and inflation of the next period. When the experiment starts, you enter the start of period 1 and have to predict the output gap and inflation for period 2. Once all participants have submitted their two predictions, the average predictions for period 2 and the shock in period 1 determine the interest rate, inflation, and output gap in period 1. You then enter period 2 and have to submit output gap and inflation predictions for period 3. This process repeats itself for the length of the session.

Additionally, you will be asked one time for a long-run inflation prediction 11 periods before a sub-session ends. That is, for the first sub-session; in period 29 you will be asked: What do you think is the long-run inflation in this economy. You can think the long-run inflation as the inflation at the last period of a sub-session.

Note: your prediction can be positive, negative, or zero, and decimal. For example, if you want to submit a prediction of 0.53% type 0.53; for a prediction of -1.75% type -1.75.

Important remarks

- In every period, the bureau averages your and others predictions to compute **the bureaus predictions**.
- **The predictions of the bureau** have an important **influence on the output gap and inflation today**.
- The central banks objective is to keep the **price gap at 5% and the output gap at 0%**.
- To keep the **price gap at 0%**, the required level of inflation may be different from the intended level of 5%.
- To **help** you with your output gap and inflation predictions, the bureau will reveal at the end of each period, **other forecasters latest output gap and inflation predictions and prediction scores** (see below how the scores are computed). This information will be anonymous but will allow you to see who the most accurate forecasters in the bureau are, on average so far in the experiment and in the last period, and what their predictions are.

Your payment

Your payment will depend on the **accuracy** of your predictions, measured as the **absolute distance** (error) between your **predictions** and the **actual values**. For each period, the prediction errors are calculated as soon as the actual values are known, that is at the end of the next period. For instance, in period 6, you predict output gap and inflation for period 7, so your prediction errors and payment for period 6 will be computed at the end of period 7. You will be paid based on your **individual** performance, not the group performance.

Your prediction score decreases as the prediction error increases. For any error, you make 100/(1+error) points. Hence, in the case of perfect prediction (zero prediction error), you get a maximum of 100 points in each period.

Example: If you predict 6 percent inflation and the actual inflation turns out to be 5 percent, your prediction error is $|6 - 5| = 1$. *If you predict 4.5 percent and the actual inflation turns out 4.9 percent, your error is* $|4.5 - 4.9| = 0.4$.

After the experiment, you will have **two total scores**, one for inflation predictions and one for output predictions. These total scores simply consist of the sum of all prediction scores you got during the experiment, separately for output gap and inflation predictions. You will be paid either for predicting inflation or for output prediction. One of the two scores will be randomly selected with equal probability for payment at a rate of **20 cents per 100 points**. The selected score is the same for all participants and the random draw is independent of how well you did in the two tasks.

For the one-time long-run inflation prediction, the payment rate is **3 dollars per 100 points**, which is 15 times higher than the short-run prediction. The score for this long-run prediction is evaluated by its absolute distance to the **last period**s inflation of the sub-session.

Computer interface and information

The computer interface is mainly self-explanatory. At the beginning of every period, prior to **submitting your predictions**. You are provided with the following information.

- The current shock of today.
- All past inflation and your predictions up until yesterday.
- All past outputs and your predictions up until yesterday.
- All past interest rates up until yesterday.

- All past prices and the price if inflation always grows at 5%.
- The **target inflation** to bring the price gap at 0, announced by the central bank, i.e. the yellow sun in the inflation chart.
- Your prediction scores of the last period and cumulative scores so far in the experiment.

Additionally, you also see

- The last period output gap and inflation predictions of your peers, and their corresponding forecasting scores in the last period only and so far in the experiment.

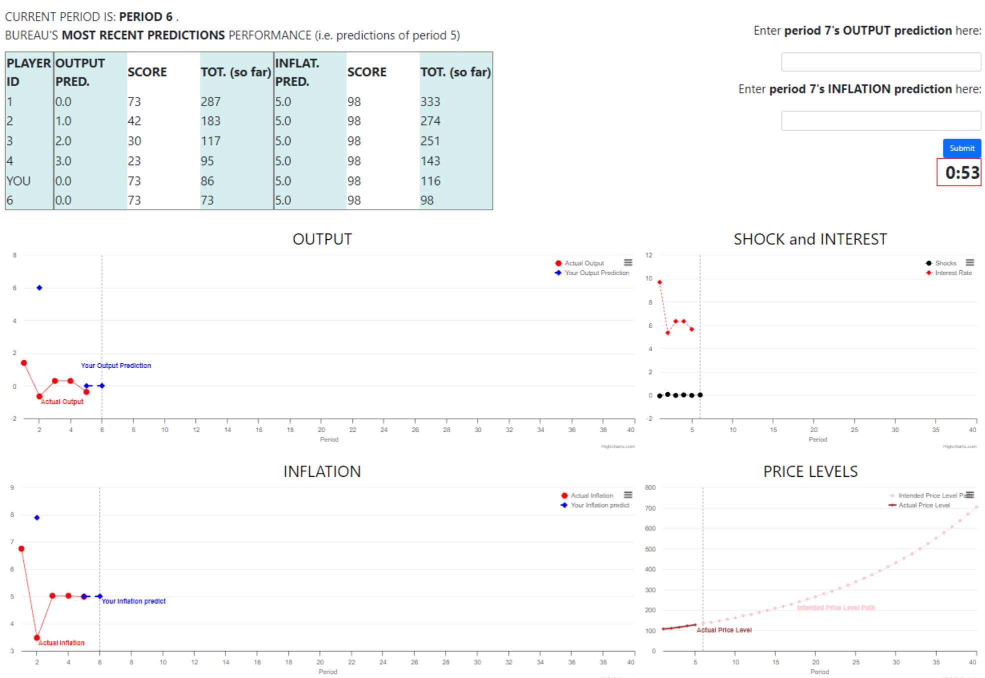
All this information may be relevant to form your predictions, but this is up to you to make use of it or not.

There will be boxes for you to enter your predictions. When submitting your predictions, use a decimal point if necessary (not a comma). For example, if you want to submit a prediction of 0.5% type 0.5; for a prediction of -1.75% type -1.75.

[the following cost-push shock description provided to subjects after the first subsession concluded is similar across treatments, see above instruction]

E.2 Examples of screens

Figure 14: GUI in `IT` and `PLT`



PLAYER ID	OUTPUT PRED.	SCORE	TOT. (so far)	INFLAT. PRED.	SCORE	TOT. (so far)
1	0.0	73	287	5.0	98	333
2	1.0	42	183	5.0	98	274
3	2.0	30	117	5.0	98	251
4	3.0	23	95	5.0	98	143
YOU	0.0	73	86	5.0	98	116
6	0.0	73	73	5.0	98	98

Figure 15: GUI in `PLT-G`