UvA-DARE (Digital Academic Repository)

Price level versus inflation targeting under heterogeneous expectations

A laboratory experiment

Hommel, C.; Makarewicz, T.

DOI

[10.1016/j.jebo.2020.11.011](https://doi.org/10.1016/j.jebo.2020.11.011)

Publication date

2021

Document Version

Final published version

Published in

Journal of Economic Behavior and Organization

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

Hommel, C., & Makarewicz, T. (2021). Price level versus inflation targeting under heterogeneous expectations: A laboratory experiment. *Journal of Economic Behavior and Organization*, 182, 39-82. <https://doi.org/10.1016/j.jebo.2020.11.011>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Price level versus inflation targeting under heterogeneous expectations: a laboratory experiment

Cars Hommes^a, Tomasz Makarewicz^{b,1,*}

^a University of Amsterdam, Tinbergen Institute and Bank of Canada, the Netherlands

^b Otto-Friedrich-Universität Bamberg and Universität Bielefeld, Germany

ARTICLE INFO

Article history:

Received 22 August 2019

Revised 9 November 2020

Accepted 10 November 2020

Available online 24 December 2020

JEL classification:

C92

D83

D84

E52

E62

Keywords:

Experimental macroeconomics

Expectations and learning

Monetary policy

ABSTRACT

Since the 2007 crisis, macroeconomists have been interested in monetary policies that could help with stabilizing inflation and output (Honkapohja, 2015). Two ideas gained particular attention: (i) that inflation should be replaced by the nominal price level (PLT) as the target for the central bank; and (ii) that the central bank should provide explicit guidance about its interest rate rule. We conduct a laboratory experiment to test the validity of these two hypotheses.

Our experiment uses a Learning-to-Forecast design based on a simple DSGE economy. Subjects are given a qualitative description of the economy and are asked to predict inflation and output gap two-periods ahead for 50 periods. There are five treatments. Baseline treatment (1) incorporates a standard inflation targeting rule. The other four treatments utilize a PLT Taylor rule and are based on a two-by-two design: a 'weak' PLT rule (2) with guidance and (3) without guidance; and a 'strong' PLT rule (4) with guidance and (5) without guidance. By guidance we mean that the central bank informs subjects about the price level deviation from its target.

We find that subjects within each treatment coordinate on similar behavior, but large differences between the treatments prevail. Guidance has a negligible effect, whereas a weak or strong Taylor rule specification turns out to be crucial for stability. PLT can be a robust monetary policy, but only if it is sufficiently responsive to the deviations of output and prices.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Until the second half of the 2000's, most economists and policy makers believed that the monetary policy, embodied by the well known Taylor rule, reached mature and satisfying level. Central banks began applying inflation target based interest rate rules in the 1980s, and this coincided with a period known as the Great Moderation, a quarter of century of relatively small inflation and output volatility in developed countries. Many empirical studies would in fact claim to identify

* Corresponding author.

E-mail address: tomasz.makarewicz@uni-bamberg.de (T. Makarewicz).

¹ The authors are thankful to Seppo Honkapohja, whose presentation during a MACFINROBODS workshop in London, January 2015, inspired this paper. We gratefully acknowledge that the research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement Integrated Macro-Financial Modeling for Robust Policy Design (MACFINROBODS, grant no. 612796). The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Canada.

the monetary policy as one of the most important factors leading to the Great Moderation, as it stabilized the propagation mechanisms of economic shocks [the so called ‘policy hypothesis’, see [14,21,20], for an overview].

Once the 2008 crisis erupted, central banks accordingly cut their interest rates, but, as soon became apparent, this response proved too slow to counter the unfolding recession. What made matters worse was that, despite the interest rates of many central banks approaching the zero lower bound, the recession remained rampant, leaving little space for conventional policies [12]. For example, the FED kept its interest rate at 0.25% level from December 2008 until December 2015, and then below 1% until March 2017. Nevertheless, the GDP growth in the US through these years was close to only 2%, and the employment has recovered to the pre-crisis level only by the end of 2016. An even more striking example comes with the events in Europe. The ECB also cut interest rates (though not as aggressively as the FED), but after a brief recovery in 2010 and 2011, the Euro Area experienced another year of negative growth. In addition, this output contraction hit some countries particularly strong, often causing major political disturbances, as in the case of Greece or Spain. See also Reifschneider and Roberts [34] for a study of a similar crisis in Japan.

The events of the crisis inspired a wave of macroeconomic literature on robust monetary policy [11,36,29,13,37]. The question is how to pull the economies from the current recession [10,38], but also how to prevent such big recessions in the future. Two policy alternatives gained particular attention: guidance and nominal price-level targeting (PLT) [27].

In this paper we report a macroeconomic laboratory experiment, in which we evaluate these two alternative monetary policies: guidance and PLT. Subjects were tasked with forecasting inflation and output gap in a simple economy, which was based on a standard Dynamic Stochastic General Equilibrium model (DSGE) with zero lower bound (ZLB) on the interest rate set by the central bank.

The usual definition of PLT is that the central bank pursues a Taylor rule, trying to steer the inflation such that the whole price level trajectory follows a preplanned path with a fixed trend [22]. Guidance in its broadest definition means that the central bank tries to manage inflation expectations by a public commitment to a certain policy [30]. For instance, during a period of recession, the central bank could acknowledge to keep its interest rate at a low value for a longer period; or it could explicate conditions under which it would raise the interest rate again.

The underlying idea of both the guidance and the PLT policies is that they give economic agents more information about the monetary policy, see Covas and Zhang [16] for a discussion about the PLT and Femia et al. [19] for an empirical evaluation of FED’s guidance. Under rational expectations (RE), this allows to curb the business cycle, and speeds up recovery from large technology shocks [18]. However, much of these results rely on the assumption of RE. The recent crisis shows that the empirical validity of RE is questionable [15], and as a result, these policy measures should be tested in a more realistic environment, which allows the agents to *learn and adapt new information* [3]. One such interesting study was conducted by Honkapohja and Mitra [27], who show that under steady state learning, a simple form of adaptive (econometric) learning, PLT policy can be stabilizing only if the agents are aware of it, that is only if the central bank provides some guidance about its policy. More precisely, when agents’ inflation expectations are based on the price level target path then adaptive learning converges to the target steady state for almost all initial states. This is an important result, as adaptive learning is still quite a small departure from perfect rationality.

The goal of our paper is to provide an experimental test for these theoretical findings on PLT and guidance. Laboratory experiments are a novel and complementary approach in empirical macroeconomics [8,26]. The idea is to take a macroeconomic model, and replace economic agents with real subjects. The subjects then are tasked to decide on consumption, or to forecast aggregate variables as in Learning-to-Forecast experiments [23,24,9]. There are several virtues of this approach. In an experiment, unlike in real economies, the researcher can easily control the setup of the markets and manipulate policy measures and fundamentals. It is possible to directly observe the decisions or expectations of the subjects, and to obtain rich samples, in terms of both the number of independent observations from a constant environment (groups within a treatment), as well as of the time span within each group. Finally, macroeconomic models with micro-foundations aim at describing the ‘everyman’, normal consumers, many of whom have little or no economic education, similar to the subjects that were invited to our experiment.

To the best of our knowledge, the experimental literature on the role of guidance for monetary policy is rather limited, and does not unequivocally support this policy. Arifovic and Petersen [7] report that subjects find it marginally easier to coordinate on a path leading from ZLB if the central bank uses qualitative instead of quantitative guidance, which simulates high-information environment. However, this effect is not statistically significant. Furthermore Mokhtarzadeh and Petersen [31] show that when the central bank tries to manage subjects’ expectations through its projections, subjects react better (and coordinate on more stable dynamics) to inflation and output projections than to direct interest rate projections. This contradicts usefulness of what is typically meant as guidance by the interest rate. A similar experiment by Ahrens et al. [1] proves that only “high-quality” inflation projections can play such a stabilizing role.

We are aware of only two in-depth experimental studies focused directly on the role of PLT. Firstly, in the aforementioned experiment by Arifovic and Petersen [7], the authors focus on a ZLB recession and find that in the treatments with a PLT Taylor rule, compared with the baseline inflation target rule treatments, the economy does not recover faster from a series of negative demand shocks. In our design, by contrast, the demand shocks are modeled as a small white noise instead of a persistent auto-regressive process, which will yield a more direct picture of the macroeconomic dynamics induced by the PLT rule. Secondly, Amano et al. [2] use a standard DSGE model with inflation target and PLT Taylor rules to generate artificial, non-persistent inflation time series. Next, subjects are asked to forecast these inflation sequences. The authors find that the subjects predict inflation as if it was persistent [c.f., 28,31], and that they under-react to the observed price level

deviation. This study, however, uses exogenous, model generated inflation time-series, allowing for no expectation feedback mechanism for subjects to learn from. Nevertheless, we will confirm the pattern of the individual behavior, which was reported by Amano et al. [2], in our laboratory study where inflation and output gap depend on the individual forecasts.

Our experiment is inspired by Honkapohja and Mitra [27] and based on a simple DSGE model with Euler equation learning. We emphasize that we use the nonlinear version of the model, which allows us to study the off-equilibrium behavior of our subjects. This links our study to the ZLB experiment by Hommes et al. [26], though, unlike that paper, we focus on linear Taylor rules and constant fiscal policy. We run five treatments in total, which vary in terms of the monetary policy and subject information. First, we look at the baseline case of inflationary targeting. All other four treatments are based on a PLT Taylor rule, with weak or strong response coefficients, and with or without guidance. The difference between the strong and weak PLT Taylor rules is that the weak rule [suggested by 27] is stable under adaptive learning, but not under naive expectations, while the strong PLT is stable under both.

Our experimental economy is based on a forward-looking Taylor rule, which is typically interpreted as the central bank trying to manage expectations of the economic agents. As Pfajfar and Žakelj [32] point out, forward-looking Taylor rule tends to exacerbate instability of experimental economies, hence our experiment provides an unstable environment to test the usefulness of the PLT monetary policy. Future research should study the robustness of this experimental design.

In our experiment, following the definition of Honkapohja and Mitra [27], we take guidance in its simplest form, focusing on whether the central bank can help the subjects by being transparent about its policy goals. To be specific, in the guidance treatments, the subjects are explicitly informed about the role of the price level deviation in the monetary rule. During the experiment, this price level deviation from its target belongs to their information set. This design choice allows us to focus on the value of clear communication between the subjects and the central bank and its target [c.f. 7], and is a direct empirical test of the theoretical setup of Honkapohja and Mitra [27]. An alternative form of guidance is to manage subject expectations by providing them with the projections (expectations) of the central bank. An example of this so called forward guidance can be found in an experiment by Mokhtarzadeh and Petersen [31], and we find these two approaches as complimentary.

The experiment results in two major findings. First, informational guidance essentially plays no role, as the difference between guidance and no-guidance treatments is negligible. This implies that the subjects could not properly use the additional information given by the central bank. Secondly, the PLT treatments were stable only under the strong rule (with inflation even more stable than under the inflation treatment). For the weak PLT, in contrast to the theoretical stability under steady-state learning, the experimental economy remains unstable. This suggests that in an unstable economy the central bank may find it very difficult to compensate weak monetary policy with guiding markets by providing additional information about its target. Our experiment also shows that the subjects coordinated on simple forecasting rules, but the type of coordination depends on the specific Taylor rule. This confirms a common insight of Learning-to-Forecast experiments that there exists a feedback between stability of the underlying experimental economy and the type of forecasting behavior which the subjects learn [8]. In the context of monetary policy, a loose Taylor rule can induce the subjects to follow observed trends more closely, thus destabilizing the economy.

The paper is organized in the following way. Section 2 discusses the experimental design and the underlying New Keynesian macro model. Section 3 shows the experimental results and discusses the individual behavior of our subjects. Section 4 concludes and finally, experimental instructions and details on the econometric analysis can be found in the appendices.

2. Experimental design

In this section we discuss the design of the experiment. Section 2.1 provides details of the design, while Section 2.2 discusses some theoretical stability results motivating the parameterizations of the inflation targeting and PLT rules. Sections 2.3 and 2.4 discuss the information, which we have provided to the subjects, as well as the specific framing of the guidance. Section 2.5 formalizes the DSGE model that served as a foundation of the experimental economy. Section 2.6 briefly discusses the sessions and software. In Section 2.7 we formalize the hypotheses to be tested in our experiment.

2.1. Experimental treatments

Our experiment is based on a standard DSGE model with a zero-lower bound on the interest rate and Rotemberg price stickiness (we provide the details of the model in Section 2.5). The central bank follows a Taylor rule with either inflation target or nominal price level target. There are two state variables in the model: output y_t , (gross) inflation π_t , and in addition the policy variable, the (gross) interest rate R_t . The realized state of the economy depends on the expectations of future inflation and output. One can show that this model has two steady states (with self-fulfilling expectations): (1) full employment with inflation equal to the central bank's target π^* and the corresponding output y^* and interest rate R^* ; and (2) the zero-lower bound steady state (ZLB) such that $R^Z = 1$, with corresponding inflation $\pi^Z < 1$ (ie., with negative net inflation), and output, which is below its potential value, ie., $y^Z < y^*$.

Regardless of the treatment, there are $I = 6$ six subjects per group² who participate in a 51 periods long session. Within each period, each subject is asked to provide his or her *net* inflation forecast two periods ahead $\pi_{i,t+1}^e$ (the index $t + 1$ indicates the period, for which the forecast is formulated), and his or her output gap forecast two periods ahead $o_{i,t+1}^e$. The relationship between the output gap and output level is given by

$$o_t = 100 \frac{y_t - y^*}{y^*}. \quad (1)$$

We decided to ask the subjects for the output gap forecasts in order to facilitate initial learning, as this variable has a simple interpretation and its level does not depend on the specific calibration (which we did not show to the subjects). In the instructions and experimental interface, the inflation and output gap are framed as percentage points, for example $\pi_{i,t+1}^e = 1$ corresponds with a forecast of one percentage point of net inflation. These forecasts are used to compute the subjects' average gross inflation forecast $\bar{\pi}_{t+1}^e$ and the subjects' average consumption level forecast $\bar{c}_{t+1}^e$, which in turn are plugged into the DSGE model to determine the realized contemporary inflation π_t , output level y_t (which equals the realized consumption c_t plus fixed government spending $g_t = g$) and interest rate R_t . Subjects observe these realized variables as new net inflation, output gap and net interest rate, and the group moves to the next period.

The information provided to the subjects is discussed in detail in Section 2.2. In summary, each subject is given only a qualitative description of the economy, and can observe only the realized aggregate variables and her or his own forecasts and payoffs.

Subjects are paid in the same way across all treatments. For every inflation and output gap forecast we compute score as

$$\text{Points}_{i,t}^{\pi} = 100 \frac{1}{1 + |\pi_{i,t}^e - \pi_t|} \in (0, 100],$$

$$\text{Points}_{i,t}^y = 100 \frac{1}{1 + |o_{i,t}^e - o_t|} \in (0, 100],$$

so for example if a subject predicts -1% inflation, whereas the realized inflation is equal to -4% , the subject would receive 25 points for this forecast. At the end of the group's session, for each subject we add separately inflation and output gap scores, and pay only for inflation forecasting accuracy, or only for output gap forecasting accuracy randomly with probability 0.5 (following a roll of dice), with the exchange rate of 0.75 Euro for every 100 points. This means that the subjects could earn anything between 0 and 37.5 Euro. We decided to pay randomly for one task only in order to motivate subjects to pay high attention to both forecasting tasks, and to discourage them from hedging.

Our experiment has five treatments. There are two treatment differences: the specification of the Taylor rule, and in addition, subjects may or may not receive guidance in the form of additional information about the monetary policy of the central bank. These five treatments are:

1. **INF** – Central Bank uses a Taylor rule with *inflation target*, given by

$$\text{IFT: } R_t = 1 + \max \left\{ 0, R^* - 1 + \psi_{\pi} (\pi_{t+1}^e - \pi^*) + \psi_y \frac{y_{t+1}^e - y^*}{y^*} \right\}, \quad (2)$$

which is bounded from below by unity and where $\psi_{\pi} = 1.5$ and $\psi_y = 1$ are two policy parameters. We used these values as they guarantee that the economy is stable under naive expectations (see the model discussion in Section 2.5). This is the baseline treatment. Note that the interest rate is subject to the Zero Lower Bound constraint (ZLB), which is a common practice in macroeconomic experiments since the 2007 financial crisis.

In the other four treatments, the Central Bank always uses a *nominal price level target* in its Taylor rule, which is given by

$$\text{PLT: } R_t = 1 + \max \left\{ 0, R^* - 1 + \psi_P \frac{P_{t+1}^e - \bar{P}_{t+1}}{\bar{P}_{t+1}} + \psi_y \frac{y_{t+1}^e - y^*}{y^*} \right\}, \quad (3)$$

where $\bar{P}_t \equiv \pi^* \bar{P}_{t-1}$ defines the trajectory which the central bank takes as the intended price path (note that it is based on the same inflation target level π^* as in the **INF** treatment and is also subject to the ZLB constraint). This leads to the following four treatments based on a 2×2 design:

1. **StrongNo** – Strong PLT rule (3) with *high* policy parameters $\psi_P = 3$ and $\psi_y = 2$;
2. **WeakNo** – Weak PLT rule (3) with *low* policy parameters $\psi_P = 0.25$ and $\psi_y = 1$;
3. **StrongGuid** – Strong PLT rule (3) with *high* policy parameters $\psi_P = 3$ and $\psi_y = 2$ and additional guidance provided by the central bank to the subjects;

² Group size of 6 to 10 subjects is quite common in experimental macroeconomics. With 6 subjects strategic behavior is difficult, while groups are still small enough to collect a reasonable sample given cost constraints. See eg. Arifovic and Duffy [6] for a discussion.

Fig. 2.1. Stability of the experimental economy under PLT Taylor rule and naive expectations $\sigma_{t+1}^e = \sigma_{t-1}$ and $\pi_{t+1}^e = \pi_{t-1}$, as a function of policy parameters (ψ_P, ψ_Y) in the Taylor rule. Stable Taylor parametrization is represented by the black area. The blue dots indicate the weak and the strong PLT rules.

4. WeakGuid – Weak PLT rule (3) with low policy parameters $\psi_P = 0.25$ and $\psi_Y = 1$ and additional guidance provided by the central bank to the subjects.

As will be explained in Section 2.5, guidance means that the subjects are given additional information about the behavior of the central bank, namely that the central bank intends to keep prices at a certain target level. On the other hand, subjects in the two PLT treatments without guidance are given the same information as those in the inflation targeting. This allows us to test whether or not the PLT policy rule is more stable or unstable than inflation targeting.

The weak PLT rule corresponds to $\psi_P = 0.25$ and $\psi_Y = 1$ [as in 27], while the strong PLT rule corresponds with $\psi_P = 3$ and $\psi_Y = 2$. Notice that the strong rule is quite harsh, as e.g., one percentage point of the output gap implies that the interest rate increases by two percentage points.

The weak rule is stable under steady state adaptive learning, and therefore was suggested by the learning literature [27]. However, other learning models may lead to unstable dynamics. Previous literature on Learning-to-Forecast experiments demonstrates that adaptive learning may be too demanding for laboratory subjects, who tend to coordinate on simple forecasting rules instead. The design of earlier laboratory experiments shows that naive expectations may provide an indication of the early stage of a learning process by human subjects in the lab. Therefore, the next subsection discusses the stability under naive expectations as a motivation for our choice of the strong and the weak PLT Taylor rules in our experimental treatments.

2.2. Stability of experimental economy under naive expectations

With the noise switched off, i.e., $\varepsilon_t^c = \varepsilon_t^\pi = 0$, and under naive expectations

$$\begin{pmatrix} \pi_{t+1}^e \\ c_{t+1}^e \end{pmatrix} = \begin{pmatrix} \pi_{t-1} \\ c_{t-1} \end{pmatrix}, \quad (4)$$

the model becomes a deterministic non-linear 2D system under the inflation target rule and a 3D system under the PLT Taylor rule. One can find the analytical expression for the relevant Jacobian matrix, and next use numerical algorithm to compute the eigenvalues of the Jacobian at the full employment steady state. These results were confirmed by a simple finite elements approximation algorithm. Note that, regardless of the monetary policy, the system has an additional ZLB steady state, which is always a saddle. These eigenvalues of all treatments are summarized below.

Inflation targeting rule In the DSGE model with the inflation-based Taylor rule (2) specified as $\psi_\pi = 1.5$ and $\psi_Y = 1$, the full employment steady state is stable with real eigenvalues given by $\lambda_1^{INF} = 0.28679$ and $\lambda_2^{INF} = 0.88583$. Hence, the INF treatment is locally stable under naive expectations.

PLT rules Consider our DSGE model with the PLT-based Taylor rule (3). The stability results for all combinations of $\psi_P, \psi_Y \in [0, 3]$ can be found in Fig. 2.1, where the black area represents Taylor policy parameters under which the model is

stable under naive expectations. As can be seen in Fig. 2.1 the stability region is surprisingly small. For the weak rule with $(\psi_p, \psi_y) = (0.25, 1)$ the full employment steady state has one real eigenvalue $\lambda_1^{PLWe} = 0.21162$ and two conjugate complex eigenvalues $\lambda_{23}^{PLWe} = 1.0770 \pm 0.20153i$ with modulus equal to $|\lambda_{23}^{PLWe}| = 1.0957$. Therefore, for the weak PLT rule the full employment steady state is unstable under naive expectations.

Fig. 2.1 demonstrates that the central bank can stabilize the economy with a PLT Taylor rule only if it uses a much stronger specification than the weak rule. For instance, if the PLT Taylor rule is parametrized with $(\psi_p, \psi_y) = (3, 2)$, the system in the full employment steady state has one real eigenvalue $\lambda_1^{PLSt} = -0.48274$ and two complex conjugate eigenvalues $\lambda_{23}^{PLSt} = 0.85067 \pm 0.52406i$ with modulus equal to $|\lambda_{23}^{PLSt}| = 0.99914$, which means that the full employment steady state is stable, but close to the edge of instability.

The dynamics under naive expectations can be visualized by phase plots in the inflation consumption space (Fig. 2.2) and sample time paths. A clear difference between the three monetary policy treatments can be observed. Under inflation targeting (Fig. 2.2a), smooth dynamics emerge. If the system is initialized close enough to the full employment steady state (such as in $(\pi_0, c_0) = (1.03, 0.78)$), it gradually converges to this equilibrium. On the other hand, if the system starts too far away (such as in $(\pi_0, c_0) = (1.14, 0.69)$), it bounces of the convergent region and falls into an inflation-output contraction spiral.

The dynamics under the two PLT Taylor rules are only remotely similar to those from the inflation rule treatment. Under the strong PLT Taylor rule (Figure 2.2b), the two initial points have the same long-run outcome as under the inflation Taylor rule. The stable initialization $(\pi_0, c_0) = (1.03, 0.78)$ seems at first to be more irregular, but it gradually converges. The unstable initialization $(\pi_0, c_0) = (1.14, 0.69)$ diverges immediately, without the transitory dynamics as under the inflation rule. Under the weak PLT Taylor rule (Fig 2.2 c), the dynamics are highly unstable and both initializations eventually diverge from the steady state.

To sum up, the difference between weak and strong PLT rules is that only the model with the strong rule is stable under homogeneous naive expectations. On the other hand, the weak rule is stable under adaptive learning, and correspondingly was suggested by the learning literature [27]. Therefore, differences between treatments 2 and 3, and between 4 and 5 will allow us to test whether subjects are closer to a simple forecasting rule, such as naive expectations, or whether they are actually able to use more sophisticated steady state adaptive learning.

2.3. Information provided to subjects

In all five treatments, subjects are told that they act as professional forecasters for a statistical bureau. We inform the subjects that they belong to a fixed group and that other subjects have the same task, but we do not specify the exact size or composition of their groups. We provide the subjects with a qualitative description of the economy. In particular, subjects are informed about the positive and negative signs of the relationship between their predictions and the realized interest rate, inflation and output gap. However, we do not show the subjects the exact mathematical formulation of the DSGE model that is underlying the experimental economy, and subjects do not know the Taylor rule, with which the central bank sets the interest rate. Furthermore, in the four PLT treatments, the subjects are not informed about whether the central bank is using the weak or the strong PLT Taylor rule.

Every subject is informed that she or he will observe the past realized inflation, output gap and interest rate, as well as her or his own past inflation and output gap forecasts, and the corresponding earnings. However, no subject is shown the forecasts or score of any other subject. Since, the experimental economy is based on a two-period ahead expectations feedback, we inform the subjects that they will receive information about their performance and the realized variables only after the first two forecasts. In order to guide them in these initial periods, we mention in the instructions that both the inflation and output gap 'have historically been between -5% and 8% '.

The difference between treatments **INF**, **StrongNo** and **WeakNo** (no guidance) on the one hand, and **StrongGuid** and **WeakGuid** (guidance) on the other hand, is that the subjects in the guidance treatments **StrongGuid** and **WeakGuid** receive additional information about the interest rate policy of the Central Bank. Namely, we explain in detail that the central banks wants the price level to follow a certain trajectory. We inform them that they cannot directly observe that path, but they will be given the exact value of the relative deviation of the prices from this intended path, that is, subjects observe the $(P_{t-1} - \bar{P}_{t-1})/\bar{P}_{t-1}$ variable. Subjects are informed that the Central Bank increases (decreases) the interest rate when the predicted price level deviation is too high (too low).

Subjects' inflation and output gap forecasts have to stay within the $[-5\%, 15\%]$ interval (regardless of the treatment), to rule out explosive dynamics. However, we did not want the framing to preclude subject coordination on the run-away trajectories, as not to interfere with the initial learning periods. Therefore, subjects were not initially informed about the constraint on their forecasts. Instead, the subject interface was programmed to issue a warning whenever a subject would try to submit a forecast outside of the allowed interval.

The instructions cover in detail this information. In addition, they contain a table and graph of the payoff function, and explain screenshots of the subject screen. During the session, subject's screen displays graphs and tables of the realized variables (inflation, output gap, interest rate; and under the guidance treatments the relative price level deviation), the individual past forecasts and scores (per period and accumulated), number of the current period and remaining number of forecasts that the subjects should submit this period. We use a number of control questions to test whether the subjects

Fig. 2.2. Phase plots of the experimental economy under three different Taylor rules and naive expectations. The state space is $\pi_t \times c_t$. Notice that under PLT rules the system is actually 3D, so these phase plots are projections with $p_0 = \hat{p}_0$. Red arrows denote the direction of the system at any point. Blue and green points denote manifolds with stable inflation and consumption respectively. The black points denote the steady state. Finally, black arrows denote two sample paths that, for each Taylor rule, start at $(\pi_0, c_0) = (1.03, 0.78)$ and $(\pi_0, c_0) = (1.14, 0.69)$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

understand the instructions. Appendix [Appendix A](#) contains the instructions and control questions for the different treatments.

2.4. Guidance specification

In our experiment, guidance implies that (1) the subjects are explicitly informed that the central bank *commits* to minimizing price level deviations, and (2) throughout the session the subjects are informed (by a graph and table) about the realized values of the relative price level deviation. This is the simplest possible version of guidance, which provides subjects with the knowledge about central bank's target and reaction to the realized price deviation path which they observe throughout their session. In contrast, under the non-guidance treatments, subjects are only informed that the central bank will increase (decrease) the interest rate when inflation and/or output gap forecasts are too high (too low), without any further details on the target of the Central Bank.³

We decided to use this simplest formulation of guidance for the following reasons. Honkapohja and Mitra [27] demonstrate that the PLT rule is stable under the adaptive learning only if the agents realize that the central bank is targeting the price level and include this variable in their perceived law of motion. On the other hand, central banks may face practical communication issues, which implies that the introduction of the PLT policy may be misunderstood by the general public. The difference between **StrongGuid** and **WeakGuid** treatments on the one hand, and **StrongNo** and **WeakNo** ones on the other hand, will allow us to measure the risk and effects of such a miscommunication and test the extent to which the laboratory subjects can learn to behave 'as-if' they knew the new policy. The second reason is that we wanted to compare the stabilizing or destabilizing effects of the inflation targeting and the PLT rules in a laboratory setting with the same information.

2.5. Experimental economy

The economy in our experiment is based on a standard DSGE model with zero lower bound and Rotemberg price stickiness, in its Euler learning form. For the derivation of the model, see for example Honkapohja and Mitra [27].⁴ The model is based on a representative Ricardian consumer-firm owner, who maximizes an infinite discounted sum of utility subject to standard production, saving and market clearing conditions. Assuming fixed government spending $g_t = g$, non-distortionary lump-sum taxes and an interest rate rule $R_t(\cdot)$, the model can be described as a two dimensional system such that in period t , the consumption level c_t and the gross inflation π_t are a function of the two-period ahead consumer expectations of these two variables. The relationship is given by the following set of nonlinear equations:

1. The aggregate consumption:

$$c_t = c_{t+1}^e \left(\frac{\pi_{t+1}^e}{\beta R_t} \right)^{1/\sigma} + \varepsilon_t^c. \quad (5)$$

2. The Phillips curve:

$$\pi_t = Q^{-1}[K(c_t, \pi_{t+1}^e)] + \varepsilon_t^\pi, \quad (6)$$

where

$$Q(\pi_t) = (\pi_t - 1)\pi_t$$

and⁵

$$\begin{aligned} \kappa(c_t, \pi_{t+1}^e) &= \beta \pi_{t+1}^e (\pi_{t+1}^e - 1) + \frac{\nu}{\alpha \gamma} (c_t + \bar{g})^{(1+\varepsilon)/\alpha} + \frac{1-\nu}{\gamma} (c_t + \bar{g}) c_t^{-\sigma}, K(c_t, \pi_{t+1}^e) \\ &= \begin{cases} \kappa(c_t, \pi_{t+1}^e) & \text{if } \kappa(c_t, \pi_{t+1}^e) > -0.25, \\ -0.25 & \text{else.} \end{cases} \end{aligned} \quad (7)$$

Notice that the interest rate rule $R_t(\cdot)$ depends on the treatment. It is either based on an inflation target (2) or price level target (3). In the latter case, the central bank keeps track of the realized price level, which increases the dimensionality of the model to three. [Table 1](#) summarizes the parameter values, and the corresponding full employment and ZLB steady

³ The term 'guidance' has different meanings in the macroeconomic literature. In particular, 'forward guidance' means that the central bank announces, and commits to an interest rate policy such that if inflation or output fall below a specific threshold, the central bank will keep the nominal interest rate at some low level for a predetermined number of periods. This version of guidance requires a number of more complex design choices (like the threshold conditions, value and duration of the 'low nominal interest rate'). For forward-guidance experiments, see Ahrens et al. [1] and [31].

⁴ Note that Honkapohja and Mitra [27] use so-called steady state learning version of the DSGE model for their adaptive learning. We decided to use the Euler learning instead, as the DSGE model under steady state learning and naive expectations is extremely unstable. As it will be apparent in the next section, naive expectations fit the behavior of our subjects better than the adaptive learning.

⁵ The lower bound $K(\cdot) > -0.25$ in [Eq. \(7\)](#) is necessary to avoid complex values of the realized inflation, and corresponds with realized gross inflation $\pi_t = 0.5$, i.e. scenario where prices fall by 50% every period.

Table 1
Experimental economy parametrization.

Parameter	Notation	Value
Number of agents/subjects	I	6
Discount factor	β	0.99
Government spending	$\bar{g}$	0.2
Output elasticity	α	0.7
Rotemberg price stickiness	γ	350
Labor supply elasticity	ϵ	1
Demand's elasticity of substitution	ν	21
Consumption elasticity	σ	1
Gross inflation target	π^*	1.05
Steady state gross interest rate	R^*	1.(06)
Steady state consumption	c^*	0.745358
Steady state output	y^*	0.945358
ZLB gross inflation	π^Z	0.99
ZLB gross interest rate	R^Z	1
ZLB consumption	c^Z	0.742765
ZLB output	y^Z	0.942765
ZLB output gap	o^Z	−0.2766%
Taylor rules		
Inflation target rule	(ψ_π, ψ_y)	(1.5,1)
Weak PLT rule	(ψ_p, ψ_y)	(0.25,1)
Strong PLT rule	(ψ_p, ψ_y)	(3,2)

states. The two small noise terms in Eqs. (5) and (6) are uncorrelated IID shocks to the consumption and inflation, such that $\varepsilon_t^c \sim NID(0, (0.0005c^*)^2)$, $\varepsilon_t^\pi \sim NID(0, 0.0005^2)$ and $Cov(\varepsilon_t^c, \varepsilon_t^\pi) = 0$. The small variance of the two noise terms was chosen so that if the agents would repeatedly forecast the full employment steady state, the realized inflation and consumption both stay within one permil point (0.1%) of the steady state 95% of time.⁶

Under Rational Expectations, the model is solved to obtain model consistent expectations such that for every period t it holds that $\pi_{t+1}^e = E_t(\pi_{t+1})$ and $c_{t+1}^e = E_t(c_{t+1})$, that is the representative consumer does not make systematic errors. However, in the experiment we take the average net inflation and output gap forecasts of the six subjects, transform them into the consumption level and gross inflation level forecasts and directly input them into the interest rate rule ((2) in treatment **INF**, and (3) in treatments **StrongNo**, **WeakNo**, **StrongGuid** and **WeakGuid**), the consumption rule (5) and finally the Phillips curve (6).⁷ Notice that the permitted forecasting intervals (between −5% and 15%) allow for coordination on the ZLB steady state, as well as on interesting dynamics around both steady states.

Previous experimental work suggests that subjects may fail to form rational expectations. Instead, we expect their forecasts to be much closer to adaptive expectations [8,26,33]. Under pure naive expectations, treatment **INF** is stable. As for the PLT treatments, guidance plays no role under naive expectations. On the other hand, the Taylor rule parametrization is crucial. The PLT Taylor rule with $(\psi_p, \psi_y) = (0.25, 1)$ [as in 27] is highly unstable under naive expectations. Therefore, we decided to run additional PLT treatments with $(\psi_p, \psi_y) = (3, 2)$, for which the system becomes stable under naive expectations. Hence, the treatment **StrongNo**, unlike **WeakNo**, is expected to yield potentially converging dynamics; and the first test for guidance will be the difference between treatments **WeakNo** and **WeakGuid**.

2.6. Experimental sessions and software

We run 6 groups for each of the 5 treatments, recruiting 180 subjects in total. The sessions were conducted at the CREED laboratory, University of Amsterdam, in November and December 2015, and January 2016. We wrote the experimental software in C++, using standard library and Wt, a C++ Web Toolkit under the standard GNU General Public License.⁸ The duration of each session was typically around two hours. We asked subjects for 51 pairs of forecasts, which results in 50 periods of data per group.⁹

⁶ Note that since the error terms are white noise instead of an autocorrelated process, under homogeneous RE the rational agents immediately converge to one of the steady states and there are no dynamics aside from the errors.

⁷ LtF macroeconomic experiments use either average forecast [eg. 8] or median forecast [eg. 31]. Given that our subjects tend to be well coordinated, our results appear to be robust to this design choice.

⁸ The Wt library is available at <http://www.webtoolkit.eu/wt>. The software (compiled for the Windows 7 operating system), as well as the source code can be provided on demand.

⁹ Due to an unexplained software or hardware failure, the last period was not recorded in the case of four groups, leaving only 49 data points. These groups are INF05, PLTStrongGuid01, PLTWeakGuid02 and PLTWeakGuid05. In addition, one subject in group INF07 was extremely slow. Despite the help of the experimenters, he was unable to act efficiently throughout the session, which then had to be terminated after period 41. We leave this group out of the following data analysis, but the group's results are presented in Appendix Appendix B.

2.7. Testable hypotheses

Under Rational Expectations, the experimental economy should immediately converge to either the ZLB or the full employment steady state. However, previous experimental evidence suggests that within the New Keynesian macroeconomic framework, the subjects do not truly converge to these steady states, or they do so only after a prolonged period of time. Therefore, our experiment will serve to directly test the following hypotheses:

- Hypothesis 1 Weak PLT rule is sufficient to stabilize the economy. All four PLT treatments (**StrongNo**, **WeakNo**, **StrongGuid** and **WeakGuid**) will exhibit stable dynamics (convergence or mild oscillations).
- Hypothesis 2 Guidance will help in stabilizing the economy, which implies that treatment **StrongGuid** is *more stable* than treatment **StrongNo**, and treatment **WeakGuid** is *more stable* than **WeakNo**.
- Hypothesis 3 Subjects will learn the fundamental steady state rather than use more complicated heuristics with trend following.

Hypothesis 1 and **Hypothesis 2** can be directly translated into statistical tests. Following the literature, we will compare the relevant treatments by testing the differences in the distribution of inflation and output gap stability. We will quantify this with the Relative Absolute Deviations [35]. Under adaptive (econometric) learning, both **Hypotheses 1** and **2** should be true [27]. In order to test **Hypothesis 3**, we will estimate simple behavioral forecasting rules for each subject and investigate the resulting distribution.

3. Experimental results

In this section we discuss the results from our experiment. We will start with a general overview of the aggregate dynamics, which was observed in the five treatments. In [Section 3.2](#), we will test the differences between the treatments in terms of stability. [Section 3.3](#) will give a brief description of the individual behavior.

3.1. Overview

The first observation is that all 30 groups in all treatments started close to each other, with both the initial inflation and the initial output gap in the vicinity of 2.5%. Second observation is that despite similar initial conditions, the experiment resulted in clear differences between the treatments. [Fig. 3.1](#) (spread over two pages) shows realized group inflation and output gap for representative groups from all treatments. [Fig. 3.2](#) shows realized inflation and output gap paths separately for each treatment, and price level deviations for the four PLT treatments. The individual figures for all groups can be found in [Appendix Appendix B](#).

Under the inflation treatment (**INF**), two types of dynamics are possible. First, mild oscillations appear in some of the groups, as in the case of group 4 from that treatment ([Fig. 3.1a](#) and b). Despite some mild instability, subjects are well coordinated and quickly learn to predict accurately both variables. Second, some of the groups exhibit convergent dynamics, as in the case of group 6 ([Appendix Appendix B](#)). Even though the economy does not always converge to the full employment steady state, this treatment seems relatively stable.

Similar dynamics appear under the strong Taylor rule PLT treatment without guidance (**StrongNo**). [Fig. 3.1c](#) and d show the results for group 4, in which we observe a high degree of coordination and converging dynamics. The groups under this treatment seem to be somewhat more unstable in the initial periods, and they tend to generate faster business cycle fluctuations than under the inflation treatment **INF**, but they all eventually settle down close to the full employment steady state. This coincides with successful monetary policy, where the price level deviation quickly falls close to zero ([Fig. 3.2g](#)). On the other hand, under the no guidance treatment with weak PLT Taylor rule (**WeakNo**), all six groups exhibited highly unstable dynamics (see [Fig. 3.1e](#) and f). Subjects repeatedly hit the upper and lower boundaries of their forecasts, which results in fast oscillations with large amplitude. Without these forecasting constraints, these economies would likely collapse to zero output or diverge. In addition, subjects find it much more difficult to coordinate in this unstable environment, they make higher forecasting errors and obtain lower earnings. All of this leads to failure of the monetary policy, where the price level deviation oscillates with a large amplitude ([Fig. 3.2h](#)).

Interestingly, the guidance seems to have little effect on the dynamics under the PLT Taylor rule. [Fig. 3.1g](#) and h show results from group 2 from the strong PLT treatment with guidance (**StrongGuid**), and [Fig. 3.1i](#) and j shows results from group 3 under the weak PLT treatment with guidance (**WeakGuid**). The realized inflation and output gap in these two groups look remarkably similar to those from their no guidance counterparts. The next subsection tests these stylized observations in a formal statistical fashion.

3.2. Measuring the stability of the treatments

We use the standard measure of Relative Absolute Deviation (RAD) to quantify the degree, to which the five different monetary policies stabilized the economy [see 35, for a general definition and discussion]. The RAD measure for variable x

Fig. 3.1. Experimental results, with one representative group per treatment. Left panels display the inflation, and right panels the output gap. In each figure, the realized variable is shown with a black line, while the corresponding six subject forecasts are shown with dashed green lines.

and group g is defined as

$$RAD_g = \frac{1}{T-10} \sum_{t=11}^T \frac{|x_{g,t} - x_t^f|}{x_t^f} \quad \text{for } x^f \neq 0, \quad (8)$$

where T denotes the length of the session of group g , $x_{g,t}$ denotes the realized value of x in period t in group g and x_t^f denotes the fundamental value of variable x in period t . In our case, the fundamental inflation is $\pi^* = 1.05$, while the RAD for the output by definition coincides with the average absolute output gap, expressed as a fraction (compare the definition of the RAD (8) with the definition of the output gap in Eq. (1)). Note that we exclude the initial 10 observations, as we interpret these as an initial learning phase.

The RAD measures can be found in Table 2. Furthermore, we test the RAD distribution differences between the treatments with Mann–Whitney U test (MWU test), and the resulting test statistics are reported in Table 3. The first clear observation is that there is little variation between the guidance and their respective no guidance treatments (**StrongGuid** versus **StrongNo**

Fig. 3.2. Realized inflation, output gap and price level deviations for each treatment. Each graph shows either realized inflation, realized output gap, or price level deviation under the four PLT treatments, in six groups in each treatment.

Table 2

Relative Absolute Deviation (RAD) of the experimental **inflation** and **output gap** for the five treatments, in percentages. *** (**) denotes groups for which the average RAD from the last 40 periods is larger than 2 percentage points on 1% (5%) significance level.

Group	INF	StrongGuid	StrongNo	WeakGuid	WeakNo
Inflation					
#1	0.85	0.37	0.207	5.934***	4.31***
#2	0.805	0.241	0.146	2.308	6.066***
#3	0.77	0.133	0.0557	5.275***	4.803***
#4	0.531	0.66	0.206	9.1***	10.13***
#5	2.937***	0.165	0.118	8.903***	4.26***
#6	0.336	0.0847	0.337	6.858***	4.742***
Average	1.038	0.2753	0.1783	6.396	5.719
Output gap					
#1	0.503	0.822	0.674	5.058	3.663
#2	0.51	1.175	0.953	5.741***	4.992
#3	0.452	0.339	0.194	4.334	5.689***
#4	0.316	1.745	0.495	9.126	11.09
#5	1.731	0.553	0.49	9.795	3.534
#6	0.221	0.235	1.28	6.191	5.025
Average	0.6219	0.8113	0.6811	6.708	5.666

and **WeakGuid** versus **WeakNo**). This is confirmed by the MWU test, according to which there is no significant difference between the distribution of inflation and output gap RAD measures between these two pairs of treatments. From this, we conclude that **Hypothesis 2** is rejected: guidance, as framed in our experiment, does not have a significant effect on the stability of the economy.

Table 3

Mann Whitney U test statistics for the differences in the distribution of the Relative Absolute Deviation of the experimental **inflation** and **output gap** for the five treatments. For a 6×6 sample the critical value for 5% p-value is equal to 5. * denotes treatments for which the difference between the treatments is statistically significant.

Treatment	StrongGuid	StrongNo	WeakGuid	WeakNo
Inflation				
INF	3*	1*	1*	0*
StrongGuid		13	0*	0*
StrongNo			0*	0*
WeakGuid				13
Output gap				
INF	12	15	0*	0*
StrongGuid		16	0*	0*
StrongNo			0*	0*
WeakGuid				10

A second observation is that there are significant differences (according to MWU tests) between the two weak PLT Taylor rule treatments and the three other treatments. Under the weak PLT Taylor rule the economies have explosive dynamics, while the strong PLT Taylor rule can enforce convergence to the full employment steady state. In addition, RAD measures imply that the PLT treatments with the strong Taylor rule outperform the inflation target rule in terms of the inflation stability, but not in terms of the output gap stability. We conclude that, in line with the theoretical literature, the price level target indeed can be a promising monetary policy, but only if sufficiently reactive. This rejects **Hypothesis 1** and implies that the weak PLT Taylor rule is not sufficiently responsive to the business cycle to stabilize it.

3.3. Individual behavior

The experimental results reject both **Hypothesis 1** and **2**, which suggest that the subjects coordinated on rules that may contain more elements than just the full-employment steady state levels. To further test **Hypothesis 3**, we use the following procedure.

In the experiment, we organized 6 groups per treatment, which in total gives us a sample of 180 subjects. Following the previous literature [eg. 9], we assume that the subject forecasting can be explained by the following (two dimensional) first-order rule:

$$\begin{aligned}\pi_{t+1}^e &= c^\pi + \alpha_1^\pi \pi_t^e + \alpha_2^\pi \pi_{t-1} + \alpha_3^\pi o_{t-1} + \beta^\pi (\pi_{t-1} - \pi_{t-2}) + \delta^\pi r_{t-1} + \gamma^\pi D_{t-1} + \varepsilon_t^\pi, \\ o_{t+1}^e &= c^o + \alpha_1^o o_t^e + \alpha_2^o o_{t-1} + \alpha_3^o \pi_{t-1} + \beta^o (o_{t-1} - o_{t-2}) + \delta^o r_{t-1} + \gamma^o D_{t-1} + \varepsilon_t^o,\end{aligned}\quad (9)$$

where the superscript e denotes forecasts, π denotes the inflation, o denotes the output gap, r denotes the net interest rate and D denotes the price level deviation. For clarity, we suppress the subject index, but in principal subjects can use very different specifications of (9). Note that D_t is used only in the guidance treatments, otherwise γ 's are set to zero. Finally,

$$\begin{pmatrix} \varepsilon_t^\pi \\ \varepsilon_t^o \end{pmatrix} \equiv \varepsilon_t \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\pi^2 & \rho_{\pi,o} \sigma_\pi \sigma_o \\ \rho_{\pi,o} \sigma_\pi \sigma_o & \sigma_o^2 \end{pmatrix}\right)\quad (10)$$

are jointly normally distributed random errors, which are pairwise independent in the dimensions of time and subjects, but not necessarily for one subject in one period ($\rho_{\pi,o} \neq 0$).

For each subject, we estimate the forecasting rule (9) jointly with a simple Maximum Likelihood estimator, which allows us to use the Likelihood Ratio test for all the following tests, for the two rules jointly. We add lags of the two forecasts until there is no evidence of auto-correlation, and hence eliminate the insignificant variables until the two rules contain only significant ones (if any). For the details on the estimation algorithm, and the results for each individual subject, see Appendix Appendix C.

We were able to estimate a two-dimensional rule for all 180 subjects, and almost all of these rules are non-trivial (that is they contain significant variables). Average estimated coefficients, as well as number of significant coefficients per treatment, are provided in Table 4. See also Appendix Appendix C for the full estimation results and their visual representation in Figs. C.1 and C.2. Given the complexity and heterogeneity of these 180 individual rules, one cannot clearly classify them into common, simple types. However, some stylized facts emerge:

1. Under Rational Expectations, subjects should quickly converge to the fundamentalist forecasting rule with $\pi_{t+1}^e = 5\%$ and $o_{t+1}^e = 0\%$ (see Footnote 5). None of the subjects learned this RE forecasting rule. However, some subjects followed richer rules with one or two constants equal to the corresponding fundamental values. This implies that some of the subjects learned the value of the steady state inflation and output gap, but they would also extrapolate short-run trends.
2. Subjects remained largely heterogeneous within groups, within treatments and between treatments. This means that they used both different values and different subsets of coefficients, even if they belonged to the same group.

Table 4

Average estimated forecasting rules. SNo, WNo, SGU and WGu denote **StrongNo**, **WeakNo**, **StrongGuid** and **WeakGuid** treatments respectively. SD denotes standard deviation across the 36 subjects from the treatment and Sign. denotes number of subjects within each treatment (out of 36), for which the coefficient was significant on 5% significance level. Remark that insignificant coefficients are treated as equal to 0 for computing the mean and SD.

	Inflation forecasting rule							Output gap forecasting rule						
	c^π	π_t^e	π_{t-1}	o_t	$\Delta\pi_{t-1}$	r_{t-1}	D_{t-1}	c^o	o_t^e	o_{t-1}	π_{t-1}	Δo_{t-1}	r_{t-1}	D_{t-1}
INF	0.568	0.373	0.431	0	0.586	0.0528	n/a	0.512	0.176	0.539	−0.0638	0.413	−0.022	n/a
(SD)	(1.3)	(0.443)	(0.528)	(0)	(0.645)	(0.279)		(1.44)	(0.3)	(0.558)	(0.263)	(0.609)	(0.174)	
Sign.	11	16	18	0	18	3		18	17	27	2	15	14	
SNo	1.72	0.193	0.484	0	0.397	−0.0229	n/a	0.633	0.178	0.653	−0.105	0.17	0.00328	n/a
(SD)	(1.94)	(0.303)	(0.517)	(0)	(0.763)	(0.097)		(2.89)	(0.249)	(0.437)	(0.524)	(0.408)	(0.351)	
Sign.	19	16	18	0	13	9		23	22	32	6	10	18	
WNo	1.88	0.131	0.0522	0.327	0.496	0.276	n/a	1.17	0.101	0.801	−0.478	0.436	0.293	n/a
(SD)	(1.92)	(0.303)	(0.742)	(0.535)	(0.312)	(0.369)		(1.44)	(0.292)	(0.47)	(0.473)	(0.277)	(0.389)	
Sign.	26	22	25	13	29	21		32	35	36	31	36	36	
SGU	1.42	0.212	0.507	−0.0967	0.299	−0.0249	−0.394	1.17	0.206	0.455	−0.193	0.216	0.00106	−0.276
(SD)	(2.7)	(0.344)	(0.665)	(0.572)	(0.542)	(0.189)	(1.01)	(3.88)	(0.305)	(1.39)	(0.998)	(0.496)	(0.486)	(2.31)
Sign.	19	16	18	1	13	9	13	23	27	31	9	18	22	22
WGu	0.616	0.15	0.632	−0.112	0.513	0.106	−0.0313	−0.32	0.127	0.467	−0.118	0.428	0.277	−0.0455
(SD)	(3.66)	(0.286)	(0.801)	(0.99)	(0.354)	(0.319)	(0.212)	(4.23)	(0.271)	(1.16)	(0.865)	(0.302)	(0.553)	(0.236)
Sign.	19	22	32	10	34	23	23	24	28	32	29	32	33	28

Fig. 3.3. Stability of the experimental economy under PLT Taylor rule, as a function of policy parameters (ψ_p, ψ_y) in the Taylor rule. ω share of the agents use the average heuristic estimated for the **StrongNo** treatment and $1 - \omega$ use the average heuristic estimated for the **WeakNo** treatment. Stable Taylor parametrization is represented by the black area. The blue dots indicate the weak and the strong PLT rules.

3. In general the inflation rules are simpler: they contain fewer significant coefficients.
4. A clear between-treatments pattern is that the subjects used simpler rules under the inflation targeting **INF** and strong Taylor rule PLT treatments **StrongNo** and **StrongGuid**. We interpret this as a sign that in a more difficult, unstable environment, subjects try to use more information.
5. Subjects learn to incorporate guidance information and the interest rate mostly under the unstable treatment **WeakGuid**. For example, δ^π is significant for 13 subjects under **StrongGuid** and for 23 subjects under **WeakGuid** treatment (out of 36 subjects per each treatment). This leads to an interesting result that people seek the monetary authority's guidance if times are unstable, but this does not counter the effect of a relatively weak response of the central bank. It also corresponds well with the previous stylized findings.
6. A typical result is that the subjects used 'close to' adaptive rules (with significant weights on past forecasts and/or past observations, not necessarily adding up to unity). In addition, trend chasing is popular, with both β 's often significant and then positive.
7. There is a clear relationship between the stability of a treatment, and degree to which subjects learn to follow trends in realized inflation and output gap. Compared to the strong PLT Taylor rule treatments **StrongNo** and **StrongGuid**, subjects under the weak PLT rule treatments **WeakNo** and **WeakGuid** are more likely to have a trend following component in their heuristics, and if they do, it is approximately twice as large. On the other hand, trend heuristics were as common under the two strong PLT rule treatments as under the **INF** treatment, however, in the latter treatment the average coefficients are approximately twice as large. This corresponds well to the stability ranking of the treatments, as explained in the previous part of this section, and therefore shows that the learning of the trend chasing behavior and the macroeconomic instability are self-reinforcing processes.

Altogether, these results reject **Hypothesis 3**. Furthermore, they are in line with Amano et al. [2], who in their experimental study demonstrate that the subjects under-react to the price-level deviation and instead use forecasting rules with a high weight on realized inflation. Finally, the observed trend following behavior, and its relationship with market stability, corroborates the literature on Learning to Forecast experiments [23,26].

3.4. Individual behavior and macroeconomic stability

The results from the previous part of this Section highlight a common finding in Learning-to-Forecast experiments: there exists a feedback between individual learning and the macroeconomic stability of the experimental economy [8]. In particular, subjects will be more likely to extrapolate the observed trends, if the underlying experimental model is less stable (under eg. naive expectations benchmark). This finding points to the crux of the argument behind the Lucas critique: in order to study the effects of a certain policy change, we need to consider how this change will effect the individual behavior. In the case of a PLT rule: how will this Taylor rule impact the dynamic properties of individual learning?

Our experiment suggests that the PLT rule has to be relatively harsh to stabilize the economy. To further emphasize this point, we performed a simple theoretical exercise. Suppose that in a model with a mass of forecasters, a share of ω agents uses a forecasting heuristic equal to the average subject rule estimated for the **StrongNo** treatment, and a share of $1 - \omega$ uses the average forecasting heuristic estimated for the **WeakNo** treatment (see Table 4 for the coefficients; we focus on the treatments without guidance for the sake of simplicity). The stability results can be found on Fig. 3.3. As long as everybody is using the heuristic from the **StrongNo** treatment ($\omega = 1$), the PLT rule supports a vast number of stable parametrizations. However, once the share of these agents starts to contract and more agents start to use the **WeakNo** forecasting heuristic, the space of stable parametrizations of the PLT rule drops quickly. In fact there exists no stable parametrization at all once ω crosses the threshold of around 0.92 (with the **WeakNo** heuristic having only an 8% share of the population). Notice that in practice this drop would likely appear even faster, since in this exercise the share of heuristics is fixed, whereas in real life it would in fact depend on the policy parametrization itself, through the aforementioned feedback. This simple exercise confirms that our strong parametrization of the rule is in fact quite close to the boundary of the stable ones. Assenza et al. [8] study the interaction between individual learning and aggregate behavior in a HSM where the fractions change over time and agents adapt their forecasting rules in response to the policy rule – a similar setup could be used in future research on the PLT rule.

4. Conclusions

In this paper we report a simple Learning-to-Forecast experiment on the stabilizing effect of Price Level Targeting (PLT) and guidance. We focus on a standard non-linear DSGE economy, in which the realized inflation and output gap depend on two-period ahead forecasts of these two variables and the specific (linear) Taylor interest rate rule. We ask the subjects to forecast inflation and output gap and pay them for their forecasting accuracy. Subjects are given only a qualitative description of the economy. We consider five treatments: (1) with an inflation target Taylor rule; and with PLT Taylor rule with (2) strong parametrization, (3) weak parametrization, (4) strong parametrization and guidance, and (5) weak parametrization and guidance. We frame the guidance as additional information: subjects (i) are explicitly informed about how the interest rate rule depends on the relative deviation of the intended price level, and (ii) can observe the realized price level deviation throughout the session. We find that the guidance has no effect on the aggregate behavior. The PLT rule can be a

robust stable monetary policy, but only if it reacts sufficiently strong to the price level deviation and output gap. Under a weak PLT rule the economy is stable under adaptive learning, but highly unstable in the lab. This shows that the subjects do not follow sophisticated adaptive (econometric) learning, but instead use simple forecasting heuristics, such as adaptive expectations and/or trend following.

After the recent financial crisis of 2007, developed economies plummeted into a prolonged recession. Central banks failed to stimulate the economies with interest rate cuts, as they hit the zero-lower bound constraint. This shows the relevance of the zero-lower bound steady state, which is characterized by a spiral of deflation and output contraction. The macroeconomic literature focused on developing monetary policies that could pull the economies from the recession, and prevent future crises of a similar magnitude. Two prominent ideas are guidance and Taylor rule based on nominal price level target (PLT). Guidance means that the central bank more openly discloses (and commits to) its monetary policy rule.

Both PLT and guidance have promising properties under Rational Expectations, since they give the economic agents more information about the monetary policy, and thus allow for a faster convergence towards the full employment steady state. Nevertheless, Rational Expectations require an unrealistic cognitive load from economic agents such as consumers or firms. On the other hand, even a small departure from Rational Expectations, can lead to different evaluation of PLT and guidance. For instance, Honkapohja and Mitra [27] show that under adaptive (econometric) learning, PLT Taylor rule actually requires the guidance to stabilize the economy.

The main goal of our paper is to provide an experimental test to the robustness of guidance and PLT monetary policies. We regard our work as complementary to the theoretical literature on learning in macroeconomics. We consider guidance framed as in Honkapohja and Mitra [27]: additional information, which the agents receive about the deviations from the target of the central bank. In addition, we look at PLT Taylor rules with two sets of parameters: (1) a weak rule as suggested in the adaptive learning literature, and (2) a much more harsh parametrization, which is required to stabilize the economy under pure naive expectations.

There are two main findings of our experiment. First, the strong PLT Taylor rule was required to stabilize the economy, and in terms of inflation stability can even outperform the classical inflation based Taylor rule. On the other hand, under the weak PLT Taylor rule the experimental sessions resulted in highly unstable dynamics. This suggest that our subjects were less sophisticated than what is required by both Rational Expectations and adaptive learning. Instead, the subjects followed much simpler forecasting rules. We find that a simple first-order heuristic – extrapolating observed trends from an anchor – can explain individual behavior well, in line with earlier literature on Learning-to-Forecast experiments.

A second main finding of our experiment is that the guidance had no visible effect on the behavior of the subjects. Estimations demonstrate that they would consider the additional information of the guidance much more often under the Weak PLT Taylor rule treatment. Nevertheless, this would not offset the under-reaction of the monetary authority to the deviations from the full employment steady state. This finding may have important implications on how the central banks should run their monetary policy, but requires further experimental and theoretical investigation.

Estimations of the individual behavior in our experiment lead to one additional interesting outcome. Subjects under the two Weak PLT Taylor rule treatments, with much more unstable dynamics, would use more complicated forecasting heuristics. We interpret this finding in the following way. The level of sophistication of individual behavior is not simply a constant that depends on individual characteristics. Instead, the subjects will increase the complexity of their behavior if they face a more complicated environment. This insight may have important consequences for economic models of learning, and should be studied more carefully in future work. In particular a complementary research should focus on whether our results can be replicated and explained by behavioral models of expectations formation, for example by Heuristic Switching Model [eg. 4] or models based on learning with Genetic Algorithms [eg. 25,5].

The design of our experiment can be easily extended to incorporate alternative monetary policies. One interesting question is whether a more involved form of guidance (such as a public commitment to a specific path of interest rate) can have a stabilizing macroeconomic effect, in particular whether a more involved communication of the central bank can anchor individual expectations. This would link our experiment with the works of Arifovic and Petersen [7], Mokhtarzadeh and Petersen [31] and Ahrens et al. [1].

Declarations of Competing Interest

None.

Appendix A. Experimental instructions and control questions

The first set of instructions was given to the subjects from the treatments **INF**, **StrongNo** and **WeakNo** (treatment without guidance). The second set of instructions was given to subjects from the treatments **StrongGuid** and **WeakGuid** (treatments with guidance).

All subjects had to answer control questions 1, 2 and 3. In addition, subjects from the treatments **StrongGuid** and **WeakGuid** (treatments with guidance) had to answer control question 4 in order to start the experiment.

Experimental instructions (treatments INF, StrongNo and WeakNo)

Welcome to this experiment! The experiment is anonymous, the data from your choices will only be linked to your station ID, not to your name. You will be paid privately at the end, after all participants have finished the experiment. After the main part of the experiment and before the payment you will be asked to fill out a short questionnaire. On your desk you will find a calculator and scratch paper, which you can use during the experiment.

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 at any time, please raise your hand and someone will come to your desk.

General information and experimental economy

All participants will be randomly divided into groups of a fixed size. The group composition will not change during the experiment. You and all other participants will take the roles of statistical research bureaus making predictions of inflation and the so-called "output gap". The experiment consists of 50 periods in total. In each period you will be asked to predict inflation and output gap for the next period.

The economy you are participating in is described by three variables: inflation π_t , output gap y_t and interest rate R_t . The subscript t indicates the period the experiment is in. In total there are 50 periods, so t increases during the experiment from 1 to 50.

Inflation (π_t) measures the percentage change in the price level of the economy. In each period, inflation depends on inflation predictions of the statistical research bureaus in the economy (that is on your own forecast as well as on the forecasts of the other bureaus in the experiment), on output gap, on interest rate and on a small random term. There is a **positive** relation between the actual inflation and (i) the inflation predictions and (ii) the actual output gap. This means that if the inflation predictions of the research bureaus or the actual output gap increase, then actual inflation will also increase (everything else equal). In economies similar to this one, inflation has historically been between -5% and 8% .

Output gap (y_t) represents the amount of goods produced by firms and consumed by households in the economy. In each period, output gap depends on inflation predictions and output gap predictions of the statistical research bureaus in the economy (that is on your own forecast as well as on the forecasts of the other bureaus in the experiment), on the interest rate and on a small random term. There is a **positive** relation between the actual output gap and both the inflation predictions and output gap predictions. This means that if the inflation predictions or output gap predictions of the research bureaus increase, then actual output gap will also increase (everything else equal). There is a **negative** relation between output gap and the interest rate. This means that if the interest rate increases, then actual output gap will instead decrease (everything else equal). In economies similar to this one, output gap has historically been between -5% and 8% .

Interest rate (R_t) measures the cost of borrowing money and is determined by the central bank. The central bank sets the interest rate in response to the inflation and the output gap. In each period, if inflation and output gap forecasts are considered too high, the central bank increases the interest rate. If inflation and output gap forecasts are considered too low, the central bank decreases the interest rate. The interest rate cannot take negative values.

Prediction task

Your task in each period of the experiment is to predict inflation and output gap in the next period. For example, in period 21 you have to predict inflation and output gap in period 22. When the experiment starts, you have to predict inflation and output gap for the first two periods, i.e. π_1^f and y_1^f and then π_2^f and y_2^f . The superscript f indicates that these are forecasts. When all participants have made their predictions for the second period, the actual inflation (π_1), the interest rate (R_1) and the actual output gap (y_1) for period 1 are announced. Then period 2 of the experiment begins. In period 2 you make inflation and output gap predictions for period 3 (π_3^f and y_3^f). When all participants have made their predictions for period 3, actual inflation (π_2), interest rate (R_2) and output gap (y_2) for period 2 are announced. This process repeats itself for 50 periods.

Thus, in a certain period t when you make predictions of inflation and output gap for the next period $t + 1$, the following information is available to you:

- realized values of inflation, interest rate and output gap, up to and including period $t - 1$;
- Your predictions up to and including your prediction for period t ;
- Your prediction scores up to and including period $t - 1$ (see below).

Payments

Your payment will depend on the accuracy of your predictions. You will be paid either for predicting inflation or for predicting output gap. The accuracy of your predictions is measured by the absolute distance between your prediction and the actual values (this distance is the prediction error). For each period the prediction error is calculated as soon as

the actual values are known; you subsequently get a prediction score that decreases as the prediction error increases. The table below gives the relation between the prediction error and the prediction score. The prediction error is calculated in the same way for inflation and output gap.

Example: If (for a certain period) you predict an inflation of 2%, and the actual inflation turns out to be 3%, then you make an absolute error of $3\% - 2\% = 1\%$. Therefore you get a prediction score of 50. If you predict an inflation of 1%, and the actual inflation turns out to be negative, for example -2% , you make a prediction error of $1\% - (-2\%) = 3\%$. Then you get a prediction score of 25. For a perfect prediction, with a prediction error of zero, you get a prediction score of 100.

The figure above shows the relation between your prediction score (vertical axis) and your prediction error (horizontal axis). Points in the graph correspond to the prediction scores in the previous table. At the end of the experiment, you will have two total scores, one for inflation predictions and one for output gap predictions. These total scores simply consist of the sum of all prediction scores you got during the experiment, separately for inflation and output gap predictions. **When the experiment has ended, one of the two total scores will be randomly selected for payment.**

Your final payment will consist of 0.75 euro for each 100 points in the selected total score (200 points therefore equals 1.50 euro). This will be the only payment from this experiment, i.e. you will not receive a show-up fee on top of it.

Computer interface

The computer interface will be mainly self-explanatory and example screenshot is presented below. The top part of the screen will tell you the current period, and how many decisions (forecasts) you still have to make in the current period (if you made all the forecasts, you will be asked to wait for other subjects). The right part of the screen will show you a table with all of the information available up to the period that you are in. That is, in period t , i.e. when you are asked to make your prediction for period $t + 1$, this will be actual inflation, interest rate and output gap until period $t - 1$, your predictions until period t , and the prediction scores arising from your predictions until period $t - 1$ for both inflation and output gap. The sum of the prediction scores over the different periods are shown in the bottom right of the screen, separately for your inflation and output gap predictions. Once the current period will become large, you may need to scroll down the table to see the early periods.

The left part of the screen will show you the information on inflation, interest rate and output gap in graphs. The axis of the inflation graph shows values in percentage points (i.e. 3 corresponds to 3%). Please note that maybe you need to scroll the graph box down to see bottom figures and the decision box for the output gap forecast (compare the last two figures).

In this panel you will also be asked to enter your predictions. When submitting your prediction, use a decimal point if necessary (not a comma). For example, if you want to submit a prediction of 2.5% type "2.5"; for a prediction of -1.75% type "-1.75". The order of the boxes in the panel is: box for inflation forecast, three graphs with inflation, output and interest rate information and box for output gap forecast.

Sample computer interface

Experimental instructions (treatments StrongGuid and WeakGuid)

Welcome to this experiment! The experiment is anonymous, the data from your choices will only be linked to your station ID, not to your name. You will be paid privately at the end, after all participants have finished the experiment. After the main part of the experiment and before the payment you will be asked to fill out a short questionnaire. On your desk you will find a calculator and scratch paper, which you can use during the experiment.

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 at any time, please raise your hand and someone will come to your desk.

General information and experimental economy

All participants will be randomly divided into groups of a fixed size. The group composition will not change during the experiment. You and all other participants will take the roles of statistical research bureaus making predictions of inflation and the so-called "output gap". The experiment consists of 50 periods in total. In each period you will be asked to predict inflation and output gap for the next period.

The economy you are participating in is described by three variables: inflation π_t , output gap y_t and interest rate R_t . The subscript t indicates the period the experiment is in. In total there are 50 periods, so t increases during the experiment from 1 to 50.

Inflation (π_t) measures the percentage change in the price level of the economy. In each period, inflation depends on inflation predictions of the statistical research bureaus in the economy (that is on your own forecast as well as on the forecasts of the other bureaus in the experiment), on output gap, on interest rate and on a small random term. There is a **positive** relation between the actual inflation and (i) the inflation predictions and (ii) the actual output gap. This

means that if the inflation predictions of the research bureaus or the actual output gap increase, then actual inflation will also increase (everything else equal). In economies similar to this one, inflation has historically been between -5% and 8% .

Output gap (y_t) represents the amount of goods produced by firms and consumed by households in the economy. In each period, output gap depends on inflation predictions and output gap predictions of the statistical research bureaus in the economy (that is on your own forecast as well as on the forecasts of the other bureaus in the experiment), on the interest rate and on a small random term. There is a **positive** relation between the actual output gap and both the inflation and output gap predictions. This means that if the inflation predictions or output gap predictions of the research bureaus increase, then actual output gap will also increase (everything else equal). There is a **negative** relation between output gap and the interest rate. This means that if the interest rate increases, then actual output gap will instead decrease (everything else equal). In economies similar to this one, output gap has historically been between -5% and 8% .

Interest rate (R_t) measures the cost of borrowing money and is determined by the central bank. The central bank sets the interest rate in response to the output gap and the relative deviation of the price level from its intended level (see below). In each period, if output gap forecasts are considered too high, the central bank increases the interest rate. If output gap forecasts are considered too low, the central bank decreases the interest rate. The interest rate cannot take negative values. Furthermore, throughout all 50 periods the central banks commits to the following inflation stabilizing policy.

The central banks wants to guide the actual inflation (price growth) in such a way that the actual **price level** P_t will not deviate from its intended path P_t^{int} , in which price growth (actual inflation) is neither too large nor too low. The intended price level, which the central bank desires for a specific period, can vary between periods. This intended price level is not known, but what is known is the *relative* deviation of the price level from this intended level. In each period, if price level forecasts are considered too high relative to the intended level, the central bank increases the interest rate. If price forecasts are considered too low relative to the intended level, the central bank decreases the interest rate. The interest rate cannot take negative values.

Prediction task

Your task in each period of the experiment is to predict inflation and output gap in the next period. For example, in period 21 you have to predict inflation and output gap in period 22. When the experiment starts, you have to predict inflation and output gap for the first two periods, i.e. π_1^f and y_1^f and then π_2^f and y_2^f . The superscript f indicates that these are forecasts. When all participants have made their predictions for the second period, the actual inflation (π_1), the interest rate (R_1), the actual output gap (y_1) and the relative deviation of the price level ($\frac{P_1 - P_1^{int}}{P_1^{int}}$) for period 1 are announced. Then period 2 of the experiment begins. In period 2 you make inflation and output gap predictions for period 3 (π_3^f and y_3^f). When all participants have made their predictions for period 3, inflation (π_2), interest rate (R_2), output gap (y_2) and $\frac{P_2 - P_2^{int}}{P_2^{int}}$ relative deviation of the price level for period 2 are announced. This process repeats itself for 50 periods.

Thus, in a certain period t when you make predictions of inflation and output gap for the next period $t + 1$, the following information is available to you:

- realized values of inflation, output gap and deviations of the price level from the one intended by the central bank, up to and including period $t - 1$;
- Your predictions up to and including your prediction for period t ;
- Your prediction scores up to and including period $t - 1$ (see below).

Payments

Your payment will depend on the accuracy of your predictions. You will be paid either for predicting inflation or for predicting output gap. The accuracy of your predictions is measured by the absolute distance between your prediction and the actual values (this distance is the prediction error). For each period the prediction error is calculated as soon as the actual values are known; you subsequently get a prediction score that decreases as the prediction error increases. The table below gives the relation between the prediction error and the prediction score. The prediction error is calculated in the same way for inflation and output gap.

Example: If (for a certain period) you predict an inflation of 2% , and the actual inflation turns out to be 3% , then you make an absolute error of $3 - 2 = 1\%$. Therefore you get a prediction score of 50. If you predict an inflation of 1% , and the actual inflation turns out to be negative, for example -2% , you make a prediction error of $1\% - (-2\%) = 3\%$. Then you get a prediction score of 25. For a perfect prediction, with a prediction error of zero, you get a prediction score of 100.

The figure above shows the relation between your prediction score (vertical axis) and your prediction error (horizontal axis). Points in the graph correspond to the prediction scores in the previous table. At the end of the experiment, you will have two total scores, one for inflation predictions and one for output gap predictions. These total scores simply consist of the sum of all prediction scores you got during the experiment, separately for inflation and output gap predictions. **When the experiment has ended, one of the two total scores will be randomly selected for payment.**

Your final payment will consist of 0.75 euro for each 100 points in the selected total score (200 points therefore equals 1.50 euro). This will be the only payment from this experiment, i.e. you will not receive a show-up fee on top of it.

Computer interface

The computer interface will be mainly self-explanatory and example screenshot is presented below. The top part of the screen will tell you the current period, and how many decisions (forecasts) you still have to make in the current period (if you made all the forecasts, you will be asked to wait for other subjects). The right part of the screen will show you a table with all of the information available up to the period that you are in. That is, in period t , i.e. when you are asked to make your prediction for period $t + 1$, this will be actual inflation, interest rate, output gap and deviation of price level from the level intended by the central bank until period $t - 1$, your predictions until period t , and the prediction scores arising from your predictions until period $t - 1$ for both inflation and output gap. The sum of the prediction scores over the different periods are shown in the bottom right of the screen, separately for your inflation and output gap predictions. Once the current period will become large, you may need to scroll down the table to see the early periods.

The left part of the screen will show you the information in graphs on inflation, output gap, interest rate and the relative deviation of the price level from its intended value. The vertical axis of the graph shows values in percentage points (i.e. 3 corresponds to 3%). Please note that maybe you need to scroll the graph box down to see bottom figures and the decision box for the output gap forecast (compare the last two figures).

In this panel you will also be asked to enter your predictions. When submitting your prediction, use a decimal point if necessary (not a comma). For example, if you want to submit a prediction of 2.5% type "2.5"; for a prediction of -1.75% type "-1.75". The order of the boxes in the panel is: box for inflation forecast, three graphs with inflation, output gap and interest rate/price deviation from the level intended by the central bank information;; and box for output forecast.

Sample computer interface

Control questions

Question 1

Suppose that the statistical bureaus predict that inflation will increase. Holding all other factors equal, including the interest rate, this means that:

- (a) the output gap will **increase**;
- (b) the output gap will **stay on the same level**;
- (c) the output gap will **decrease**;

Question 2

Suppose that the statistical bureaus predict that output gap in period decrease. Holding all other factors equal, this means that:

- (a) the central bank will **increase** the interest rate, which in turn has a **positive** impact on the output gap;
- (b) the central bank will **decrease** the interest rate, which in turn has a **negative** impact on the output gap;
- (c) the central bank will **increase** the interest rate, which in turn has a **negative** impact on the output gap.
- (d) the central bank will **decrease** the interest rate, which in turn has a **positive** impact on the output gap;

Question 3

Suppose that your inflation prediction for period 9 is -1%, and the realized inflation in that period is 3%. For this forecast you will receive score of

- (a) 10 points;
- (b) 100 points;
- (c) 20 points;
- (d) 33.33 points.

Question 4 – only for guidance treatment

Suppose that the prices fall below the intended level of the central bank. Holding all other factors equal, this means that:

- (a) the central bank will **increase** the interest rate.
- (b) the central bank will **decrease** the interest rate.
- (c) the central bank will **not change** the interest rate.
- (d) it is not possible to say what the central bank will do.

Solution:

1a, 2d, 3c, 4b

Appendix B. Experimental results (graphical representation)**B.1. Inflation targeting**

Fig. B.1.1. Group number 1 (Inflation Targeting).

Fig. B.1.2. Group number 2 (Inflation Targeting).

Fig. B.1.3. Group number 3 (Inflation Targeting). **Fig. B.1.4.** Group number 4 (Inflation Targeting). **Fig. B.1.5.** Group number 5 (Inflation Targeting).

Fig. B.1.6. Group number 6 (Inflation Targeting).

Fig. B.1.7. Group number 7* (Inflation Targeting). The group in which one of the subjects was extremely slow. The group was terminated after period 41 and not used in the econometric analysis.

B.2. PLT: Stable with Guidance

Fig. B.2.1. Group number 1 (PLT: Stable with Guidance).

Fig. B.2.2. Group number 2 (PLT: Stable with Guidance).

Fig. B.2.3. Group number 3 (PLT: Stable with Guidance).

Fig. B.2.4. Group number 4 (PLT: Stable with Guidance).

Fig. B.2.5. Group number 5 (PLT: Stable with Guidance).

Fig. B.2.6. Group number 6 (PLT: Stable with Guidance).

B.3. PLT: Unstable with Guidance

Fig. B.3.1. Group number 1 (PLT: Unstable with Guidance).

Fig. B.3.2. Group number 2 (PLT: Unstable with Guidance).

Fig. B.3.3. Group number 3 (PLT: Unstable with Guidance).

Fig. B.3.4. Group number 4 (PLT: Unstable with Guidance).

Fig. B.3.5. Group number 5 (PLT: Unstable with Guidance).

Fig. B.3.6. Group number 6 (PLT: Unstable with Guidance).

B.4. PLT: Stable with No Guidance

Fig. B.4.1. Group number 1 (PLT: Stable with No Guidance).

Fig. B.4.2. Group number 2 (PLT: Stable with No Guidance).

Fig. B.4.3. Group number 3 (PLT: Stable with No Guidance).

Fig. B.4.4. Group number 4 (PLT: Stable with No Guidance).

Fig. B.4.5. Group number 5 (PLT: Stable with No Guidance).

Fig. B.4.6. Group number 6 (PLT: Stable with No Guidance).

B.5. PLT: Unstable with No Guidance

Fig. B.5.1. Group number 1 (PLT: Unstable with No Guidance).

Fig. B.5.2. Group number 2 (PLT: Unstable with No Guidance).

Fig. B.5.3. Group number 3 (PLT: Unstable with No Guidance).

Fig. B.5.4. Group number 4 (PLT: Unstable with No Guidance).

Fig. B.5.5. Group number 5 (PLT: Unstable with No Guidance).

Fig. B.5.6. Group number 6 (PLT: Unstable with No Guidance).

Appendix C. Estimated individual behavior

For every subject, we independently estimate the two-dimensional forecasting rule (9) with the algorithm discussed below. All the estimations are based on a straightforward two-dimensional ML approach (with BFGS maximization algorithm), while tests are performed with LR test on 5% significance level. We wrote the econometrics code in matrix algebra language Ox [17]. The code is available on request.

Variable selection algorithm **Start** with all coefficients in the coefficient pool.

1. Test significance of each individual coefficient, which was not yet thrown out of the coefficient pool. If all are significant, **stop**. Otherwise, go to point 2.
2. Test the joined significance of all coefficients, which were found insignificant in the previous step. If test rejects their joint significance, throw them all out of the pool and go to point 1. Otherwise go to point 3.
3. From the coefficients, which were found insignificant in point 1, select exactly one to be thrown out of the pool according to this criterion:
 - If no coefficient was so far thrown out from the pool at point 3; or if the last coefficient that was thrown out of the pool at point 3 was a coefficient from the output gap rule, select a coefficient from the inflation rule;
 - Otherwise, select a coefficient from the output gap rule;
 - For the relevant rule, throw one coefficient out of the pool which was deemed insignificant in point 1, and which appears as first in the following enumeration (superscripts have been suppressed for the sake of brevity): α_3 , β , γ (only if one of the two guidance treatments), δ , α_2 , α_1 , constant.

Afterward, go back to point 1.

Remark that at no stage one throws out higher lags of the explained variables, as to make sure that there is no autocorrelation in the data. See also Table 4 for the average coefficients.

C.1. Treatment INF

Table 5
Estimated individual rules for Treatment 1 (Inflation forecasting).

Subject	Inflation forecasting rule						Output gap forecasting rule					
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	c^v	α_1^v	α_2^v	α_3^v	β^v	δ^v
Group 1												
11		0.553	0.43		0.995		2.41	0.305				−0.403
12		0.292				0.551	1.94	0.28			−0.601	−0.286
13			0.972		1.38				1.01		1.16	
14	0.528	0.874			1.24		1.06	−0.0197	0.693			−0.163
15	0.465	0.888			1.16		1.02	0.241	0.403			−0.159
16			0.994						1.02			
Mean	0.165	0.434	0.399	0	0.796	0.0918	1.07	0.134	0.521	0	0.0932	−0.168
(SD)	(0.235)	(0.368)	(0.44)	(0)	(0.575)	(0.205)	(0.9)	(0.142)	(0.423)	(0)	(0.525)	(0.145)
Group 2												
21	4.44				1.04		−0.958	0.201	0.528		1.23	0.205
22			1.01		1.01		−2.12		1.36		1.08	0.365
23	2.25	0.478			1.37		−0.098	0.264	0.532		1.29	0.0238
24	4.24				1.4		0.777	0.0913	0.00873		1.53	−0.102
25			0.997		0.642				0.964		0.742	
26	−0.898		−0.837			1.59	0.0146	0.00215	1.15		0.783	0.0255
Mean	1.67	0.0797	0.195	0	0.91	0.265	−0.397	0.0931	0.757	0	1.11	0.0863
(SD)	(2.11)	(0.178)	(0.644)	(0)	(0.479)	(0.594)	(0.92)	(0.105)	(0.451)	(0)	(0.278)	(0.154)
Group 3												
31			1.01						1.04			
32	0.335	0.929			1.26		1.24	0.182	0.486			−0.219
33			0.976				−0.113		0.99		1.03	
34			1.01				3.37					−0.558
35		0.995			1.18				1.11			
36			1						1.02			
Mean	0.0559	0.321	0.667	0	0.408	0	0.75	0.0304	0.774	0	0.172	−0.13
(SD)	(0.125)	(0.454)	(0.472)	(0)	(0.577)	(0)	(1.26)	(0.068)	(0.402)	(0)	(0.384)	(0.208)

(continued on next page)

Table 5 (continued)

Subject	Inflation forecasting rule						Output gap forecasting rule					
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	c^ν	α_1^ν	α_2^ν	α_3^ν	β^ν	δ^ν
Group 4												
41		1.01							0.95			
42		1.01						0.932				
43		1.01							0.907			
44		1.01						0.947				
45		1						0.977				
46		1.01						0.941				
Mean	0	1.01	0	0	0	0	0	0.633	0.309	0	0	0
(SD)	(0)	(0.00242)	(0)	(0)	(0)	(0)	(0)	(0.448)	(0.438)	(0)	(0)	(0)
Group 5												
51			1.03		1.08		6.29		−1.24	−1.08	0.814	
52			0.94		0.688			0.283	0.612		0.788	
53			0.992		1.63				0.849		2.01	
54			0.974		0.649		1.27		0.526		0.565	−0.159
55	1.59		0.998		1.63	−0.243	3.37	−0.0395	−0.422	−1.21	1.57	0.459
56			1		0.607				1.02		0.868	
Mean	0.265	0	0.989	0	1.05	−0.0405	1.82	0.0405	0.223	−0.383	1.1	0.05
(SD)	(0.592)	(0)	(0.0277)	(0)	(0.44)	(0.0905)	(2.33)	(0.109)	(0.798)	(0.543)	(0.511)	(0.192)
Group 6												
61		1.01							1.25			
62			0.997									
63	1.5	0.675			2.11		−1.06	0.417	0.899			0.179
64			0.998									
65	4.64						0.031		0.85			
66	1.37	0.701					−0.0455	0.349	0.9			
Mean	1.25	0.397	0.333	0	0.352	0	−0.178	0.128	0.65	0	0	0.0298
(SD)	(1.64)	(0.411)	(0.47)	(0)	(0.787)	(0)	(0.393)	(0.182)	(0.478)	(0)	(0)	(0.0667)

C.2. Treatment StrongNo

Table 6

Estimated individual rules for Treatment 2 (Stable price level targeting (no guidance)).

Subject	Inflation forecasting rule						Output gap forecasting rule					
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	c^ν	α_1^ν	α_2^ν	α_3^ν	β^ν	δ^ν
Group 1												
11			1.11			−0.0882	2.2		0.735			−0.286
12			1.18			−0.15			0.719			
13			0.992						0.944			
14			0.993						0.941			
15	1.03	0.796			1.17		1.58	0.209	0.736			−0.259
16			0.953				−8.72		1.28	0.699	1.02	0.847
Mean	0.171	0.133	0.87	0	0.195	−0.0397	−0.823	0.0349	0.892	0.117	0.17	0.0503
(SD)	(0.382)	(0.296)	(0.397)	(0)	(0.436)	(0.0589)	(3.64)	(0.0781)	(0.197)	(0.261)	(0.38)	(0.377)
Group 2												
21	3.35	0.308			2.5		0.0945	0.215	0.646		0.856	
22	4.96				0.89		7.62	0.136	1.08	−1.63	−0.0981	0.0959
23	6.06				1.03	−0.174	−3.53	−0.129	1.62	−0.84	1.76	1.26
24	2.43	0.818			1.2	−0.237	0.486	0.584	0.479		0.134	−0.0675
25		0.997						0.511	0.489			
26	5						0.452	0.569				

(continued on next page)

Table 6 (continued)

Subject	Inflation forecasting rule						Output gap forecasting rule					
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	c^ν	α_1^ν	α_2^ν	α_3^ν	β^ν	δ^ν
Mean	3.63	0.354	0	0	0.936	−0.0686	0.854	0.314	0.719	−0.411	0.443	0.214
(SD)	(2.01)	(0.41)	(0)	(0)	(0.844)	(0.0987)	(3.33)	(0.263)	(0.513)	(0.624)	(0.67)	(0.468)
Group 3												
31	5							−0.106	1.13			
32	5						0.105		1.05			
33			1									
34	1.72	0.171	0.473				3.49	0.812	0.0597			−0.562
35			0.999									
36	5.01						0.0569	0.541	0.234			
Mean	2.79	0.0285	0.412	0	0	0	0.608	0.208	0.413	0	0	−0.0937
(SD)	(2.29)	(0.0638)	(0.448)	(0)	(0)	(0)	(1.29)	(0.343)	(0.486)	(0)	(0)	(0.21)
Group 4												
41	2.16		0.565				2.25	0.223	0.534			−0.333
42			0.992						1.19			
43			1.01				0.411		0.66			
44	3.38	0.32			1.3		2.39	0.437	0.24			−0.343
45	3.11	0.383			2.72		4.44	0.37	0.307			−0.702
46	2.83	0.431					0.136	−0.0313	0.72			
Mean	1.91	0.189	0.428	0	0.671	0	1.61	0.167	0.608	0	0	−0.23
(SD)	(1.4)	(0.192)	(0.452)	(0)	(1.03)	(0)	(1.59)	(0.188)	(0.311)	(0)	(0)	(0.26)
Group 5												
51	1.48	0.347				0.293	−0.547	0.259	0.308			0.0807
52		0.63	0.382						0.267			
53	2.11	0.579			1.34		−0.104	−0.231	1.28			0.0364
54			1.01					0.357	0.71			
55			1.25		−1.21	−0.206	9.42		1.64	−2.52		0.53
56		−0.462	1.47					0.589				
Mean	0.599	0.182	0.685	0	0.021	0.0144	1.46	0.162	0.702	−0.419	0	0.108
(SD)	(0.867)	(0.38)	(0.587)	(0)	(0.736)	(0.145)	(3.57)	(0.27)	(0.585)	(0.937)	(0)	(0.191)
Group 6												
61	3.27	0.318			1.38		−3.24	0.0223	0.55	0.255	1.14	0.427
62	2.23	0.344			0.698	0.155	3	0.271	0.348		−0.0979	−0.435
63	1.8	0.612			0.908		−1.8	0.394	0.592		0.593	0.355
64			1.16		0.372	−0.176		0.398	0.401	0.253	0.662	−0.136
65			0.964						0.971		0.151	
66		0.359	0.923			−0.239	2.61		0.657			−0.386
Mean	1.22	0.272	0.507	0	0.559	−0.0433	0.0955	0.181	0.586	0.0847	0.408	−0.0291
(SD)	(1.29)	(0.215)	(0.513)	(0)	(0.495)	(0.13)	(2.22)	(0.179)	(0.202)	(0.12)	(0.432)	(0.332)

C.3. Treatment WeakNo

Table 7

Estimated individual rules for Treatment 3 (Unstable price level targeting (no guidance)).

Subject	Inflation forecasting rule						Output gap forecasting rule					
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	c^ν	α_1^ν	α_2^ν	α_3^ν	β^ν	δ^ν
Group 1												
11			0.361		0.855	0.427	−0.958	−0.109	0.993	−0.585	0.683	0.614
12			0.284		0.445	0.256		−0.0777	0.748	−0.647	0.359	0.591
13	2.37			0.315	0.648	0.262	1.93	0.25	0.394	−0.183	0.515	−0.0118
14			−0.695	0.747	0.5	0.907		−0.255	0.898	−0.715	0.647	0.585

(continued on next page)

Table 7 (continued)

Subject	Inflation forecasting rule						Output gap forecasting rule					
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	c^y	α_1^y	α_2^y	α_3^y	β^y	δ^y
15	5.75		−2	1.76		1.14	3.08	−0.103	1.38	−1.43	−0.0866	0.634
16	4.95	0.181	−0.247	0.471			3.11	0.0278	0.414	−0.186	0.0468	−0.203
Mean	2.18	0.0301	−0.383	0.549	0.408	0.498	1.19	−0.0446	0.804	−0.624	0.361	0.368
(SD)	(2.41)	(0.0674)	(0.805)	(0.602)	(0.316)	(0.396)	(1.59)	(0.155)	(0.341)	(0.417)	(0.291)	(0.341)
Group 2												
21	2.14	−0.295	0.172	0.855		0.513	1.53	−0.355	1.65	−0.927	0.204	0.66
22	4.81	−0.372		0.859	0.65	0.437	3.37	0.292	0.544	−0.264	0.461	−0.0281
23	2.24			0.645	0.348	0.623	2.01	0.0601	1.3	−0.825	0.289	0.466
24	4.69	0.544	−1.26	1.4		0.612	2.11	−0.0336	1.96	−1.95	−0.00253	1.17
25	2.16	0.694		0.226			3.57	0.627	0.291	−0.0239	−0.0589	−0.224
26	2.67	−0.368		0.615	1.09	0.611	2.43	−0.231	1.24	−0.62	0.675	0.368
Mean	3.12	0.0339	−0.181	0.767	0.348	0.466	2.5	0.06	1.16	−0.769	0.261	0.402
(SD)	(1.17)	(0.434)	(0.485)	(0.354)	(0.41)	(0.219)	(0.737)	(0.327)	(0.584)	(0.614)	(0.254)	(0.455)
Group 3												
31	2.12	0.532			0.564		2.1	0.604	0.2	−0.276	0.503	−0.0353
32	0.592	−0.4	1.03		0.56	0.251	−0.314	−0.374	1.57	−0.911	0.483	0.779
33	1.15	0.165	0.18		0.776	0.241	1.21	0.169	0.18	0.0631	0.856	0.182
34	0.745	0.473					0.103	0.341	0.618	−0.716	0.0799	0.486
35	1.32	0.586			0.728		1.12	0.361	0.53	−0.459	0.712	0.218
36		−0.179	0.687		0.683	0.559	−0.474		0.698	−0.396	0.553	0.78
Mean	0.987	0.196	0.316	0	0.552	0.175	0.623	0.184	0.633	−0.449	0.531	0.402
(SD)	(0.658)	(0.374)	(0.401)	(0)	(0.259)	(0.204)	(0.923)	(0.31)	(0.464)	(0.311)	(0.24)	(0.307)
Group 4												
41		0.32	0.47		0.587	0.118	−1.27	0.214	1.14	−0.953	0.367	0.761
42	0.66		0.589		0.268	0.258	−0.542	0.0292	1.46	−1.34	0.109	0.971
43		0.223	0.607				−0.419	0.334	0.801	−0.586	0.272	0.508
44			0.724		0.747	0.163	−1.3	−0.376	1.55	−1.05	0.418	1.02
45		0.26	0.522		0.748	0.142	0.57	0.793	0.134		0.317	−0.101
46		0.511	0.361		0.782			0.3	0.435		0.399	0.0399
Mean	0.11	0.219	0.546	0	0.522	0.114	−0.493	0.216	0.922	−0.655	0.314	0.534
(SD)	(0.246)	(0.179)	(0.114)	(0)	(0.292)	(0.0912)	(0.662)	(0.351)	(0.518)	(0.513)	(0.104)	(0.434)
Group 5												
51	5.5	0.352	−1.82	1.61	0.476	0.85	2.3	0.268	0.779	−0.578	0.6	0.159
52			−0.584	0.522	0.524	0.79		−0.141	0.371		0.417	0.088
53	6.69	−0.301	0.887		0.575	−0.459	4.45	−0.128	0.513	−0.0205	0.737	−0.247
54	3.58	0.229	−1.81	1.73	0.253	1.23	2.39	0.31	0.607	−0.531	0.37	0.198
55	1.31		0.759		0.347		0.605	0.0133	0.765	−0.137	0.3	0.177
56	4.55	0.251			0.45		2.78	−0.0786	0.595	−0.321	0.312	0.098
Mean	3.6	0.0884	−0.428	0.645	0.437	0.401	2.09	0.0405	0.605	−0.265	0.456	0.0789
(SD)	(2.32)	(0.217)	(1.1)	(0.752)	(0.108)	(0.591)	(1.46)	(0.183)	(0.141)	(0.23)	(0.16)	(0.151)
Group 6												
61	1.76	0.586			0.892		1.53	0.278	0.519	−0.0833	0.854	−0.135
62	0.58		0.802		0.414		−0.00439	0.0818	0.624		0.486	0.18
63	1.54		0.575		0.905		0.949	−0.224	1.19	−0.543	1.08	0.338
64	1.34		0.741		0.738		1.75	−0.086	1.15	−0.225	0.541	−0.0832
65	0.811	0.731			0.31		1.14	0.659	0.218		0.228	−0.184
66	1.66		0.543		1.01		1.4	0.201	0.383	0.215	0.977	−0.266
Mean	1.28	0.219	0.443	0	0.711	0	1.13	0.152	0.68	−0.106	0.694	−0.0252
(SD)	(0.44)	(0.313)	(0.326)	(0)	(0.261)	(0)	(0.568)	(0.282)	(0.367)	(0.235)	(0.299)	(0.213)

C.4. Treatment StrongGuid

Table 8
Estimated individual rules for Treatment 4 (Stable price level targeting with guidance).

Subject	Inflation forecasting rule							Output gap forecasting rule						
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	γ^π	c^ν	α_1^ν	α_2^ν	α_3^ν	β^ν	δ^ν	γ^ν
Group 1														
11	–1.61		1.3				–0.44	4.53	0.5	–1.36		–0.618	–0.704	–2.63
12	0.973		0.798					–1.17	0.399	0.236		1.26	0.247	
13	4.93				0.936			8.71	0.119	0.933	–1.77	–0.179	0.032	0.497
14	–4.14		1.76				–1	5.23	0.475	–1.65		–0.365	–0.84	–2.89
15			0.977					2.26		–0.728			–0.377	–2.17
16		0.626	0.355				–0.365		–0.268					–2.06
Mean	0.0255	0.104	0.865	0	0.156	0	–0.301	3.26	0.204	–0.428	–0.294	0.0158	–0.274	–1.54
(SD)	(2.74)	(0.233)	(0.58)	(0)	(0.349)	(0)	(0.361)	(3.33)	(0.28)	(0.906)	(0.658)	(0.595)	(0.399)	(1.3)
Group 2														
21	2.07	0.507					–0.632	2.41	0.329	–0.665		0.041	–0.268	–1.4
22	4.28							0.609	0.577					
23		0.917					–0.684		0.901					
24			0.876				–0.574		0.467	0.276		0.61	0.0544	
25		0.314	0.663		0.874			–0.765		0.265			0.126	–1.61
26		0.833					–1.24		0.626					–0.821
Mean	1.06	0.429	0.257	0	0.146	0	–0.521	0.376	0.483	–0.0207	0	0.109	–0.0147	–0.638
(SD)	(1.63)	(0.363)	(0.368)	(0)	(0.326)	(0)	(0.427)	(0.994)	(0.277)	(0.312)	(0)	(0.225)	(0.122)	(0.68)
Group 3														
31			1.01					0.568	0.229	0.48			–0.0745	0.379
32	6.89	–0.213			1.08	–0.138	–0.986	1.35	–0.105	1.12		–0.446	–0.18	1.25
33	6.83				1.28	–0.29		–12.4	–0.293	2.39	1.41	1.05	0.909	0.387
34	2.63	0.499						0.0161	0.491	0.477				
35		0.995			0.573			–2.51		1.12			0.426	
36 [†]	4.95				1.08				0.0954	–0.769		1.36	0.00826	–2.78
Mean	3.55	0.214	0.168	0	0.669	–0.0715	–0.164	–2.16	0.0695	0.802	0.235	0.328	0.181	–0.126
(SD)	(2.88)	(0.411)	(0.375)	(0)	(0.519)	(0.11)	(0.367)	(4.71)	(0.249)	(0.947)	(0.526)	(0.648)	(0.376)	(1.26)
Group 4														
41		0.241	0.644		0.767	0.0835				0.551				–1.2
42	2.88	0.358			1.43			9.54		3.84	–3.13	0.886	1.07	5.35
43*		0.477				0.358	–0.332		–0.271	–2.25	1.57		–1.1	–5.12
44	–5		3.13	–3.48	0.603	–0.904	–5.99	13	0.244	5.35	–4.19	0.628	1.44	8.94
45	4.08				1.18	0.137		2.77	0.429	–0.546	0.426	–0.0214	–0.677	–1.27
46	1.43	–0.381	0.977		0.47	0.0886		3.88	0.0909	1.73	–0.796	0.507	0.0931	2.01
Mean	0.565	0.116	0.792	–0.58	0.74	–0.0395	–1.05	4.86	0.082	1.45	–1.02	0.333	0.138	1.45
(SD)	(2.89)	(0.282)	(1.11)	(1.3)	(0.466)	(0.402)	(2.21)	(4.83)	(0.218)	(2.56)	(2.01)	(0.358)	(0.893)	(4.64)
Group 5														
51		0.629	0.365							0.828				
52			0.976							0.844				
53	4.94				1.5			–0.802	0.251	–0.923	0.592	1.63	–0.355	–2.83
54			1.01							0.828		0.25		
55 [†]	3.66					0.195			–0.267	1.91	–1.07	0.534	0.855	0.961
56	1.27	0.303	0.429				–0.479	0.293	0.212	–0.0945		0.665		–1.12
Mean	1.64	0.155	0.464	0	0.251	0.0325	–0.0798	–0.0849	0.0325	0.565	–0.0797	0.514	0.0834	–0.498
(SD)	(1.97)	(0.239)	(0.409)	(0)	(0.561)	(0.0727)	(0.178)	(0.338)	(0.17)	(0.882)	(0.493)	(0.559)	(0.369)	(1.2)
Group 6														
61	5							0.0388	0.0711	0.692				
62			0.996							1				
63			0.994						0.969					
64		0.986					–0.815	0.147	0.402	0.755				
65			0.989				–0.65	1.9	0.35	0.406			–0.277	
66	5.12	0.547			–0.998	–0.427		2.48	0.379	–0.648		–0.00864	–0.37	–1.82
Mean	1.69	0.256	0.496	0	–0.166	–0.0712	–0.244	0.762	0.362	0.368	0	–0.00144	–0.108	–0.304
(SD)	(2.39)	(0.383)	(0.496)	(0)	(0.372)	(0.159)	(0.349)	(1.03)	(0.313)	(0.552)	(0)	(0.00322)	(0.155)	(0.68)

C.5. Treatment WeakGuid

Table 9
Estimated individual rules for Treatment 5 (Unstable price level targeting with guidance).

Subject	Inflation forecasting rule							Output gap forecasting rule						
	c^π	α_1^π	α_2^π	α_3^π	β^π	δ^π	γ^π	c^v	α_1^v	α_2^v	α_3^v	β^v	δ^v	γ^v
Group 1														
11*		0.487	0.457			0.589	−0.0781	−5.73				0.554	0.937	−0.258
12* [†]		0.491	0.384	−0.308	0.258	0.39	−0.203		0.405			0.493	0.214	−0.158
13* [†]		−0.478	1.31	−0.28	0.592	0.181				0.517		1.02	0.232	
14 [†]	13.7	0.258	−1.95	3.09	0.456		0.674		−0.000506	1.19	−0.81	0.56	0.716	0.000803
15 [†]		−0.186	0.698		0.313	0.196	−0.0579		−0.213	0.937	−0.565	0.549	0.535	
16		0.501	0.189		−0.109	0.266	−0.19	−0.852	0.444	−0.156			0.404	−0.234
Mean	2.28	0.179	0.182	0.417	0.252	0.27	0.0243	−1.1	0.106	0.415	−0.229	0.529	0.506	−0.108
(SD)	(5.11)	(0.38)	(1.02)	(1.2)	(0.243)	(0.184)	(0.299)	(2.09)	(0.238)	(0.509)	(0.332)	(0.295)	(0.259)	(0.113)
Group 2														
21* [†]		0.106	0.906	−0.2	1.15	−0.0797			0.448	0.432	−0.157	0.356	0.0823	
22	2.56	0.639	0.205				−0.0583	−11.1	0.102	−2.21	1.31	−0.144	1.19	−0.75
23* [†]		−0.368	1.59		0.758	−0.696	0.337		0.462	2.28	−1.09			0.486
24	1.3		0.765		1.07			7.72	0.374	2.21	−1.28	0.441	−0.364	0.402
25 [†]	1.65		0.919		0.461		−0.0334		0.183	−0.329	−0.526	−0.0641	1.2	−0.474
26*		0.544	0.876	−1.28	−0.794	0.551	−0.401	2.83	0.373	0.801			−0.403	0.264
Mean	0.917	0.154	0.877	−0.246	0.441	−0.0374	−0.026	−0.089	0.323	0.53	−0.291	0.0982	0.285	−0.0121
(SD)	(0.99)	(0.344)	(0.402)	(0.466)	(0.673)	(0.362)	(0.214)	(5.63)	(0.135)	(1.54)	(0.852)	(0.219)	(0.668)	(0.457)
Group 3														
31* [†]			1.45	−0.969	1.08	−0.276	−0.0893				0.67	0.961	−0.193	−0.0591
32		−0.197	0.497		0.644	0.546		−14	−0.179	−2.83	2.3	0.524	0.835	−0.586
33 [†]	7.48	0.484	0.721		0.564	−0.872	0.114		0.09	−0.981	1.05	0.96	−0.474	−0.137
34* [†]		0.324			0.362	0.447	−0.0799		0.166	0.71	−0.743	0.347	0.66	−0.0511
35	0.863	0.514	0.254		0.398			3.94	0.525	0.499		0.353	−0.435	0.102
36	−13.9	0.29	3.92	−4.69	0.723		−0.862	−1.51	0.491	−1.52	1.56	0.647	−0.595	−0.198
Mean	−0.929	0.236	1.14	−0.944	0.628	−0.0257	−0.153	−1.92	0.182	−0.687	0.805	0.632	−0.0338	−0.155
(SD)	(6.39)	(0.256)	(1.32)	(1.71)	(0.238)	(0.471)	(0.324)	(5.63)	(0.254)	(1.24)	(0.993)	(0.254)	(0.567)	(0.214)
Group 4														
41	−1.26		0.522		0.299	0.478	−0.0491	−2.24	−0.119	1.32	−1.07	0.217	1.12	−0.0599
42			0.671		0.556	0.223		−0.488		1.15	−0.727	0.37	0.663	
43		0.647	0.192		0.462		−0.0378	−1.4	0.451		0.331	0.411		−0.0669
44	−2.3	0.577			0.659	0.417	−0.0788	−0.286	0.64	−0.3	0.632	0.266	−0.378	−0.0637
45	1.07		0.76		0.412			1.15	−0.184	1.13	−0.346		0.275	0.011
46	0.675		0.784		0.64			−0.301		1.47	−0.819	0.482	0.636	
Mean	−0.303	0.204	0.488	0	0.505	0.186	−0.0276	−0.542	0.0812	0.797	−0.333	0.291	0.386	−0.0299
(SD)	(1.15)	(0.289)	(0.295)	(0)	(0.127)	(0.202)	(0.0302)	(1.07)	(0.344)	(0.685)	(0.62)	(0.157)	(0.487)	(0.0338)
Group 5														
51	−0.679	0.275		0.391	0.598	0.437		−0.81	0.118	1.22	−0.885	0.3	0.71	
52	−1.35		0.728		0.706	0.222	−0.0283	−1.85	−0.164	1.55	−1.11	0.233	0.91	−0.00805
53			0.589		0.469	0.315	−0.0107	−1.61		0.73	−0.358	0.288	0.651	−0.0293
54	1.07	0.241	1.25	−0.653	0.69	−0.424		0.956	0.182	−0.796	1.46	0.463	−0.658	−0.0348
55		−0.285	1.01		0.326	0.156				0.651		0.312	0.102	
56		0.263	0.531		0.676	0.0858	−0.0126	−0.51		1.16	−0.732	0.53	0.574	
Mean	−0.16	0.0822	0.684	−0.0437	0.577	0.132	−0.0086	−0.638	0.0227	0.753	−0.27	0.354	0.382	−0.012
(SD)	(0.738)	(0.202)	(0.393)	(0.307)	(0.138)	(0.273)	(0.0102)	(0.951)	(0.109)	(0.756)	(0.853)	(0.105)	(0.525)	(0.0145)
Group 6														
61	1.55		0.517		0.862	0.151	−0.0771	4.23	−0.0461	1.44	−0.696	1.05	0.173	0.0849
62			0.213		0.804	0.524	−0.0684	5.36		1.61	−0.873	0.742		0.202
63	1.32		0.639		0.651			−6.28	0.0755	−0.745	0.777	0.877	0.563	−0.3
64	2.3	0.287	0.327		0.826		−0.0372	6.91	0.744	0.436	0.181	0.656	−0.895	0.149
65	4.96			0.878	0.552		0.2	7.15	−0.222	2.52	−1.42	0.26	0.132	0.272
66	1.19		0.834		0.351			−3.17	−0.255	0.708	−0.322	0.388	0.85	−0.141
Mean	1.89	0.0478	0.422	0.146	0.674	0.112	0.00291	2.37	0.0494	0.994	−0.392	0.661	0.137	0.0442
(SD)	(1.53)	(0.107)	(0.276)	(0.327)	(0.18)	(0.192)	(0.0931)	(5.18)	(0.332)	(1.03)	(0.717)	(0.27)	(0.544)	(0.201)

Fig. C.1. Distribution of estimated individual inflation forecasting rules. Stars denote individual rules (if significant), diamonds indicate group averages and squares denote average in treatment.

Fig. C.2. Distribution of estimated individual output gap forecasting rules. Stars denote individual rules (if significant), diamonds indicate group averages and squares denote average in treatment.

References

- Ahrens, S., Lustenhouwer, J., Tettamanzi, M., 2017. The stabilizing role of forward guidance: a macro experiment.
- Amano, R., Engle-Warnick, J., Shukayev, M., 2011. Price-level Targeting and Inflation Expectations: Experimental Evidence. Bank of Canada working paper.
- Anufriev, M., Assenza, T., Hommes, C., Massaro, D., 2013. Interest rate rules and macroeconomic stability under heterogeneous expectations. *Macroecon. Dyn.* 17 (08), 1574–1604.
- Anufriev, M., Hommes, C., 2012. Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *Am. Econ. J.* 4 (4), 35–64.
- Anufriev, M., Hommes, C., Makarewicz, T., 2019. Simple forecasting heuristics that make us smart: Evidence from different market experiments. *J. Eur. Econ. Assoc.* 17 (5), 1538–1584.
- Arifovic, J., Duffy, J., 2018. Heterogeneous agent modeling: experimental evidence. In: *Handbook of Computational Economics*, 4. Elsevier, pp. 491–540.
- Arifovic, J., Petersen, L., 2017. Stabilizing expectations at the zero lower bound: experimental evidence. *J. Econ. Dyn. Control* 82, 21–43.
- Assenza, T., Heemeijer, P., Hommes, C., Massaro, D., 2020. Managing self-organization of expectations through monetary policy: a macro experiment. *J. Monet. Econ.*
- Bao, T., Hommes, C., Makarewicz, T., 2017. Bubble formation and (in) efficient markets in learning-to-forecast and optimise experiments. *Econ. J.* 127 (605), F581–F609.
- Bech, M.L., Gambacorta, L., Kharroubi, E., 2014. Monetary policy in a downturn: are financial crises special? *Int. Financ.* 17 (1), 99–119.
- Boivin, J., Kiley, M.T., Mishkin, F.S., 2010. How has the monetary transmission mechanism evolved over time? In: *Handbook of Monetary Economics*, 3. Elsevier, pp. 369–422.
- Chen, Q., Filardo, A., He, D., Zhu, F., 2016. Financial crisis, us unconventional monetary policy and international spillovers. *J. Int. Money Financ.* 67, 62–81.

- Chung, H., Herbst, E., Kiley, M.T., 2015. Effective monetary policy strategies in new Keynesian models: a reexamination. *NBER Macroecon. Annu.* 29 (1), 289–344.
- Cogley, T., Sargent, T.J., 2005. Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Rev. Econ. Dyn.* 8 (2), 262–302.
- Cornea-Madeira, A., Hommes, C., Massaro, D., 2019. Behavioral heterogeneity in US inflation dynamics. *J. Bus. Econ. Stat.* 37 (2), 288–300.
- Covas, F., Zhang, Y., 2010. Price-level versus inflation targeting with financial market imperfections. *Can. J. Econ.* 43 (4), 1302–1332.
- Doornik, J., 2007. *Object-oriented Matrix Programming Using Ox*, third ed. Timberlake Consultants Press, London.
- Evans, C., 2012. Monetary policy in a low-inflation environment: developing a state-contingent price-level target. *J. Money Credit Bank.* 44, 147–155.
- Femia, K., Friedman, S., Sack, B.P., 2013. The Effects of Policy Guidance on Perceptions of the FED's Reaction Function. FRB of New York Staff Report.
- Gali, J., Gambetti, L., 2009. On the sources of the great moderation. *Am. Econ. J.* 1 (1), 26–57.
- Giannone, D., Lenza, M., Reichlin, L., 2008. Explaining the great moderation: it is not the shocks. *J. Eur. Econ. Assoc.* 6 (2–3), 621–633.
- Giannoni, M.P., 2010. Optimal Interest-rate Rules in a Forward-looking Model, and Inflation Stabilization Versus Price-level Stabilization. Working paper no. w15986, National Bureau of Economic Research.
- Hommes, C., 2011. The heterogeneous expectations hypothesis: some evidence from the lab. *J. Econ. Dyn. Control* 35 (1), 1–24. doi:10.1016/j.jedc.2010.10.003.
- Hommes, C., 2020. Behavioral and experimental macroeconomics and policy analysis: a complex systems approach. *J. Econ. Lit.*
- Hommes, C., Makarewicz, T., Massaro, D., Smits, T., 2017. Genetic algorithm learning in a new Keynesian macroeconomic setup. *J. Evol. Econ.* 27 (5), 1133–1155.
- Hommes, C., Massaro, D., Salle, I., 2019. Monetary and fiscal policy design at the zero lower bound: Evidence from the lab. *Econ. Inq.* 57 (2), 1120–1140.
- Honkapohja, S., Mitra, K., 2014. Targeting nominal GDP or prices: guidance and expectation dynamics. *Bank of Finland*.
- Huber, J., Kirchler, M., Stöckl, T., 2010. The hot hand belief and the gamblers fallacy in investment decisions under risk. *Theory Decis.* 68 (4), 445–462.
- Krugman, P., 2014. Four observations on secular stagnation. *Secular Stagnation* 61.
- Magill, M., Quinzii, M., 2014. Term structure and forward guidance as instruments of monetary policy. *Econ. Theory* 56 (1), 1–32.
- Mokhtarzadeh, F., Petersen, L., 2017. Coordinating Expectations Through Central Bank Projections. Simon Fraser University Department of Economics Working Paper 17-03.
- Pfajfar, D., Žakelj, B., 2014. Experimental evidence on inflation expectation formation. *J. Econ. Dyn. Control* 44, 147–168.
- Pfajfar, D., Žakelj, B., 2016. Inflation expectations and monetary policy design: evidence from the laboratory. *Macroecon. Dyn.* 1, 1–41.
- Reifschneider, D.L., Roberts, J.M., 2006. Expectations formation and the effectiveness of strategies for limiting the consequences of the zero bound. *J. Jpn. Int. Econ.* 20 (3), 314–337.
- Stöckl, T., Huber, J., Kirchler, M., 2010. Bubble measures in experimental asset markets. *Exp. Econ.* 13 (3), 284–298.
- Summers, L.H., 2014. US economic prospects: secular stagnation, hysteresis, and the zero lower bound. *Bus. Econ.* 49 (2), 65–73.
- Woodford, M., 2012. Inflation targeting and financial stability. National Bureau of Economic Research.
- Yellen, J.L., 2013. Challenges confronting monetary policy. In: *Speech at the 2013 National Association for Business Economics Policy Conference*, Washington, DC, March 4, 2013.