The Stabilizing Role of Forward Guidance: A Macro Experiment

Steffen Ahrens* Joep Lustenhouwer† Michele Tettamanzi ‡

May 12, 2017

Abstract

Expectations are among the main driving forces for economic dynamics. Therefore, managing expectations has become a primary objective for monetary policy seeking to stabilize the business cycle. In this paper, we study if central banks can manage market expectations by means of forward guidance in a New Keynesian learning-to-forecast experiment. Forward guidance takes the form of one-period ahead inflation projections that are published by the central bank in each period. Subjects in the experiment observe these projections along with the historic development of the economy and subsequently submit their own one-period ahead inflation forecasts. In this context, we find that the central bank can significantly manage market expectations through forward guidance and that this management strongly supports monetary policy in stabilizing the economy. Moreover, forward guidance drastically reduces the probability of a deflationary spiral after strong negative shocks to the economy.

JEL classification: C92, E32, E37, E58.

Keywords: learning-to-forecast experiment, forward guidance, heterogeneous expectations.

*Technische Universität Berlin, Strasse des 17. Juni 135, 10623 Berlin, Germany. Email: steffen.ahrens@tu-berlin.de

†Universiteit van Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. Email: J.E.Lustenhouwer@uva.nl

‡Università Cattolica del Sacro Cuore Milano, Via Necchi 5, 20123 Milano, Italy. Email: michele.tettamanzi@unicatt.it

We thank Tiziana Assenza, Frank Heinemann, Cars Hommes, Paul Hubert, Domenico Massaro, Julián A. Parra-Polanía, and seminar participants of the CREED Seminar at Universiteit van Amsterdam, the DEFAP PhD Seminar at Univeristà Cattolica di Milano, the Macroeconomics Seminar at University of Illinois at Urbana-Champaign, and the Nuremberg Reserach Seminar in Economics at Friedrich-Alexander-Universität Erlangen-Nürnberg for valuable comments. We also thank participants of the 2015 Barcelona LeeX Experimental Economics Summer School in Macroeconomics at Universitat Pompeu Fabra, the 21st WEHIA Annual Workshop in Castellón de la Plana, and the EU FP7 project "Integrated Macro-Financial Modeling for Robust Policy Design" (MACFINROBODS) for fruitful discussions.

The authors gratefully acknowledge the financial support of NWO (Dutch Science Foundation) Project No. 40614011 "Monetary and Fiscal Policy under Bounded Rationality and Heterogeneous Expectations" and the DFG (German Research Foundation) through CRC 649 "Economic Risk" and CRC TRR 190 "Rationality and Competition."

1 Introduction

In economics, expectations can be seen as one of the main driving forces for economic dynamics. As Hommes (2011) puts it, "individual expectations about future aggregate outcomes are the key feature that distinguishes social sciences and economics from the natural sciences. Daily weather forecasts, either by the public or by experts, do not affect the probability of rain." Therefore, managing market expectations has become a primary objective for economic policy makers seeking to actively influence the economic development.[1]

For monetary policy makers, market expectations not only determine the effectiveness of their main conventional monetary policy instrument (i.e., the short-term nominal interest rate) in normal times; they are also central to the transmission of unconventional monetary policy (e.g. quantitative easing and forward guidance) when the short-term nominal interest rate is restricted by the zero lower bound (as we currently witness in many of the leading industrialized economies). So it has come that central banks worldwide have become increasingly communicative, providing the public with detailed information about their views of monetary policy and the fundamental factors driving their monetary policy decisions (Blinder et al., 2008).

A pivotal aspect in this regard is the central bank practice to publish inflation projections. This practice, which qualifies as a tool of forward guidance,[2] intends to provide superior information about future macroeconomic developments to the private sector and thereby to reduce private-sector uncertainty (Campbell et al., 2012). But central banks may also use this tool to strategically influence private-sector expectations by intentionally over- or underreporting the projected level of inflation (Gomez-Barrero and Parra-Polania, 2014; Charemza and Ladley, 2016; Jensen, 2016). Independent of the central banks' motive to publish inflation projections, ample empirical evidence reveals that this practice considerably impacts on private-sector expectations (Hubert, 2014, 2015a,b).

While the publication of central bank inflation projections might be a powerful tool for private-sector expectations management, the central bank must consider its effects on the endogenous credibility of future central bank inflation projection[3] (Blinder, 2000). Publishing accurate inflation projections strengthens the central bank's reputation as a good, accurate forecaster but deters the central bank from the ability to steer private-sector expectations into a different direction if preferred. Conversely, strategic inflation projections might allow the central bank to steer private-sector expectations, but may be damaging to credibility if they systematically over- or underestimate inflation. Thus, the

[1] See, for instance, the speech held by Janet Yellen at "The Elusive 'Great' Recovery: Causes and Implications for Future Business Cycle Dynamics" (60th annual economic conference) on October 14, 2016.

[2] There is no official definition of forward guidance. Commonly, forward guidance is understood as a (commitment to a) projected path of the nominal interest rate. This understanding of forward guidance is, however, very restrictive. In this paper, the term "forward guidance" refers to the rather vague concept of Delphic forward guidance as defined by Campbell et al. (2012). According to Campbell et al. (2012), (Delphic) forward guidance publicly states a forecast of macroeconomic fundamentals and the likely future course of monetary policy. As we show below, in our experiment, this form of forward guidance can be approximated by public central bank inflation projections, as they convey information also about the expected future interest rate.

[3] Throughout this paper, the term credibility refers exclusively to the central bank's inflation projections, and not to the central bank as the monetary authority.

central bank faces the risk of diminishing its ability to strategically influence private-sector expectations in the future. This trade-off between short term gains and potential long term losses raises the question how the central bank's ability to manage expectations via publishing inflation projections depends on the credibility of the central bank forecasts and how in turn credibility depends on the central bank's past forecasting performance.

In this paper, we study to what extent the central bank can influence economic activity through market-expectations management via the publication of (strategic) official central bank inflation projections[4], i.e., by using forward guidance, both in normal times and in times of severe economic stress (i.e. periods where there is a high probability of the zero lower bound on the nominal interest rate becoming binding), and how this influence depends on the endogenous degree of credibility given to the central bank projections by the private sector.

The analysis is conducted by means of a laboratory experiment. For the question at hand, a laboratory experiment has several advantages over traditional empirical or theoretical approaches.[5] First, it allows us to study the expectation formation process of the subjects and its interaction with monetary policy design, without having to rely on prescribed expectations formation processes, as e.g., rational or adaptive expectations. Second, we are able - in a very natural way - to depart from the representative agent hypothesis commonly put forth in macroeconomics and to allow for substantial heterogeneity. Finally, we can control the subjects' incentives and information sets in the laboratory.

The underlying economic environment of the experiment is given by a standard forward-looking New Keynesian model. The experimental task for the subjects is a learning-to-forecast experiment as pioneered by Marimon and Sunder (1993). All but one subject are "professional forecasters" in the private sector who are asked repeatedly to form one-period ahead expectations about future inflation, having only a limited understanding of the true data generating process. The remaining subject is assigned the role of the "central bank forecaster." Apart from the control treatment, at the beginning of each period the central bank publishes an official one-period ahead central bank inflation projection. Depending on the treatment, this public central bank inflation projection is produced either by the central bank forecaster or by a computerized algorithm. In any case, the central bank is provided with superior information that can be used in the forecasting process. Professional forecasters are presented with the public central bank projection before they submit their own inflation forecasts.

[4]The focus on the publication of inflation projections rather than interest rate projections is motivated by the work of Ferrero and Secchi (2010), who study the effect of different central bank communication strategies in a standard New Keynesian model when agents are learning. They find that the communication of interest rate projections can be destabilizing, while the communication of inflation projections is stabilizing. Although, the model attributes a stabilizing role also to output gap projections, we choose to abstract from output gap projections entirely based on institutional and empirical grounds. Institutionally, it is inflation stabilization which has traditionally been the core mandate of many central banks. Empirically, the relationship between output gap predictions and private-sector expectations is rather vague. E.g., in the United States, the FOMC's central bank output gap projections neither have an informational advantage over private-sector output gap forecasts (Romer and Romer, 2000), nor do they significantly influence private-sector output gap expectations (Hubert, 2014).

[5]For a thorough discussion about the potential advantages of laboratory experiments for the conduct of monetary policy analysis, see Cornand and Heinemann (2014).

The novelty of the proposed experiment is that, through a series of different treatments, we can study the impact of (strategic) forward guidance on the subjects' expectation formation process and the resulting dynamic evolution of the underlying theoretical economy. We are mainly interested in answering the questions: if a central bank can influence or even manage private-sector expectations via the publication of (strategic) central bank inflation projections in such a laboratory environment, if such expectations management can successfully be applied as an additional monetary policy instrument to stabilize the economy in normal times and in times of severe economic stress, and how effectiveness of forward guidance depends on the endogenous degree of central bank credibility.

For normal times, we find that the publication of inflation projections strongly affects private-sector expectations. Instead of simply following trends, under forward guidance subjects put a large weight on the public inflation projection when forming their expectations about future inflation. We show that the macroeconomic consequences of this influence depend on the quality of the published projection. Reasonable, informative public projections act as focal points which decrease the dispersion among individual professional forecasters and increases their forecasting performance. They stabilize the economy, i.e., they unambiguously mitigate the mean squared errors of the fundamentals by bringing the economies faster and closer towards the steady state and by reducing the volatility of inflation, output, and the interest rate.

In times of severe economic stress, the publication of optimistic central bank projections reduces the risk of deflationary spirals. Noisy inflation projections, by contrast, are generally harmful to the economy as they unleash disturbing forces which give rise to more dispersed and less precise individual private-sector forecasts. Finally, credibility of central bank inflation projections, at least in times of severe economic stress, seems to be of minor importance for the stabilizing role of forward guidance.

The paper is organized as follows. Section 2 reviews the relevant literature. Section 3 describes our experimental design. In Section 4 we study the influence of forward guidance on economic stability and the dynamic performance of inflation expectations. Section 5 analyzes the expectation formation processes of the subjects, and Section 6 briefly discusses central bank projection credibility and its interaction with forward guidance. Finally, Section 7 concludes.

2 Related Literature

Laboratory experiments on monetary policy have become increasingly popular in recent years (see Cornand and Heinemann (2014) for a survey). A considerable fraction of this newly developed literature deals with learning-to-forecast experiments in New Keynesian models. Adam (2007) shows that in such an environment subjects expectation formation process generally fails to be rational, but can rather be described by simple forecasting rules based on lagged inflation. Pfajfar and Zakelj (2014) and Assenza et al. (2013) study the expectation formation process of the subjects and its interaction with conventional monetary policy design. They find a stronger mandate for price stability to better stabilize private-sector expectations and thereby the economy. Kryvtsov and Petersen (2015) show that much of this stabilizing power is through the effect on private-sector expectations. Close to the zero lower bound, however, Hommes et al. (2015) find that conventional monetary policy is generally not very effective in stabilizing the economy and insulating it from the risk of falling into an expectation driven liquidity trap.

The effects of forward guidance on economic stability in New Keynesian learning-to-forecast experiments are mixed. While Cornand and M'Baye (2016a,b) find that the communication of the central bank's inflation target can reduce the volatility of the economy in normal times, Arifovic and Petersen (2015) find that it does not provide a stabilizing anchor in crisis times, e.g. in a liquidity trap. Mokhtarzadeh and Petersen (2016) find that providing the economy with the central bank's projections for inflation and the output gap stabilizes the economy, while Kryvtsov and Petersen (2015) find that providing the expected future interest rate path diminishes the effectiveness of monetary policy in stabilizing the economy.

The contribution of this project to the literature is twofold. First, we analyze the stabilizing role of central bank forward guidance in the form of inflation projections in normal times and in times of severe economic stress. Publishing inflation projections is common practice for central banks, but has yet received very little attention in the context of New Keynesian learning-to-forecast experiments. To the best of our knowledge, the only exception is Mokhtarzadeh and Petersen (2016). Second, we investigate how the central bank's effectiveness to influence expectations depends on its endogenous degree of credibility.

The papers closest to ours are Mokhtarzadeh and Petersen (2016) and Goy et al. (2016). Mokhtarzadeh and Petersen (2016) also study the effect of public central bank projections on expectation formation, future credibility and the stabilizing role of forward guidance in a New Keynesian learning-to-forecast experiment. Yet there are substantial differences in the methodology. Mokhtarzadeh and Petersen (2016) study the publication of a larger set of macroeconomic projections comprising five-period ahead projections of inflation, the output gap, and the nominal interest rate. The projections are generated by a computerized central bank that assumes agents to form expectations rationally or adaptively. The current paper focuses on the publication of one-period ahead inflation projections, i.e., we abstract from output gap and interest rate projections. The inflation projections in this paper are generated either by a student subject who does not have to follow any specific expectation formation mechanism, or by a computerized algorithm that assumes that agents form expectations according to a Heuristic Switching Model as presented in Assenza et al. (2013). The perhaps largest difference between the two studies arises from the motive for forward guidance. Mokhtarzadeh and Petersen (2016) study the stabilizing role of informative forward guidance, whereas the present paper studies the stabilizing role of strategic forward guidance in normal times and at the zero lower bound. Goy et al. (2016) present a theoretical analysis under learning - instead of a laboratory experiment - of central bank forward guidance in a New Keynesian model with zero lower bound and boundedly rational and heterogeneous agents. Despite the differences in methodology, the general conclusions of Mokhtarzadeh and Petersen (2016), Goy et al. (2016), and this paper are reconfirming. All three contributions find that the publication of inflation forecasts is a helpful tool to anchor private sector expectations and to stabilize the economy.

The experimental setup mainly follows the work by Assenza et al. (2013), with two major differences: (i) subjects face a public central bank projection which they can utilize in forming their own expectations and (ii) output gap expectation are not subject-based but model-based. The latter assumption is made in order to keep the experimental task for the subjects simple and to focus this study entirely on inflation expectations. Output gap expectations are endogenously determined by the model following a Heuristic Switching Model, which has proven to fit well learning-to-forecast experiments in New Keynesian frameworks (e.g. Assenza et al., 2013).

3 Experimental Design

Subjects interact with the economy through expectations of inflation, which affect the outcome of the economy through a positive feedback[6] of the form:

$$\pi_t = f\left(\bar{E}_t \pi_{t+1}\right),$$

where π_t and $\bar{E}_t \pi_{t+1}$ denote inflation and aggregate private-sector expected future inflation, respectively, and f is a functional form, which is specified below. Note that subjects do not yet know the realization of π_t when they form their expectation about π_{t+1}. We follow Arifovic and Petersen (2015) and Kryvtsov and Petersen (2015) and define aggregate inflation expectations as the median[7] of the individual inflation expectations, i.e., $\bar{E}_t \pi_{t+1} = median(\mathbf{E_t \pi_{t+1}})$, where $\mathbf{E_t \pi_{t+1}}$ is a vector collecting all $j = 1, ...J$ professional forecasters' individual inflation expectations $E_t^{fc,j} \pi_{t+1}$ of period t for period $t+1$.

3.1 The New Keynesian Economy

The underlying economy evolves according to a New-Keynesian model under heterogeneous expectations.[8]

$$y_t = \tilde{E}_t y_{t+1} - \frac{1}{\sigma}\left(r_t - \bar{E}_t \pi_{t+1} - \bar{r}\right) + \varepsilon_t, \tag{1}$$

$$\pi_t = \beta \bar{E}_t \pi_{t+1} + \kappa y_t + \eta_t, \tag{2}$$

$$r_t = max\left[0, \bar{r} + \pi^T + \phi_\pi\left(\pi_t - \pi^T\right) + \phi_y y_t\right], \tag{3}$$

where y_t is the aggregate output gap, r_t is the nominal interest rate, $\bar{r} = \frac{1}{\beta} - 1$ is the steady state interest rate, and $\tilde{E}_t y_{t+1}$ is the aggregate expected future output gap. The parameter π^T denotes the central bank's target value for inflation. Finally, the economy is perturbed by stochastic i.i.d demand and supply shocks, denoted η_t and ε_t, respectively.[9]

[6] Positive feedback means that the derivative of the function $f(\cdot)$ is positive. Note that although the nominal interest rate rule (3) adds some negative feedback to the economy, the overall feedback of inflation expectations on current inflation remains positive, independent of the coefficients in this interest rate rule.

[7] When the aggregate is determined as the mean of all forecasts, an individual could cast an extreme forecast, in order to obtain an extreme aggregate, which would then feed back into the economy. Such individual strategic power that does not reflect the real world is eliminated when the aggregate is instead determined by the median of all forecasts.

[8] Microfoundations for this model under heterogeneous expectations can be found, for instance, in Branch and McGough (2009), Kurz et al. (2013), and Hommes and Lustenhouwer (2015).

[9] There are six economies (groups) in each treatment. Therefore, there are six random shock processes each for η_t and ε_t. These are applied to all treatments so that each shock sequence

The calibration of the constant model parameters follows Clarida et al. (2000). I.e., we set the quarterly discount factor $\beta = 0.99$, implying an annual risk-free interest rate of four percent. The coefficient of relative risk aversion is set to $\sigma = 1$ and the output elasticity of inflation is $\kappa = 0.3$. The quarterly inflation target is set to $\pi^T = 0.00045$, implying an annual inflation rate of 0.18 per cent.[10] The Taylor rule coefficients are chosen to be $\phi_\pi = 1.25$ and $\phi_y = 0.3$, which is well within the range of values that are common in related experiments.[11]

Equation (1) refers to an optimized IS curve, equation (2) is the New Keynesian Phillips curve and equation (3) is the rule for the nominal interest rate set by the central bank. We assume the central bank follows a Taylor (1993) type interest rate rule, where it adjusts the interest rate in response to inflation and output gap. Furthermore, equation (3) also shows that the nominal interest rate is subject to a zero lower bound.[12] Under rational expectations this model has two equilibria. A determinate equilibrium equal to the target steady state[13] that has values of inflation and output (close to) $\pi_t = y_t = 0$ given that π^T is (close to) zero, and an indeterminate equilibrium where the zero lower bound on the nominal interest rate is binding and $(\pi_t, y_t) = (-\bar{r}, -\frac{1-\beta}{\kappa}\bar{r})$ (Benhabib et al., 2001). Under adaptive learning the target steady state is locally stable (if the Taylor principle is satisfied), while the zero lower bound steady state is an unstable saddle-point. Therefore, depending on initial conditions, either convergence to the target steady state occurs or the economy falls into a deflationary spiral (Evans et al., 2008).

Since we focus on how the central bank can stabilize the economy by publishing inflation projections, we do not make any assumptions on the way inflation expectations are formed, but ask the subjects in the lab for their inflation expectations. $\bar{E}_t \pi_{t+1}$ is therefore an aggregation of elicited expectations. In contrast, $\tilde{E}_t y_{t+1}$ is endogenously determined by the model. $\tilde{E}(y)$ follows a Heuristic Switching Model[14], that was originally developed to fit a learning-to-forecast experiment in an asset price setting (Anufriev and Hommes, 2012), but has proven its robustness to fit also learning-to-forecast experiments in New Keynesian frameworks (e.g. Assenza et al., 2013).

The Heuristic Switching Model can be summarized by the following equations:

is applied once in each treatment. In particular, the following pairings arise: E1-E7-E13-E19, E2-E8-E14-E20, E3-E9-E15-E21, E4-E10-E16-E22, E5-E11-E17-E23, E6-E12-E18-E24.

[10] We choose a value of the inflation target near zero to be in line with the zero inflation steady state that is assumed when log-linearizing the macro economic model to obtain equations (1) and (2). We choose however a value slightly different from zero in order not to present subjects with a round number on which they can easily coordinate.

[11] Standard values for comparable experiments range from $\phi_\pi \in (1, 2)$ and $\phi_y \in (0, 0.5)$, e.g., Cornand and M'Baye (2016b) and Arifovic and Petersen (2015) among others.

[12] Note that under commitment to a Taylor rule, setting the nominal interest rate is not part of the task attributed to the subject with the role as central bank forecaster. Rather the nominal interest rate is influenced implicitly, through the effects of forward guidance on private-sector expectations and their feedback on the economy. Information about likely feedback effects and the corresponding prescribed reaction of future interest rates are provided to the central bank (described in detail in Section 3.3.2) as input for the inflation projection. Thereby, forward guidance and the nominal interest rate are in practice not chosen independent of each other.

[13] The rational expectations equilibrium coincides with the steady state because shocks are not autocorrelated.

[14] Heuristic Switching Models were introduced by Brock and Hommes (1997).

$$\begin{cases} \text{Adaptive Rule} & \rightarrow \quad E_t^{ada} y_{t+1} = 0.65 y_{t-1} + 0.35 E_{t-1}^{ada} y_t \\ \text{Weak Trend} & \rightarrow \quad E_t^{wtr} y_{t+1} = y_{t-1} + 0.4\left(y_{t-1} - y_{t-2}\right) \\ \text{Strong Trend} & \rightarrow \quad E_t^{str} y_{t+1} = y_{t-1} + 1.3\left(y_{t-1} - y_{t-2}\right) \\ \text{Learn and Anchor} & \rightarrow \quad E_t^{laa} y_{t+1} = \frac{\left(y_{t-1}^{av} - y_{t-1}\right)}{2} + \left(y_{t-1} - y_{t-2}\right) \end{cases} \tag{4}$$

$$U_{t-1}^h = \frac{100}{1 + \left|y_{t-1} - E_{t-2}^h y_{t-1}\right|} + \eta U_{t-2}^h \tag{5}$$

$$n_t^h = \delta n_{t-1}^h + (1-\delta) \frac{exp\left(\gamma U_{t-1}^h\right)}{\sum_{j=1}^4 exp\left(\gamma U_{t-1}^j\right)} \tag{6}$$

$$\tilde{E}_t y_{t+1} = E_t^{ada} y_{t+1} n_t^{ada} + E_t^{wtr} y_{t+1} n_t^{wtr} + E_t^{str} y_{t+1} n_t^{str} + E_t^{laa} y_{t+1} n_t^{laa} \tag{7}$$

Equation (4) lists the set of heuristics available to the agents when forming their expectations. The variable y_{t-1}^{av} denotes the average of past output gaps. Once heuristics are used, the agents weight their past performance following equation (5), with η denoting the parameter describing the preference for the past. Equation (6) updates the probability of using heuristic h when forecasting for period $t+1$. Notice that γ captures the sensitivity of agents to heuristic performances and δ denotes the fraction of agents that in period t stick to the heuristic they used in period $t-1$. Then, using, (7) the expectation are aggregated and $\tilde{E}_t y_{t+1}$ is determined. The calibration of the Heuristic Switching Model follows Assenza et al. (2013), i.e., we set $\eta = 0.7$, $\delta = 0.9$, and $\gamma = \left(0.4 \cdot 4^2\right) = 6.4$.[15]

3.2 The Experiment

We apply a learning-to-forecast experiment following the approach of Assenza et al. (2013). The general setup is as follows: subjects in the laboratory are randomly divided in groups of 7. Subjects either take the role as a professional forecaster or as a central bank forecaster. Professional forecasters are employed at the forecasting department of a company which needs predictions about future inflation as input for the management's operative decisions. Professional forecasters' job is to generate these inflation forecasts and to communicate them to the management. Professional forecasters are provided with some qualitative knowledge of the economy,[16] the direction of the feedback on their expectations (i.e. positive feedback), and a public central bank projection. The professional forecasters' payoffs are determined according to their forecasting performance, measured by the following payoff function from Assenza et al. (2013):

$$\Pi_{fc,j} = \frac{100}{1 + \left|\pi_{t+1} - E_t^{fc,j} \pi_{t+1}\right|}. \tag{8}$$

[15] We multiply γ by 4^2 relative to the calibration of Assenza et al. (2013) because we use a Heuristic Switching Model with quarterly rather than annualized data.

[16] This is a common assumption shared among all studies cited in Section 2 of this paper, except for Adam (2007), who does not provide any information about the working of the economy. We abstract from providing the subjects with the fully quantified set of equations, as real world economists neither know the full set of specific equations nor their quantitative relations in the real world economy.

The central bank forecaster is employed at the forecasting department of the central bank and the central bank forecaster's job, too, is to generate inflation forecasts, which we denote $E_t^{cbf}\pi_{t+1}$. However, this forecast does not enter the vector $\mathbf{E_t}\pi_{\mathbf{t+1}}$ from which the aggregate inflation expectation is determined. The incentives for the central bank forecaster in determining her inflation forecasts, therefore, are different from the incentives of professional forecasters and also differ strongly between treatments, as explained below.

Whether a subject is assigned the role of a professional forecaster or a central bank forecaster is the outcome of a preliminary stage (henceforth: Stage I). Independent of the treatment, in Stage I, all subjects of a group play 8 initial rounds of the experiment as professional forecasters in the absence of any public central bank inflation projection. To level the playing field, all participating subjects are presented with an identical three-period history (for periods $t = -2$, $t = -1$, and $t = 0$) for inflation, the output gap and the interest rate, which initializes the economy off the steady state.[17] Subjects are ranked according to their relative forecasting performance. The role of the central bank forecaster for the remaining rounds of the experiment (period 9-37) is assigned to the best ranked subject. This is common knowledge.

Apart from the control treatment (Treatment 1), as the economy enters period 9, at the beginning of each period the central bank publishes an official central bank inflation projection, denoted by $E_t^{pub}\pi_{t+1}$. Depending on the treatment, this official central bank inflation projection is produced either by the central bank forecaster so that $E_t^{cbf}\pi_{t+1} = E_t^{pub}\pi_{t+1}$ (Treatment 2) or by a computer algorithm (Treatments 3 and 4). In any case, the central bank is provided with superior information that can be used in the forecasting process. Professional forecasters are subsequently presented with the official central bank projection before they submit their own forecasts.

Since we are interested in the expectations channel of monetary policy both in normal times and in times when the zero lower bound on the nominal interest rate may become binding, in the spirit of Arifovic and Petersen (2015), there is a series of negative fundamental shocks, which hit the economy. In this experiment, the series of fundamental shocks appears in a very late stage of the experiment, in particular it starts in period 29 and prevails for four periods. This series of fundamental shocks is chosen such that it is likely to induce a liquidity trap and therewith the possibility of a deflationary spiral.

With this subdivision, the economy is fairly stable in the first part of the actual experiment (periods 9-28; henceforth: Stage II) and it is investigated whether central bank forward guidance can influence private-sector expectations and actively stabilize the economy. In the latter part of the experiment (periods 29-37; henceforth: Stage III), on the other hand, it is investigated whether the central bank can prevent or reverse a deflationary spiral by means of forward guidance.

The timing of the experiment is as follows: In $t = 1, ..., 8$ (Stage I), all subjects submit their inflation forecast $E_t^{fc,j}\pi_{t+1}$ simultaneously. In $t = 9, ..., 37$ (Stages II and III), first the central bank forecaster submits her forecast $E_t^{cbf}\pi_{t+1}$. With the exception of Treatment 1, afterwards the official central bank projection $E_t^{pub}\pi_{t+1}$ is published. Professional forecasters observe the public inflation

[17]The history is displayed in Figure 3 in Appendix C. It comprises the first three observations.

projection of the central bank and subsequently submit their own inflation forecasts $E_t^{fc}\pi_{t+1}$. After all professional forecasters have submit their forecast, the aggregate inflation forecast $\bar{E}_t\pi_{t+1}$ is determined and the values for the variables in period t are computed. The economy proceeds to the next round.

While the objective of the professional forecasters remains the same in all treatments throughout the whole experiment, the objectives of the central bank forecaster differ across treatments. These differences are described in detail in the following subsection.

3.3 Treatments

We consider four treatments in this experiment.

3.3.1 Treatment 1: Control treatment

In this treatment, the control treatment, no central bank projections are published, i.e., there is no central bank forward guidance. The central bank forecaster produces forecasts, but the forecasts of the central bank forecaster are not publicly shown. Therefore, the central bank forecaster has no ability to influence the professional forecasters' expectations and thereby no incentive to produce strategic forecasts.

In each period, the central bank forecaster is provided with a data-driven forecast $E_t^{ddf}\pi_{t+1}$, which she can choose to consider or to ignore when forming her own inflation forecasts. The data-driven forecast uses model equations (1) to (7) and data up to period $t-1$ to predict what level of inflation is likely to prevail in period $t+1$. However when the data-driven forecast is made, it is not yet known what aggregate inflation expectations formed in periods t and $t+1$ will be, which are important determinants of inflation in period $t+1$. These expectations therefore need to be modeled. This is done by assuming a Heuristic Switching Model for inflation expectations analogous to equations (4) to (7).

In this treatment, all subjects (including the central bank forecaster) share the same incentives arising from equation (8); i.e., even the central bank forecaster's goal is simply to predict inflation accurately.

3.3.2 Treatment 2: Forward Guidance from a Human Central Bank Forecaster

In this treatment, the central bank forecaster publishes official central bank inflation projections, i.e., $E_t^{pub}\pi_{t+1} = E_t^{cbf}\pi_{t+1}$. Hence, there is central bank forward guidance. The other subjects of her group are informed (i) that there is a central bank forecaster publishing official central bank inflation projections in this economy, (ii) that the central bank forecaster is the subject that predicted inflation best in Stage I, (iii) that the central bank forecaster has additional information about the economy without specifying this any further, and (iv) that the central bank has an inflation target without quantifying this target.

In this setup, the central bank forecaster may have an ability to influence the professional forecasters' expectations and thereby may have an incentive to produce strategic projections. Note that it is not a priori clear whether it is optimal for professional forecasters to use the published projection when forming their own forecasts or to ignore it. This depends on what a subject believes about how the central bank forms its projection and about how other subjects form their expectations.[18]

As in Treatment I, the central bank forecaster is provided with a data-driven forecast. The data-driven forecast algorithm is however somewhat different than in the control treatment. Now it must account for the potential self-fulfilling properties a published central bank projection can have on the economy. This works as follows: when the central bank publishes a projection, this is likely to affect, to some extent, the inflation expectations of the professional forecasters. Since the main determinant of current inflation is inflation expectations, aggregate expectations of professional forecasters in turn affect realized inflation. This implies that when the published projection is high, this is likely to also lead to somewhat higher aggregate inflation expectation, and therefore to a higher inflation realization. The task of the data-driven forecast in this treatment is to find, given this possibly self-fulfilling feedback mechanism, those expectation values that, when published, are most likely to come true. This is done by including a fifth heuristic to the Heuristic Switching Model that is used to model inflation expectations. This heuristic can be termed "Follow the Published projection" and is defined by $E_t^{fpp}\pi_{t+1} = E_t^{pub}\pi_{t+1}$. The fitness of this heuristic is calculated analogously to equation (5), and 5 heuristics are now considered when calculating fractions as in equation (6). This implies that the data-driven forecast assumes that aggregate inflation expectations will be more in line with the published projection when the past published projections have been relatively accurate. The data-driven forecast then performs a grid search to choose the forecast that is most likely to be accurate, taking account of the effects that such a forecast are likely to have on aggregate expectations.[19]

Moreover, the central bank forecaster is provided with information about which aggregate inflation expectations for the following period need to prevail for inflation to be (in expectations) at target level π^T already in the current period. This specific aggregate inflation expectation is calculated by doing a grid search on $\bar{E}_t\pi_{t+1}$ in the model defined by equations (1) to (7). This information tells the central bank forecaster in what direction she should steer aggregate expectations about next period in order to get closer to her inflation target in this period. We label this piece of information "required for target" and denote it by $E_t^{rft}\pi_{t+1}$.

Finally, the central bank forecaster is presented with a credibility index I_t^{cred},

[18] For example, it is optimal for a subject to predict exactly the published forecast when she thinks that the central bank is able to foresee what the median forecast will be and that the central bank will use all its information to publish a truthful forecast. If, on the other hand, the subject believes that the central bank is not good in predicting the median forecast of the professional forecasters or if she believes that the central bank is more concerned with strategically trying to steer the economy rather then publishing accurate projections, then the subject is better of ignoring the published forecast.

[19] Since the published forecast about $t+1$ affects realizations in period t, and the published forecast about $t+2$ affects realizations in $t+1$, an assumption needs to be made about what the published forecast about $t+2$ will be, in order to evaluate whether the forecast made about $t+1$ is likely to come true. The data driven forecast simply assumes here that the published forecast about $t+2$ will be the same as the published forecast about $t+1$. Since both inflation and the published forecast turn out to be highly persistent, also in our experimental sessions, this is arguably not a very restrictive assumption.

given by

$$I_t^{cred} = \frac{1}{4}\sum_{i=1}^{4}\left[\frac{1}{6}\sum_{j} exp\left(-3\cdot\left(E_{t-i}^{pub}\pi_t - E_{t-i}^{fc,j}\pi_t\right)^2\right)\right], \tag{9}$$

where, in the spirit of Cecchetti and Krause (2002), the central bank's credibility towards professional forecaster j is defined by the distance between the central bank's inflation projection and the inflation forecasts of professional forecaster j.[20] The credibility index is based on the distance between the projections of the central bank forecaster and of the individual professional forecasters in the last four periods. We take the exponent of the negative of each squared distance in order to make sure that the index is between 0 and 1. The scale parameter 3 is calibrated based on pilot data in such a way that the index is not too easily close to 1. When the index equals 1, every individual forecaster predicted exactly the same as the central bank forecaster in each of the last four periods. When all professional forecasters made forecasts that were quite far away from the projections of the central bank forecaster in the last four periods, the index is close to zero.

Having available these three sources of information, the central bank forecaster must decide to what extent she follows the data-driven forecast or to what extent she publishes a strategic projection based on the "required for target", taking into account her credibility.

The central bank forecaster's objective, in this treatment, is twofold; i.e., there are two payoff functions. On the one hand she has to stabilize inflation, i.e., minimize the deviations of inflation from her target values, while on the other hand her inflation projections have to remain maximally credible, as measured by the credibility index. We consider central bank credibility explicitly, as it is of utmost importance for the functioning of monetary policy and thereby enjoys a lot of attention of monetary policy makers (Blinder, 2000; Bordo and Siklos, 2014). In line with this strategy, Gomez-Barrero and Parra-Polania (2014) present a theoretical model of strategic central bank forecasting which explicitly considers reputational concerns of central bank credibility in the central bank's loss function. The payoff functions of the central bank forecaster have the following form:

$$\begin{array}{lcl} \Pi_{cbf}^{stability} & = & \max\left(0, 100 - 44.4\left(\pi_t - \pi^T\right)^2\right) \\ \Pi_{cbf}^{credibility} & = & \max\left(0, 100 - 400\left(1 - I_t^{cred}\right)^2\right) \end{array}. \tag{10}$$

Equation (10) is calibrated such that in each period the central bank forecaster receives a payoff of zero for stability if inflation deviates from target by more than 1.5 percentage points and receives a payoff of zero for credibility of the projection if the credibility index is below 0.5. At the end of the experiment, one of these two objectives is chosen randomly by the computer and the central

[20]In the econometric analysis, additionally we consider two alternative approaches to measure credibility. First, credibility as the weight of the information content from public announcements of the central bank (in our environment the public forecasting signals) attached by professional forecasters to their expectation formation process (Bomfim and Rudebusch, 2000). Second, as subjectively elicited measure, asking them a Lickert question at the beginning, in the middle, and at the end of the actual experiment.

bank forecaster is paid according to the total payoff of the chosen objective. The randomization eliminates any incentives to focus on only one of the two goals or to strategically play one goal of another in any other way.

3.3.3 Treatment 3: Forward Guidance from a "good" Computerized Central Bank Forecaster

In this treatment, the official central bank projection is published by a computer algorithm, i.e., again there is central bank forward guidance. The computer algorithm publishes strategic inflation projections, which intend to steer the economy back to target via the manipulation of private-sector expectations. The extent to which the projections try to manipulate rather than to purely inform private-sector expectations depends primarily on the current state of the economy (in particular, whether previous inflation was (i) close to, (ii) above or (iii) below its target value) and secondarily on the credibility of recent central bank inflation projections.

The computer algorithm works as follows: (i) If previous inflation was close to the target (within ± 0.5 percentage points), the central bank tries to initiate long term coordination on its inflation target through projections equal to the inflation target. (ii) If previous inflation was far above target (for more than 0.5 percentage points), the algorithm makes a trade-off between building credibility and steering the economy. If past projections have been only little credible, the algorithm aims at building credibility through accurate inflation projections by publishing projections close to the data driven forecast (which is calculated in the same way as in Treatment 2). If projections have already been credible, the algorithm leans more towards the "required-for-target" information. (iii) If previous inflation was far below target (for more than 0.5 percentage points) the economy faces the risk of a binding zero lower bound and a deflationary spiral. Now, building up credibility by following the data-driven forecast becomes dangerous as the data-driven forecast may predict a deflationary spiral. Therefore, the algorithm balances forecasting the target with forecasting the last observed inflation level, where the latter can improve on credibility without amplifying the downturn in inflation. The weight on the last observed value is relatively high when there is a downward trend in inflation, because then it might not be credible that inflation will suddenly go up by much. On the other hand, if there is an upward trend in inflation it might be more credible that inflation will go up more, so the computer algorithm can put more weight on the target.

The explicit algorithm is spelled out below:
"close to target": $E_t^{pub}\pi_{t+1} = \pi^T$

"sufficiently above target": $E_t^{pub}\pi_{t+1} = I_t^{cred} * E_t^{rft}\pi_{t+1} + (1 - I_t^{cred})E_t^{ddf}\pi_{t+1}$

"sufficiently below target": if $\pi_{t-1} < \pi_{t-2} : E_t^{pub}\pi_{t+1} = 0.5\pi^T + 0.5\pi_{t-1}$
if $\pi_{t-1} > \pi_{t-2} : E_t^{pub}\pi_{t+1} = 0.8\pi^T + 0.2\pi_{t-1}$

For reasons of comparability, in this treatment, the central bank forecaster subject takes the same role as in Treatment 1, however, she is not provided with any additional information. This allows us to learn more about the expectation formation process of the central bank forecaster, especially the dependence on additional information for the forecasting performance and the expectation formation process of the central bank forecaster in case this subject cannot interact with the economy.

3.3.4 Treatment 4: Forward Guidance from a "bad" Computerized Central Bank Forecaster

This treatment is similar to Treatment 3, but with a different computer algorithm in Stage II. In Stage II of this treatment, the computer algorithm publishes inflation projections, which it randomly draws from a uniform distribution with support from -5 to 5, i.e., $E_t^{pub}\pi_{t+1} \sim Unif(-5, 5)$. The support is chosen according to the support of realized inflation throughout the first three treatments of this experiment. In Stage III of this treatment, the computer algorithm is the same as in Treatment 3. This twist after Stage II allows us to analyze the influence of credibility on the central bank's ability to stabilize the economy in times of severe economic stress.

3.4 Hypotheses

Our experimental design allows us to address several hypothesis. The following hypotheses distinguish between "informative" and "random" forward guidance. We consider forward guidance informative, if the central bank inflation projection lies systematically (i.e. most of the time) inside the interval between the data-driven forecast and the "required for target" information. Analogously, forward guidance is considered "random," if the central bank inflation projection lies systematically (i.e. most of the time) outside the interval between the data-driven forecast and the "required for target" information. According to this criterion, forward guidance from a human central banker forecaster (Stages II and III of Treatments 2) and from the "good" computerized central bank forecaster (Stages II and III of Treatment 3 and Stage III of Treatment 4) are considered "informative" and forward guidance from the "bad" computerized central bank forecaster (Stage II of Treatment 4) is considered "random." [21]

Hypothesis 1: Informative forward guidance stabilizes the economy (a) in normal times and (b) in times of severe economic stress; random forward guidance does not.

Although from an empirical point of view published central bank inflation projections seem beneficial for macroeconomic stability (Chortareas et al., 2002), from a theoretical point of view, the effects of published central bank inflation projections on macroeconomic stability are generally ambiguous and depend on the quality of the projections (see Geraats (2002) for an extensive survey).

For instance, Gersbach (2003) and Jensen (2002) find that publishing inflation projections may be destabilizing, as it carries information about future

[21] For the central bank forecaster subjects, more than 85% of all public central bank projections lie within the required interval; for the "good" computer algorithm it is more than 80% (and above 90% if the predictions of the target inflation rate when the economy is "close to target" are considered as well). For the "bad" computer algorithm, less than 4% of all public central bank projections lie within the required interval.

shocks, which are internalized by private-sector expectations and therefore cannot be stabilized by the central bank anymore. Geraats (2002), Amato and Shin (2006), Walsh (2007) argue that central bank communication can be destabilizing as potentially noisy public information crowds out accurate private information. By contrast, Tarkka and Mayes (1999) find that publishing central bank projections conveys information about the central bank's targets as well as about the central bank's belief about private-sector expectations, which enhances the predictability of monetary policy actions and reduces output volatility. Along similar lines, Geraats (2005) finds that publishing inflation projections reduces the inflation bias. In a standard New Keynesian model with learning, Eusepi and Preston (2010) and Ferrero and Secchi (2010) find a stabilizing role of public central bank inflation projections through an anchoring effect on private-sector inflation expectations in normal times. Goy et al. (2016) reach similar conclusions in a New Keynesian model with boundedly rational heterogeneous agents, however, not only in normal times but also in times of severe economic stress. At the zero lower bound, the authors show that the publication of inflation projections can lower the likelihood of a deflationary spiral if central bank projections are sufficiently credible.

Hypothesis 2: Informative forward guidance anchors private-sector inflation expectation; random forward guidance does not.

In their seminal theoretical contribution, Morris and Shin (2002) show that public central bank information can act as a coordination device by anchoring private-sector expectations and thereby reduce the dispersion of private-sector expectations. Empirical support for such an anchoring effects for expectations (especially in the context of public central bank projections) is given by Hubert (2014) for the Federal Reserve, by Fujiwara (2005) for the Bank of Japan, and by Ehrmann et al. (2012) for 12 advanced economies (including the former two).

Hypothesis 3: Informative forward guidance increases the forecasting accuracy of all market participants; random forward guidance does not.

Dale et al. (2011) show in a stylized model of imperfect knowledge and learning that if the central bank has an informational advantage with respect of the functioning of the economy and if this informational advantage is perceived correctly by the private sector, publishing inflation projections can improve the accuracy of private-sector expectations. By contrast, if the central bank projections are imprecise and noisy, the publication of these projections might unleash distracting forces which deteriorate the accuracy of private-sector expectations.

Hypothesis 4: The degree of "strategic-ness" of a public central bank projection depends on its credibility.

Recent empirical evidence gives rise to the assumption that central bank projections are not just a purely informational tool but are also used as a strategic instrument to influence private-sector expectations, which manifests in biased projections. Indicative evidence for such a claim is presented by Romer and Romer (2008), who find that the forecasting accuracy of the Federal Open Market Committee (FOMC) is systematically lower relative to the projection of their own research staff (the so-called Greenbook projections), even though these forecasts are available to the FOMC when publishing their projections. For a sample of ten inflation targeting central banks (Australia, Canada, Chile, Czech Republic, Korea, new Zealand, Mexico, Norway, Poland, and Sweden) Charemza and Ladley (2016) find that central bank projections are biased towards their inflation targets.

From a theoretical point of view, Jensen (2016) shows that - if credible - optimal inflation projections are indeed misleading, whereas non-misleading inflation projections are time-inconsistent in an augmented Barro-Gordon type game featuring a New Keynesian sticky price model. In a related augmented Barro-Gordon type game, Gomez-Barrero and Parra-Polania (2014) show that the degree of the inflation projection bias is endogenous, as it optimally solves the trade-off between the benefits (i.e. the enhanced stabilizing effects) and the costs (i.e. the loss of credibility for the inflation projection) of the strategic bias. Empirical evidence from six inflation targeting central banks (Brazil, Canada, England, Iceland, New Zealand, and Sweden), presented by Gomez-Barrero and Parra-Polania (2014), turns out to be consistent with their theoretical predictions.

In the context of this experiment, the degree of "strategic-ness" is measured as the relative weight given to the "required for target" over the data-driven forecast in the published inflation projection. This measure is consistent with the notion of an intentional over- or underreporting of the projected level of inflation.

Hypothesis 5: (a) The credibility of the central bank projections depends positively on their past performance and (b) the ability of the central bank to stabilize the economy by means of its projections depends positively on the past credibility of the central bank projections.

In a survey among 84 central bank presidents worldwide, Blinder (2000) finds that the most important matter for credibility is believed to be a consistent track record. With respect to inflation projections and projection of inflation in particular, such a consistent track record is established primarily by a sustained projection accuracy. Loss in credibility of the central bank's projections can therefore be attributed to a (systematic) failure to produce accurate projections (Mishkin, 2004). Following this line of reasoning, also the two most closely related studies to this paper, Goy et al. (2016) and Mokhtarzadeh and Petersen (2016), determine central bank credibility by looking at past central bank forecasting performance.

A good deal of credibility, in turn, is necessary for forward guidance to be effective in stabilizing the economy (Filardo and Hofmann, 2014). A particularly illustrative example in this respect is provided by Svensson (2015) for the Swedish case (although with respect to interest path projections). While credible projections remarkably influenced market behavior towards stabilization in 2009, in 2011 non-credible projections left the market unimpressed and without any response in market behavior.

3.5 Experimental Procedure

Each treatment of this experiment consists of six economies with seven subjects each. Thus, the experiment has a total of $4 * 6 * 7 = 168$ subjects. Subjects were recruited from a variety of academic backgrounds using ORSEE (Greiner, 2015). The subject population comprised undergraduate students (64%), graduate students (34%), and non students (2%). Subjects were mostly from the natural sciences (61%) and the social sciences (16%). Around two thirds of the subjects were male (62%) and one third were female (38%). During the experiment, subjects earned experimental currency units (ECU) according to their respective payoff functions. At the end of the experiment, subjects were paid €1 for every 85 ECU; that is, each ECU paid approximately €0.012. The average payment was €31.66. The experimental software was programmed in oTree (Chen et al., 2016). The experiment was conducted in May and June 2016 at the experimental lab of the Technische Universität Berlin.

4 Macroeconomic Results

In this section, we address Hypotheses 1, 2, and 3, i.e. we analyze the role of central bank forward guidance for the macroeconomy. To fix ideas, first we juxtapose the median economic dynamics arising from the actual experiment in each of the four treatments and their statistical properties.

Figure 1 shows the median evolution of inflation, the output gap, and the interest rate for all four treatments; Treatment 1 is depicted by the solid lines, Treatment 2 by the dashed lines, Treatment 3 by the dotted lines, and Treatment 4 by the dashed-dotted lines.[22] The figure shows that all four treatments share a common pattern for the evolution of the macroeconomy over much of the 37 rounds of the experiment. First, there is slow convergence towards the steady state. Second, starting in period 29 (the second vertical, gray line), a deep recession takes place, which drives the economy towards the zero lower bound at which it remains for an extended period of time. However, while median economies recover from the recession under central bank forward guidance (Treatments 2-4), the median economy produces a deflationary spiral in the absence of central bank forward guidance (Treatment 1).

Although, at first sight, the general pattern looks very similar across all four treatments (with the exception of Stage III), we find considerable effects of central bank forward guidance on the economy. Tables 9 to 12 in Appendix C summarize descriptive statistics for all 24 economies in Treatments 1, 2, 3, and 4, respectively. Comparing the descriptive statistics shows that in the preliminary stage (i.e., Stage I) medians and variances[23] of all three macroeconomic variables inflation, the output gap, and the interest rate are very close across treatments. We test for equality of the medians and variances for pairwise comparison of treatments using the non-parametric Mann-Whitney-Wilcoxon-test and Siegel-Turkey-test, respectively. The results of these tests are presented in Tables

[22]Figures 4 to 7 in Appendix C show all 6 individual economies for each treatment, respectively.

[23]For completeness, we also present means in these tables. All results for medians qualitatively carry over to means. Therefore, for the rest of the analysis we do not consider them explicitly. Furthermore, comparing means statistically necessitates parametric tests which given the small number of observations are not appropriate.

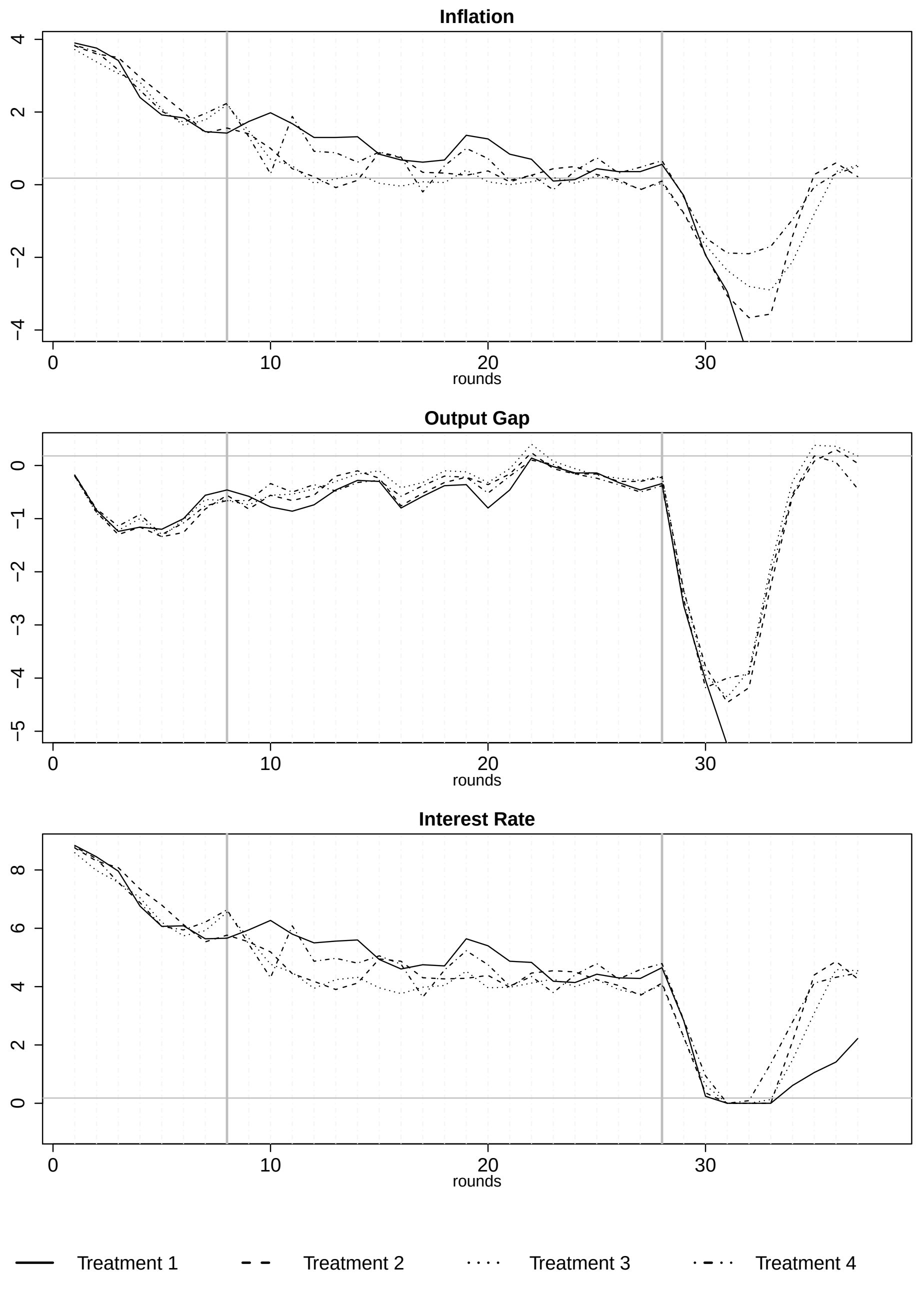


Figure 1: Median responses of inflation (upper panel), the output gap (middle panel), and the interest rate (lower panel) for all four treatments. For each treatment, median responses are generated by taking the median of each inflation, the output gap, and the interest rate from all six economies at each period $t = 1, ..., 37$.
Note that for Treatment 1 the median interest rate leaves the zero lower bound despite a deflationary recession. This abnormal artifact is a result from the aggregation procedure (median) as three economies of Treatment 1 remain at the zero lower bound, while three economies leave the zero lower bound (see Figure 4 in Appendix C.)

13 and 14 in Appendix C. They show that the Null hypothesis of equality in medians and variance cannot be rejected for any pairwise comparison of treatments.

In Stage II, by contrast, median and variance of all three variables in Treatments 2 and 3 (i.e., under informative forward guidance) are considerably and statistically significantly closer to the (determinate) target steady state compared to Treatments 1 and 4 (i.e., without informative forward guidance). Hence, informative forward guidance has a significant influence on private-sector expectations which helps reduce the economy's volatility and drives it closer to the steady state. Random forward guidance (Treatment 4), by contrast, has rather averse effects on the economy. We find a marginal but statistically significant increase in median inflation and the median interest rate compared to Treatment 1, while a slight reduction in the median output gap is statistically insignificant. Moreover, random forward guidance is without significant effect on the volatility of the economy.

During severe economic stress (Stage III), informative forward guidance (in Stage III this is Treatments 2-4) keeps the economy closer to the steady state and strongly reduces the volatility of the economy, as it strongly reduces the occurrence of deflationary spirals.

Taken together, these results point towards an important role of forward guidance for the stability and predictability of the macroeconomy, which we scrutinize more deeply in the following. In Section 4.1 we analyze the stabilizing role of forward guidance in normal times (Hypothesis 1(a)), in Section 4.2 we focus on the stabilizing role of forward guidance at the zero lower bound (Hypothesis 1(b)). In Section 4.3 we analyze the anchoring effect of forward guidance (Hypothesis 2). Subsequently, in Section 4.4 we study the influence of forward guidance on the predictability of the economy (Hypothesis 3).

4.1 Macroeconomic Stability in Normal Times

In this section, we focus on Hypothesis 1(a), i.e. we analyze in more detail the stabilizing role of central bank forward guidance for the economy in normal times, i.e., we focus entirely on Stage II. Macroeconomic stability is of utmost importance, as it can be directly linked to welfare in the economy. Woodford (2001) shows that minimizing the squared deviations of inflation and the output gap from zero, maximizes expected household utility and thereby welfare. Consequently, for each experimental economy $i = 1, ..., 24$, we evaluate macroeconomic stability by the mean squared deviations of inflation and the output gap from zero

$$S_i^\pi = \frac{1}{20}\sum_{t=9}^{28}\pi_t^2, \tag{11}$$

$$S_i^y = \frac{1}{20}\sum_{t=9}^{28}y_t^2. \tag{12}$$

The lower S_i^π and S_i^y the more stable the economy i. The results are summarized in Table 1. The last column of Table 1 shows the average mean squared error for each respective treatment. Informative forward guidance (Treatments

	Treatment 1						
Economy	E1	E2	E3	E4	E5	E6	Avg
S_i^π	1.9420	0.5922	2.1814	1.1505	3.1821	2.0837	1.8553
S_i^y	0.5560	0.2535	0.5261	0.2254	0.3396	0.3684	0.3782
	Treatment 2						
Economy	E7	E8	E9	E10	E11	E12	Avg
S_i^π	0.4984	0.1935	0.4289	1.4772	0.5132	0.7994	0.6518
S_i^y	0.2715	0.1937	0.2369	0.2759	0.2598	0.1610	0.2331
	Treatment 3						
Economy	E13	E14	E15	E16	E17	E18	Avg
S_i^π	0.3233	0.1720	0.1993	0.3390	0.3413	1.5116	0.4811
S_i^y	0.2441	0.1532	0.2420	0.1198	0.2967	0.3278	0.2306
	Treatment 4						
Economy	E19	E20	E21	E22	E23	E24	Avg
S_i^π	0.5679	0.1830	1.3334	1.8194	0.8692	3.0710	1.3073
S_i^y	0.1614	0.1534	0.4611	0.1230	0.4730	0.3410	0.2855

Table 1: Stage II mean squared deviations of inflation and the output gap. The table shows the stability measures given by equations (11) and (12) for all 24 economies, as well as their respective treatment averages.

2 and 3) dramatically reduces the average mean squared error for inflation by two thirds and the output gap by one third. These differences are statistically significant: the p-values of the Mann-Whitney-Wilcoxon-test for pairwise comparisons of Treatment 1 with Treatments 2 and 3 are $p_{(S^\pi_{T1},S^\pi_{T2})} = 0.0152$ and $p_{(S^\pi_{T1},S^\pi_{T3})} = 0.0087$ for inflation and $p_{(S^y_{T1},S^y_{T2})} = 0.0931$ and $p_{(S^y_{T1},S^y_{T3})} = 0.0649$ for the output gap. Random forward guidance, by contrast, has no statistically significant effect on macroeconomic stability, i.e., ($p_{(S^\pi_{T1},S^\pi_{T4})} = 0.2403$ and $p_{(S^y_{T1},S^y_{T4})} = 0.3095$). The stabilizing role of informative forward guidance manifests itself impressively through a much faster convergence of inflation towards the steady state of the economy. In Treatments 2 and 3, inflation reaches the close neighborhood of the steady state, say an interval of ± 25 basis points around the steady state, on average within 5 periods. In Treatments 1 and 4, time to convergence triples, with a third of the economies not reaching convergence at all during Stage II. All these results carry over, if stability is measured by the squared deviation from target rather than from zero.

The stabilizing role of informative forward guidance becomes even more pronounced when taking into consideration the influence of the first stage developments on the starting point of Stage II. In Stage I, Treatment 1 economies are on average at least as stable (measured analogously to (11) and (12)) as Treatment 2 and 3 economies. Although not statistically significant, they hand over the economy to Stage II even at slightly lower mean and median levels of inflation with the consequence that, if at all, Treatments 2 and 3 face a marginally unfavorable situation upon entering Stage II.

To account for the influence of Stage I stability on Stage II stability, for each economy i we normalize the Stage II mean squared error of inflation and the output gap by their respective mean squared errors from Stage I

$$R_i^\pi = \frac{S_i^\pi}{\frac{1}{8}\sum_{t=1}^{8}\pi_t^2}, \quad (13)$$

$$R_i^y = \frac{S_i^y}{\frac{1}{8}\sum_{t=1}^{8}y_t^2}. \quad (14)$$

The results are presented in Table 15 in Appendix C. First, the table implies that for each of the 24 economies medians are lower for all variables in Stage II relative to Stage I, which manifests in values smaller than unity. This results, however, is not surprising as the Taylor rule slowly drives the economy towards the steady state. Second, the table generally confirms our results from above. Informative forward guidance strongly helps stabilize the economy with respect to inflation and output gap relative to the absence of forward guidance, albeit not statistically significantly for the output gap in the case for Treatment 3. In the latter case the p-value of the Mann-Whitney-Wilcoxon-test is $p_{(R^y_{T1},R^y_{T3})} = 0.1797$. The remaining p-values are $p_{(R^\pi_{T1},R^\pi_{T2})} = 0.0411$ $p_{(R^\pi_{T1},R^\pi_{T3})} = 0.0043$ for Treatment 2 and 3 inflation and $p_{(R^y_{T1},R^y_{T2})} = 0.0649$ for the Treatment 2 output gap. Concerning random forward guidance, the picture changes somewhat. Relative to the Stage I development, random forward guidance stabilizes inflation slightly better compared to no forward guidance at all, but stabilizes the output gap less effectively. However, neither of these differences is statistically significant, i.e., $p_{(R^\pi_{T1},R^\pi_{T3})} = 0.4848$ and $p_{(R^y_{T1},R^y_{T4})} = 0.5887$. Again, all these results carry over, if stability is measured by the squared deviation from target rather than from zero.

The analysis above implies that informative forward guidance is an effective instrument to increase welfare through its stabilizing role in the economy. Random forward guidance, by contrast, remains without statistically significant effects on stabilization. As a result, the above analysis confirms Hypothesis 1(a).

4.2 Forward Guidance at the Zero Lower Bound

Now, we focus on Hypothesis 1(b), i.e. we analyze the impact of informative forward guidance in times of severe economic stress. To do so, we look at Stage III (periods 29-37) of the experiment.[24] Between periods 29 and 32, a series of severe shocks to the output gap (ε_t takes a value of -2.5% annually in periods $t = 29, ..., 32$) hits all 24 economies alike. Figure 1 and Figures 4 to 7 in Appendix C show the reaction of the macroeconomies to these shocks. In each case, a deflationary recession takes place, which drives the economy to the zero lower bound on the nominal interest rate. The severity of the economic downturn, however, can be mitigated when the central bank conducts informative forward guidance. This can be seen from Table 2 where we summarize important key indicators describing the median severity of the economic downturn in each of the four treatments. In the results description below, p-values of Mann-Whitney-Wilcoxon-tests are reported only if differences in medians are statistically significant.

[24] Be reminded that in Stage III of Treatment 4 the public inflation projection is produced by the "good" computer algorithm instead of the random number generator.

	T1	T2	T3	T4
Periods at zero lower bound	5	3	2.5	1.5
Length of recession	7.5	4.5	4	3
Depth of recession	-239.76	-4.28	-4.18	-4.22
Periods of deflation	8	6	7	6.5
Deflationary spirals	3	1	0	1
Credibility Index in $t = 28$	—	0.88	0.89	0.13

Table 2: Important key indicators for Stage III. The table shows treatment medians of key indicators describing the severity of the recession and the accompanying liquidity trap in Stage III.

Table 2 shows that informative forward guidance on average halves the median time spent at the zero lower bound, from 5 periods in Treatment 1 to less than 2.5 periods on average in Treatments 2-4. Secondly, the length of the recession[25] is significantly ($p_{(T1,T2:T4)} = 0.090$) reduced from 8 periods in Treatment 1 to less than 4 periods on average in Treatments 2-4. Also, the depth of the recession radically reduces in the presence of informative forward guidance. We measure the depth of the recession by comparing the latest pre-crisis output gap with the largest negative output gap during the crisis. In Treatment 1, the median depth is a loss in output gap of approximately -240 percent, whereas this loss is around -4 percent on average for Treatments 2-4. Prices, in all economies, fall, i.e., there is deflation. However, with 6.5 periods on average in Treatments 2-4 median deflation episodes are reduced by 1.5 periods relative to Treatment 1. All qualitative results carry over for pairwise comparisons of Treatment 1 to Treatments 2, 3, and 4.

Despite binding zero lower bounds and prolonged deflationary episodes, deflationary spirals are rare. However, they occur much more often in the absence of informative forward guidance than in the presence of informative forward guidance. In Treatment 1 three out of six economies result in a deflationary spiral after a series of severe fundamental shocks. While deflationary spirals can be avoided successfully in all six economies of Treatment 3, in both Treatments 2 and 4 one out of six economies result in a deflationary spiral. Therefore, forward guidance significantly ($p_{(T1,T2:T4)} = 0.060$) reduces the occurrence of deflationary spirals.

In the following, we examine the deflationary spiral in economy E11 of Treatment 2 in more detail. We believe that it provides an informative counterfactual that help understand the stabilizing role of forward guidance at the zero lower bound.

We argue that in this particular case the central bank forecaster not only failed to prevent the deflationary spiral, but to a large part powered the deflationary spiral through inadequate central bank forward guidance. Indicative evidence for this assertion can be found in Figure 2. The figure shows the time series for the data-driven forecast (black solid line), the required-for-target value (gray solid line), the published central bank projection (dashed line), realized inflation (dotted line), and the inflation target (horizontal line) for all six economies of Treatment 2. As is apparent from the lower left panel of Figure 2,

[25] According to the NBER, a recession is a drop in economic activity between peak and trough.

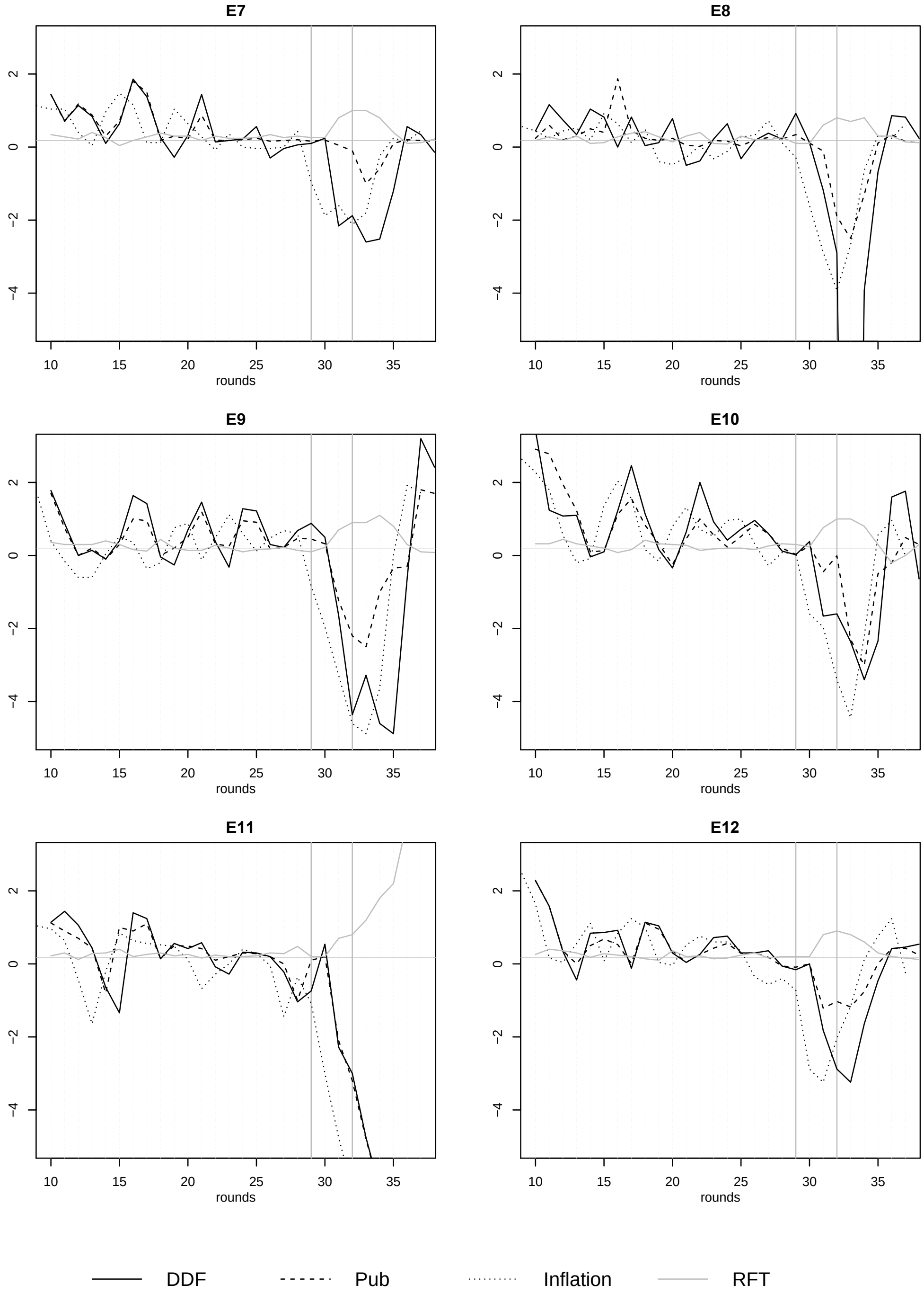


Figure 2: Time series for data-driven forecast (black solid line), public central bank inflation projection (dashed line), aggregate inflation (dotted line), and the "required for target" information (gray solid line) for all six experimental economies of Treatment 2.

Economy \ Period	31	32	33	34	35	36	37	Median	SD
E7	0.62	0.44	0.58	0.81	0.78	0.73	0.83	0.73	0.14
E8	0.27	0.89	0.56	0.81	0.93	0.99	1.00	0.89	0.27
E9	0.41	0.19	0.63	0.80	0.36	0.45	0.31	0.41	0.21
E10	0.61	0.02	0.09	0.70	1.00	0.72	1.00	0.70	0.40
E11	−0.05	−0.01	−0.01	−0.00	0.31	−0.24	−0.30	−0.01	0.20
E12	0.49	0.51	0.39	0.62	0.00	0.10	0.71	0.49	0.26

Table 3: "Strategic-ness" measure of Equation (15) for periods $t = 31, ..., 37$ of Stage III (as well as their respective medians and standard deviations) for each of the six economies of Treatment 2

rather than trying to strategically stabilize the economy through publishing over-optimistic inflation projections - that are inflation projections moving into the direction of the inflation target or even the necessary aggregate forecast as prescribed by the "required for target" criterion -, the central bank forecaster followed the advice of the data-driven forecast and thereby publicly predicted the deflationary spiral. The central bank forecasters of the other five economies, by contrast, do not publicly predict a deflationary spiral, but resort to over-optimistic inflation projections and pull the economy out of the slump.

To quantify the above argument, we define a measure of "strategic-ness" of the published central bank projection, denoted by SP_t below. The measure illustrates in any given period to what extent the published inflation projection follows the data-driven forecast (non-strategic behavior) and to what extend it tries to strategically influence inflation expectations in the direction consistent with the economy's determinate target steady state. The latter can be measured by how much the central bank uses either the "required for target" tool or the inflation target. In the following, we report results applying the "required for target" tool. All results are robust to applying the inflation target instead. The explicit "strategic-ness" measure takes the following form

$$SP_t = \frac{E_t^{pub}\pi_{t+1} - E_t^{ddf}\pi_{t+1}}{E_t^{rft}\pi_{t+1} - E_t^{ddf}\pi_{t+1}}. \tag{15}$$

If $SP_t = 1$ the published projection coincides with the "required for target" forecast, whereas if $SP_t = 0$ the published projection coincides with the data-driven forecast.[26] Table 3 presents the measure of "strategic-ness" for the 7 periods after the shocks have died out in all six economies of Treatment 2 along with their median and standard deviation. For all five economies without a deflationary spiral SP_t for $t = 31, ..., 37$ is mostly substantially above zero, with median values in the interval $[0.41, 0.89]$, which are always significantly different from zero given the Wilcoxon signed rank test ($p \leq 0.02$ in each case). This implies that in these economies the central bank forecaster does not just follow

[26]The index can also take values above unity and below zero. If $SP_t > 1$ the published projection lies outside the interval of the data driven forecast and the "required for target" information, on the side of the "required for target". This implies that the central bank is trying to steer more than necessary to achieve the target. If $SP_t < 0$, the published projection lies outside the band of the data driven forecast and the "required for target" information, on the side of the data driven forecast. This implies that the central bank CB tries to drive expectations "away" from target. A proof of these claims is presented in Appendix B.

trends or the data-driven forecast, but that she considerably tries to steer the economy towards target. By contrast, in the economy with deflationary spiral SP_t is mostly very close to zero with a median value of -0.01. A median of zero cannot be rejected by the Wilcoxon signed rank test ($p = 0.4$) in this case. This implies that the central bank forecaster is not concerned with steering the economy back to the target, but solely with giving "accurate" projections. Moreover, the negative sign of the "strategic-ness" measure indicates that the central bank forecaster predicts an even more extreme deflationary spiral relative to what is predicted by the data-driven forecast (see Appendix B).

The stabilizing role of forward guidance at the zero lower bound is particularly surprising, since at the zero lower bound an overoptimistic (or strategic) inflation projection must be considered cheap talk. At the zero lower bound, the central bank has no means to actively support the public projection using the interest rate. We believe that the evidence presented in this section is a confirmation of Hypothesis 1(b). Furthermore, our results support the finding by Duffy and Heinemann (2014), that cheap talk can be a very successful strategy for the central bank to achieve its stabilization goals.

4.3 Anchoring Effect of Forward Guidance

In the following we address Hypothesis 2, i.e. we analyze the anchoring effect of forward guidance. Eusepi and Preston (2010) argue that in economies with potentially self-fulfilling expectations and learning, such as the economy in the present experiment, central bank projections generate macroeconomic stabilization through the anchoring of private-sector expectations on a path consistent with monetary policy. In the absence of inflation projections, expectation may be unanchored and even be inconsistent with monetary policy. Therefore, unlike under rational expectations, the Taylor principle alone does not guarantee macroeconomic stability.

The anchoring effect manifests itself in a lower disagreement among individual professional forecasters in the presence of forward guidance relative to the absence of forward guidance. We measure disagreement among individual professional forecasters by the cross-sectional dispersion of individual professional forecasts in each period t, using three alternative dispersion measures which are commonly found in the literature; the variance as in Fujiwara (2005), the distance between the highest and the lowest forecast (henceforth: range) as proposed by the FED, and the inter-quartile range of forecasts in any given period as in Ehrmann et al. (2012); Hubert (2014). Table 4 presents the median values of all three measures for each of the single economies and the respective averages for each of the four treatments.[27] The table shows that for all three measures considered informative forward guidance (Treatments 2 and 3) reduces the disagreement among individuals roughly by one third. Not surprisingly, the random forward guidance (Treatment 4) increases disagreement considerably, almost doubling the dispersion of individual forecasts. According to the Mann-Whitney-Wilcoxon test the above mentioned differences are highly statistically significant at the 1% significance level. To account for Stage I influences, analogously to the previous stability analysis, Table 16 in Appendix C presents the dispersion measures in Stage II relative to their counterparts in Stage I.

[27] The results are the same for the mean in stead of the median of the dispersion measures.

	Dispersion Measure: Variance						
Economy	1	2	3	4	5	6	Avg
Treatment 1	0.06	0.07	0.22	0.07	0.40	0.27	0.18
Treatment 2	0.09	0.01	0.09	0.24	0.04	0.11	0.10
Treatment 3	0.15	0.06	0.04	0.02	0.07	0.19	0.09
Treatment 4	0.45	0.05	0.12	0.41	0.41	2.02	0.58
	Dispersion Measure: Range						
Economy	1	2	3	4	5	6	Avg
Treatment 1	0.61	0.75	1.30	0.66	1.76	1.42	1.08
Treatment 2	0.82	0.27	0.65	1.25	0.54	0.84	0.73
Treatment 3	0.94	0.70	0.55	0.40	0.62	1.13	0.73
Treatment 4	1.68	0.62	0.85	1.64	1.89	3.56	1.71
	Dispersion Measure: Interquartile Range						
Economy	1	2	3	4	5	6	Avg
Treatment 1	0.25	0.24	0.39	0.19	0.43	0.60	0.35
Treatment 2	0.29	0.12	0.30	0.56	0.19	0.30	0.29
Treatment 3	0.26	0.16	0.24	0.10	0.29	0.44	0.25
Treatment 4	0.69	0.21	0.47	0.81	0.61	1.23	0.67

Table 4: Medians of Stage II Dispersion Measures. The table shows the medians of three dispersion measures from periods $t = 9, ..., 28$ for all six economies in each treatment and the respective treatment averages. Dispersion in economy i in period t is either the variance, range, or interquartile range of the period t inflation forecasts for period $t+1$ from all six individual professional forecasters in economy i.

The numbers in Table 16 confirm the previous results that informative central bank forward guidance more successfully reduces the disagreement among the individual professional forecasters.

To quantify the anchoring effect of public inflation projections, in the spirit of Ehrmann et al. (2012) and Hubert (2014), we resort to a simple regression analysis of the form

$$\sigma_{fc,t} = constant + \beta_1 PP_t + \beta_2 \sigma_{fc,t-1} + \beta_3 X_{t-1} + \varepsilon_t, \tag{16}$$

where $\sigma_{fc,t}$ is the cross-sectional dispersion of the professional forecasters in period t, PP_t is a dummy variable which takes value 1 when a public inflation projection is present and X_t is a vector of macroeconomic controls. The macroeconomic controls X_{t-1} comprise the lagged interest rate, the lagged output gap, and lagged inflation uncertainty defined by $IU_{t-1} = |\pi_{t-1} - \pi_{t-2}|$, which is the absolute error of a random walk forecast (Ahrens and Hartmann, 2015). We expect a positive relationship between lagged inflation uncertainty and cross-sectoral dispersion. The higher lagged inflation uncertainty, the harder the prediction of inflation and thereby the greater the cross-sectional dispersion (Capistrán and Timmermann, 2009; Dovern and Hartmann, 2016). Concerning the remaining control variables, first, we expect cross-sectoral dispersion to be positively influenced by the lagged interest rate. According to the Taylor rule the interest rate increases in inflation. Mankiw et al. (2004) show that a higher level of inflation, in turn, yields more disagreement in inflation expectations. For the lagged output gap we expect a negative relationship, since Dovern et al.

	[1]	[2]	[3]	[4]	[5]
constant	-0.229	0.118	0.179**	0.080	-0.248
	[0.162]	[0.083]	[0.078]	[0.102]	[0.203]
PP_t	0.044	-0.101***	-0.126***	-0.084**	0.286***
	[0.056]	[0.037]	[0.043]	[0.034]	[0.085]
$\sigma_{fc,t-1}$	0.228***	0.130**	0.136**	0.125*	0.138*
	[0.054]	[0.059]	[0.068]	[0.075]	[0.079]
r_{t-1}	0.089***	0.025	0.012	0.030	0.108***
	[0.030]	[0.016]	[0.015]	[0.021]	[0.038]
y_{t-1}	-0.028	-0.026	0.004	-0.018	0.045
	[0.037]	[0.030]	[0.036]	[0.036]	[0.074]
IU_{t-1}	0.298***	0.340***	0.371***	0.382***	0.289***
	[0.067]	[0.064]	[0.082]	[0.091]	[0.100]

Table 5: Anchoring effect of forward guidance. The table shows the results from estimating equation (16) for different subsamples of the experimental data. The respective samples are: [1] T1-T4; [2] T1-T3; [3] T1 vs T2; [4] T1 vs T3; [5] T1 vs T4

(2012) and Hubert (2014) document a higher disagreement in recessions. The parameter estimates are summarized in Table 5.

Column [1] in Table 5 shows the results when all four treatments (T1,T2,T3,T4) are considered. In this case, the table shows that the publication of inflation projections per se has no anchoring effect, i.e., PP_t is close to zero and statistically insignificant as are the interest rate and output gap coefficients. For the complete set of data, cross-sectional dispersion is a persistent phenomenon which is mainly driven by inflation uncertainty. In Columns [2]-[5] we distinguish between informative and random forward guidance. While Columns [2]-[4] show variants which abstract from random forward guidance (Treatment 4), Column [5] abstracts from informative forward guidance. Consider Columns [2]-[4] first. Column [2] shows the parameter estimates of (16) using data from Treatments 1 to 3, Column [3] using data from Treatments 1 and 2 and Column [4] using data from Treatments 1 and 3. First, the table shows that parameter values generally have the expected sign. More importantly, informative forward guidance unambiguously reduces the cross-sectional dispersion of individual expectations. The reduction is statistically significant. The influence of inflation uncertainty on the cross-sectional dispersion remains statistically significant. The interest rate and the output gap coefficients again are negligible and statistically insignificant. Finally, Column [5] shows the parameter estimates of (16) using data from Treatments 1 and 4 only. Now, the effect of publishing inflation projections is positive and statistically significant. Random forward guidance increases the cross-sectional dispersion by approximately 29%. The results are similar, if we consider contemporaneous macroeconomic controls X_t, as applied in the original studies by Ehrmann et al. (2012) and Hubert (2014).

Taken together, the above results give rise to the notion that informative forward guidance acts as an anchor for private-sector inflation expectations, while random forward guidance unleashes disturbing forces driving private-sector expectations apart. Therefore, our evidence confirms Hypothesis 2. Furthermore, the evidence documents a substantial influence of central bank forward guidance on private-sector expectation formation. Section 5 analyzes this influence in more detail.

4.4 Forecasting Performance

In this section, we address Hypothesis 3, i.e. we analyze the relative forecasting performance of the subjects, the data-driven forecast, and the computer-algorithms. We evaluate the relative forecasting performance by means of the mean squared error in Stage II of the experiment.

Table 17 in Appendix C presents the mean squared errors of the published central bank projection, the central bank forecaster subject's forecast, the data-driven forecast, and the aggregate forecast of the professional forecasters for each economy of all four treatments. Analogously, Table 18 in Appendix C presents the mean square errors of the individual professional forecasters. Following Romer and Romer (2000) and Hubert (2015c), p-values for pairwise comparisons of the mean squared errors are calculated by estimating

$$(\pi_{t+1} - E_t^a \pi_{t+1})^2 - (\pi_{t+1} - E_t^b \pi_{t+1})^2 = c + u_{t+1}, \tag{17}$$

where $E_t^a, E_t^b \in \{E_t^{pub}\pi_{t+1}, E_t^{cbf}\pi_{t+1}, \bar{E}_t\pi_{t+1}, E_t^{ddf}\pi_{t+1}\}$. The p-values test the null hypothesis that $c = 0$. The standard errors are corrected for autocorrelation and heteroskedasticity using the Newey-West HAC method (Newey and West, 1987).

The tables show that for Treatment 1, the central bank forecaster, the aggregate inflation forecast, and the data-driven forecast seem to do equally well in most cases, as well as on average. In four out of six economies, there cannot be found a significant ranking of the forecasts. For the remaining two economies, there can be found the worst forecasting entity, but not the best forecasting entity. On average, in Treatment 1 there is no (or only very little) evidence for the superiority of one forecasting entity over the other. If at all, the central bank forecaster subject does worst. Given the fact that the central bank has more and better information about potential future inflation than the professional forecasters, it is somewhat surprising that the central bank forecaster performs no better (rather slightly worse) than the aggregate forecast of the professional forecasters. Two explanations come to mind. First, being rather persistent, the aggregate forecast becomes highly self-fulfilling and thereby accurate by construction. A second potential explanation can be found in the "wisdom of the crowd," which describes the phenomenon that groups can achieve higher forecast accuracy by taking the group average or median compared to their individual forecasts, as the mean or median filters out idiosyncratic noise (Surowiecki, 2005). The latter argument is also supported by the fact that (in all four treatments) the mean squared error of the aggregate forecast of the professional forecasters for a treatment is always below the average mean squared error of all individual forecasters within that treatment. Moreover, in 21 out of 24 economies at most one individual forecaster performs better individually (has a lower mean squared error) than the aggregate forecast in that respective economy.[28] The remaining three economies feature at least 2 and at most 3 individual forecasters who perform better individually than the aggregate forecast.

[28]In 9 out of these 21, no individual performs better than the aggregate forecast.

A similar insignificant pattern amongst the central bank forecaster, the aggregate inflation forecast, and the data-driven forecast arises in Treatment 4. The computerized random published projection in Treatment 4, not surprisingly, is considerably less accurate compared to all other forecasting entities with these differences being statistically significant.

By contrast to Treatment 4, for Treatments 2 and 3 the published inflation projection improves substantially and performs significantly better than any other forecasting entity, both on average and for most of the individual economies. The performance of the data-driven forecast remains basically unchanged (for pairwise comparisons the p-value is never below 0.10). The aggregate inflation forecast of the professional forecasters improves significantly only when a good computerized central bank projection is provided ($p_{afc_{T1},afc_{T3}} = 0.0579$), whereas the improvement is not significant when the human central bank provides the projections ($p_{afc_{T1},afc_{T2}} = 0.5542$). For the individual professional forecasters, the average reduction of the mean forecast error under informative forward guidance is substantial (approximately one third) and statistically significant independent of whether the central bank is computerized or not, i.e. $p_{fc_{T1},fc_{T2}} = 0.000$ and $p_{fc_{T1},fc_{T3}} = 0.000$. The central bank forecaster improves her forecasting performance considerably (by more than 50%) and significantly ($p_{cbf_{T1},cbf_{T2}} = 0.0745$) when her forecast is published. Summing up, the above evidence confirms Hypothesis 3.

The result that the central bank forecasting performance is improved when its forecast is published is particularly interesting in the light of analyzing the potential source of the well-documented superiority of central bank projections over private-sector expectations. Romer and Romer (2000) put forth the hypothesis that the FOMC is able to produce superior inflation projections from publicly available information projections simply by committing far more resources to forecasting compared to the private sector. The results from Treatments 1 and 2 in our experiment reject this hypothesis. In both treatments, an equal amount of resources is invested to process the publicly available information and to provide it to the central bank forecaster in the form of the data-driven forecast,[29] yet the forecasting performance of the central bank forecaster differs substantially. Three potential explanations may be put forward for this difference: the quality of the information supplied to the central bank forecaster, the different incentive structure for the central bank forecaster across both treatments, and the publication of the inflation projection. Since the predictive power of the data-driven forecast is similar across Treatments 1 and 2, information quality cannot explain the difference. Also, the central bank forecaster's incentive cannot explain this result. The central bank forecaster's incentive to predict inflation accurately is stronger in Treatment 1 relative to Treatment 2 (as in Treatment 2 there is a trade-off between predicting correctly and strategically). Consequently, a better forecasting performance should be expected in Treatment 1 and not the other way around. By contrast, the publication of the inflation projection is a potential explanation for the improved forecasting performance. If credible, central bank projections are self-fulfilling. That is, when the central bank publishes a high (low) projection and professional forecasters respond by giving a forecast in the same direction, then a high (low)

[29]Note that the "required for target" information supplied to the central bank forecaster in Treatment 2 has no predictive content for potential future inflation and thereby should not be included in the information set used for accurate forecasting.

rate of inflation realizes. Consequently, even ex-ante incorrect projections, if believed by the professional forecasters, are evaluated as quite accurate projections ex-post. Whether or not central bank projections indeed positively affect private sector expectations and hence are (at least in part) self-fulfilling, we explore in more detail in the next section.

5 Expectation Formation

In this section we analyze how individual subjects form their expectations. We consider the professional forecasters (Section 5.1) and the central bank forecasters (Sections 5.2) separately. In the latter section, we also address Hypothesis 4.

5.1 Professional Forecaster Expectation Formation

In the four treatments of the experiment we can distinguish two types of professional forecasters: professional forecasters that were not exposed to a published projection, and professional forecasters that did see a published projection prior to submitting their own forecast. In this section we investigate (i) whether these two groups of forecasters formed expectations in a qualitatively different way and (ii) to what extent the expectation formation of the forecasters that did see a central bank projection depended on the quality of this projection. Since we treat Stage I as learning stage in all treatments and Stage III presents subjects with an inherently unstable environment, we focus this analysis on Stage II only.

First, consider professional forecasters that did not see a published projection when making their own forecast. This group consists of the professional forecasters in Treatment 1 (the control treatment). We follow Assenza et al. (2013) and Pfajfar and Zakelj (2014) and regress each subject's inflation forecast on a general linear forecasting rule of the form

$$E_t^{fc,j}\pi_{t+1} = c^j + \sum_{i=1}^{2} \alpha_i^j E_{t-i}^{fc,j}\pi_{t+1-i} + \sum_{i=1}^{2} \beta_i^j \pi_{t-i} + \gamma^j y_{t-1} + \varepsilon_t^j, \tag{18}$$

where ε^j is the error term of each individual regression. The results are summarized in the second column of Table 6, which show the percentage of individually significant regressors and the median estimated parameter values for each treatment, respectively.[30] The second column of Table 6 shows that 92% of subjects consider the first lag of inflation when forming their expectation about future inflation. 36% of subjects also consider the second lag of inflation. Given that the sign of the coefficient on the first lag is generally positive with a median of 1.11, while the sign on the second lag of inflation is generally negative with median of -1.14 it appears that the professional forecasters engaged in trend following behavior when forecasting inflation. In line with early evidence from Adam (2007) only few subjects consider past realizations of the output gap to predict future inflation.

Next, we consider all subjects which were shown a public central bank projection prior to submitting their own forecast. This group consists of all subjects

[30] In the estimation we follow Massaro (2012) by iteratively eliminating all insignificant regressors. The details of the procedure are presented in Appendix A.

Treatment	1	2	3	4
constant	39%	36%	56%	50%
	(0.431)	(0.381)	(0.162)	(0.807)
$E_{t-1}^{fc,j}\pi_t$	14%	19%	14%	19%
	(0.429)	(0.140)	(0.395)	(0.550)
$E_{t-2}^{fc,j}\pi_{t-1}$	3%	11%	17%	8%
	(-0.734)	(-0.479)	(-0.385)	(-0.369)
π_{t-1}	92%	47%	56%	42%
	(1.105)	(0.617)	(0.744)	(0.813)
π_{t-2}	36%	25%	17%	11%
	(-1.140)	(-0.553)	(-0.006)	(-0.586)
y_{t-1}	14%	11%	14%	25%
	(-1.055)	(0.971)	(0.348)	(1.350)
$E_t^{pub}\pi_{t+1}$		69%	31%	31%
		(0.818)	(1.441)	(0.216)
avg. R^2	0.76	0.72	0.66	0.46
#Sign.Coeff	1.97	2.19	2.03	1.86

Table 6: Percentages of significant regressors and the median regression coefficients (in parentheses) estimating equations (18) and (19) for all professional forecasters per treatment. Additionally, the table shows the average R^2 and the average number of significant coefficients per forecaster for each treatment.

in Treatment 2, 3, and 4. We follow the same procedure as above, but with one difference. Now, we include the published central bank inflation projection as an additional regressors in the set of possible regressors. Thus, the subject's new general linear forecasting rule has the form

$$E_t^{fc,j}\pi_{t+1} = c^j + \sum_{i=1}^{2}\alpha_i^j E_{t-i}^{fc,j}\pi_{t+1-i} + \sum_{i=1}^{2}\beta_i^j\pi_{t-i} + \gamma^j y_{t-1} + \delta^j E_t^{pub}\pi_{t+1} + \varepsilon_t^j. \quad (19)$$

Column 3 of Table 6 shows the results of the regression on the subjects with a published projection provided by a human central banker. It can be seen that for 69% of the subjects the published projection has a statistically significant effect on their expectations. This is more than for the first lag of inflation which is now statistically significant for less than half of the subjects. The significance of the second lag of inflation is also reduced considerably. Surprisingly, when the public central bank projection is given by a computer algorithm, it is statistically significant only for 31% of the subjects (Columns 4 and 5). Nevertheless, the first and second lag of past inflation lose significance compared to the control treatment. We conclude from this that when subjects are presented with a published central bank projection, many subjects let their own forecast be affected by the public projection. In this case, subjects put less weight on past inflation and trend behavior in inflation in particular.

The bottom row of Table 6 presents the average number of significant regressors used in the expectation formation process in each of the four treatments. Interestingly, this number is around two for all of the four treatments. This leads to the conclusion that subjects rather substitute the public central bank inflation projection for another source of information than complement their information set in the expectation formation process.

Another interesting observation is that the fraction of subjects using the published inflation projection is the same, independent of whether the computer algorithm is sophisticated and good or bad (Treatment 3 versus 4).[31] However, even though the fraction of subjects considering the public projection in Treatments 3 and 4 are the same, we find that the weight subjects put on this projection when forming their expectations is much higher for Treatment 3 compared to Treatment 4. The median of the significant coefficients on the published projection in Treatment 3 is 1.44, whereas it is 0.22 in Treatment 4. The coefficient in Treatment 3 is furthermore also bigger than in Treatment 2, where the median coefficient is equal to 0.818.

5.2 Central Bank Forecaster Expectation Formation

Now, consider the expectations of the central bank forecaster subjects. Analogous to the previous analysis, for each central bank forecaster subjects j we estimate a general linear forecasting rule. To account for the additional information supplied to the central bank forecaster subject, the general forecasting rule takes the following form:

$$\begin{aligned} E_t^{cbf,j}\pi_{t+1} =& c^j + \sum_{i=1}^{2} \alpha_i^j E_{t-i}^{cbf,j}\pi_{t+1-i} + \sum_{i=1}^{2} \beta_i^j \pi_{t-i} + \gamma^j y_{t-1} \\ & + \delta_1^j E_t^{ddf}\pi_{t+1} + \delta_2^j E_t^{rft}\pi_{t+1} + \varepsilon_t^j. \end{aligned} \tag{20}$$

Note that for subjects in Treatments 1 $\delta_2^j = 0$ and that for subjects in Treatments 3 and 4 $\delta_1^j = \delta_2^j = 0$.[32] Table 7 summarizes the percentages of significance, which is now based on 6 observations per treatment.

The central bank forecasters of Treatments 3 and 4 face the same decision as the professional forecasters in Treatment 1, with the only difference that the central bank forecasters have no influence whatsoever on the economy. Since the influence of an individual professional forecaster is relatively small, we should expect to see a similar expectation formation process. Indeed, from columns 4 and 5 of Table 7, it can be seen that most subjects have a significant (positive) coefficient (with medians of 1.01 in Treatment 3 and a median of 0.93 in Treatment 4) on the first lag of inflation and that many subjects also have a negative coefficient (with medians of -0.73 in Treatment 3 and a median of -0.68 in Treatment 4) on the second lag of inflation.

When facing the same decision as in Treatments 3 and 4, but when presented with the data-driven forecast (Treatment 1), the central bank forecaster resorted strongly to this source of information in their expectation formation process. This is in line with the result of the previous section that professional forecasters that got a published central bank projection in each period partly substituted this forecast for their own trend extrapolations. This results seems even stronger for the central bank forecaster in Treatment 2. Five out of the six subjects get

[31] Interestingly, when we split the sample between subjects with high cognitive ability and subjects with low cognitive ability this is no longer the case (see Table 19 in Appendix C). We measure cognitive ability with the three-item "cognitive reflection test" of Frederick (2005). Subjects with a higher CRT use the published projection more often when it is informative (Treatments 2 and 3), while they use it less often when it is random (Treatment 4). Therefore, subjects with a high CRT score can be seen as more "rational."

[32] The estimation procedure again follows the procedure described in the Appendix. The order of removal for equation (20) is: $\alpha_2, \gamma_1, \beta_2, \delta_2, \delta_1, \beta_1, \alpha_1, c$.

Treatment	1	2	3	4
constant	17%	33%	17%	67%
	(0.338)	(0.200)	(0.123)	(0.682)
$E^{cbf,j}_{t-1}\pi_t$	33%	33%	0%	0%
	(0.597)	(0.323)	(0.000)	(0.000)
$E^{cbf,j}_{t-2}\pi_{t-1}$	0%	0%	3%	17%
	(0.000)	(0.000)	(-0.155)	(-0.373)
π_{t-1}	50%	0%	100%	83%
	(0.792)	(0.000)	(1.014)	(0.927)
π_{t-2}	33%	0%	50%	33%
	(-0.887)	(0.000)	(-0.678)	(-0.773)
y_{t-1}	0%	0%	0%	0%
	(0.000)	(0.000)	(0.000)	(0.000)
$E^{ddf}_{t-1}\pi_t$	50%	83%		
	(0.984)	(0.660)		
$E^{rft}_{t-1}\pi_t$		0%		
		(0.000)		
avg. R^2	0.66	0.63	0.64	0.52
#Sign.Coeff	1.83	1.5	2	2

Table 7: Percentage of significant regressors and the median regression coefficients (in parentheses) from estimating equations (18) and (19) for all central bank forecasters per treatment. Additionally, the table shows the average R^2 and the average number of significant coefficients per forecaster for each treatment.

a significant coefficient on the data-driven forecast and past inflation is never significant.[33]

It is further noteworthy that not a single significant coefficient on the "required for target" is obtained. To investigate in more detail to what extent the central bank forecaster in Treatment 2 made use of the data-driven forecast rather than the "required for target", we pool the observations of all 6 subjects together and estimate the following regression.[34]

$$E^{cbf}_t\pi_{t+1} = b_1 E^{rft}_t\pi_{t+1} + (1-b_1)E^{ddf}_t\pi_{t+1} \quad (21)$$

This regression can give insight in the conditions under which subjects expectations where more in line with the data-driven forecast or more with the "required for target", and sheds light on Hypothesis 4. Specifically, we look at how this depends on credibility, as measured by the credibility index, and on how close inflation was to the target in the previous period. We do this by splitting the sample of observations according to the credibility index and according to the distance from target.

Rows 2 and 3 show the results for estimating equation (21) split in two subsamples. The subsample for Row 2 collects all observations for which the credibility index is larger than its sample median. Row 3 collects all observations for which the credibility index is lower than or equal to the sample median. As

[33]For the sixth subject we do not obtain a significant coefficient on any of the regressors.

[34]We also estimate regressions where we add past inflation and a constant as extra regressors. However, both of them do not obtain coefficients that are significant at a 5% level, while the data-driven forecast and the required for target both are significant at a 1% level. It therefore seems that it is not necessary to add these controls.

Sample	b_1	$n.obs$
Full Sample	0.297***	120
	$(0.215, 0.379)$	
High Credibility	0.301***	56
	$(0.187, 0.414)$	
Low Credibility	0.295***	64
	$(0.178, 0.412)$	
Far From Target	0.287***	59
	$(0.160, 0.413)$	
Close To Target	0.374***	61
	$(0.256, 0.492)$	
CTT + LC	0.394***	27
	$(0.321, 0.468)$	
CTT + HC	0.354**	33
	$(0.139, 0.570)$	
FFT + LC	0.261**	37
	$(0.088, 0.435)$	
FFT + HC	0.365***	23
	$(0.224, 0.507)$	

Table 8: Weights on "required for target" and data-driven forecast in Expectation Formation Process. The table shows the results of estimating equation (21) for the full sample and for several subsamples.

can be seen, the relative weights on the "required for target" and the data-driven forecast change only negligibly. Thus, credibility seems to have only a negligible influence on the central bank forecasters' decision whether to follow the data-driven forecast or the "required for target."

Rows 4 and 5 show the relative weights, when the sample is split according to the distance of inflation from the target. For observations close to the target, the estimated weight on the "required for target" (0.374) is considerably higher than for observations further away from the target (0.287). This may indicate that when inflation is far from the target, subjects are mainly occupied with predicting correctly, but that when the target comes in sight they start aiming more to achieve this target. Due to a limited number of observations the difference is however not statistically significant.

Finally, for the last four rows of Table 8 the sample is split in four quadrants according to both the credibility index and the distance of inflation from target. This analysis seems to indicate that when inflation is close to the target credibility does not play much of a role, but when inflation is far from the target, less credibility implies less weight on "required for target" and more weight on the data-driven forecast. The differences are, yet again, not statistically significant.

The above above seems to indicate that there might be some truth to Hypothesis 4, at least when inflation is far away from its target. However, because of a small sample size, we are not able to confirm or reject the hypothesis.

6 Credibility of the Central Bank Projections

In the stability analysis in Section 4 we have not explicitly considered the role of the credibility of the published central bank projections. We now turn to this issue in more detail. In Section 6.1 we consider whether credibility of central bank projections depends on past performance of the projections (Hypothesis 5(a)), while in Section 6.2 we look at the influence of credibility on the central bank's ability to stabilize the economy (Hypothesis 5(b)).

6.1 How Past Performance Shapes Future Credibility

First, we address Hypothesis 5(A) by studying whether past performance of public inflation projections determines the credibility of future inflation projections. To analyze this, we follow Mokhtarzadeh and Petersen (2016) and estimate a series of probit models, where the dependent variable $Utilize_t$ is binary taking value 1 if individual professional forecasters utilized the central bank projection and 0 if not. A central bank projection is said to be utilized if an individual professional forecasters forecast is within 5 basis points of the respective central bank projection.[35] In accordance with Mokhtarzadeh and Petersen (2016), our explanatory variable is past forecasting performance of the central bank projections, measured by the absolute[36] forecast error from the previous period. As controls we employ the absolute deviation of previous inflation from the central bank's inflation target, the professional forecasters previous absolute forecast error, period $t-2$ utilization of the central bank projection, and the interaction of the latter two. The interaction term measures the degree to which past shaken confidence in the central bank projection influences the willingness to utilize the central bank projection in the future. Additionally, we control for past aggregate credibility of the central bank projection measured by the period $t-1$ credibility index, and the subjects cognitive ability measured by the three-item "cognitive reflection test" of Frederick (2005). The estimation results for Stage II from Treatments 2, 3, and 4 are presented in Table 20 and for Stage III in Table 21. Both tables are located in Appendix C.

The tables show that central bank projections are more likely to be adopted in the future, if they were accurate in the past, independent of whether the economy functions in normal times or in times of severe economic stress. Consequently, credibility increases in past forecasting performance, confirming Hypothesis 5(a). Additionally, credibility also seems to be a persistent phenomenon. If a professional forecaster adopted the central bank projection in the past or if it was credible in the past, the professional forecasters are more likely to adopt the central bank projection in the future. Even if the adoption of a past projection ex-post turns out to be a disappointment, i.e., it resulted in an own large forecast error, the willingness of the professional forecaster to adopt future central bank projections remains unchanged, which can be read from the insignificance of the interaction term in Tables 20 and 21. Subjects seem to pay more attention to the performance of the central bank projections than to reflect on their past behavior and its outcomes. Finally, in line with evidence presented in footnote

[35] Mokhtarzadeh and Petersen (2016) choose a band of 2 basis points to identify utilization of the central bank projection, which yields approximately 20% of private forecasts to utilize the central bank projection in their experiment. In our experiment, a 2-basis-point band yields a utilization of only around 7.5%, whereas a 5-basis-point band yields around 17.5% utilization. The increased number of observations in the 5-basis-point case does not change the qualitative results of the estimation, but results in stronger statistical significance.

[36] Results do not change if forecast errors are squared. Only exception is that the interaction term gains significance. Results are available from the authors on request.

31, the probit regressions reveal that cognitive ability increases the likelihood to adopt central bank projection.

6.2 The Role of Credibility for Stabilizing Forward Guidance

Next, we analyze how credibility of central bank projections affects the central bank's ability to stabilize the economy.

Comparing Treatments 3 and 4 when entering Stage III yields a natural test for Hypothesis 5(b), at least for times of severe economic stress. Right before entering the recession in period 29, central bank projections in Treatment 3 are highly credible (credibility index is $I_{28}^{cred} = 0.89$ and the median subjective survey measure is 7), whereas they are not credible at all in Treatment 4 (credibility index is $I_{28}^{cred} = 0.125$ and the median subjective survey measure is 3). Therefore, credibility of central bank projections is highly different (and statistically significantly so $p_{T3,T4}^{survey} = 0.005$ and $p_{T3,T4}^{I^{cred}} = 0.005$) upon entering the deep recession initiated by the series of fundamental shocks. Despite this large difference in credibility, the behavior of the economy in the time of severe economic stress is basically identical, as documented by the important key indicators describing the median severity of the economic downturn in Treatments 3 and 4 from Table 2. Therefore, low credibility of the inflation projection upon entering the recession does not hamper the stabilizing role of informative forward guidance at the zero lower bound. Consequently, this evidence rejects Hypothesis 5(b) for times of severe economic stress. Whether or not credibility has an influence on the stabilizing power of forward guidance in normal times is subject to further research.

7 Conclusion

In this paper, we study the stabilizing role of central bank forward guidance in a standard New Keynesian learning-to-forecast laboratory experiment. Subjects take the role of "professional forecasters" in the private sector who form one-period ahead inflation forecasts. Subjects are provided with a limited understanding of the true data generating process and a public central bank inflation projection, i.e. central bank forward guidance. We show that central banks can manage private-sector expectations via the publication of (strategic) central bank inflation projections and that such expectations management can successfully be applied as an additional monetary policy instrument to stabilize the economy.

In particular, we show that such central bank forward guidance considerably influences the subjects' expectations formation process. In the absence of forward guidance, subjects expectation formation process is well characterized as mostly backward-looking with simple trend following. In the presence of forward guidance, by contrast, the public inflation projection becomes an influential piece of information which starkly diminishes the prevalence of backward-looking expectation formation. The utilization of the central bank projections, albeit very persistent, is not unconditional. Weak past performance, which manifests in large forecast errors, reduces future credibility of the public projection and therewith the subjects' probability of future utilization.

The documented influence on expectations by means of forward guidance strongly impacts on macroeconomic activity. We show that whether this impact is stabilizing or destabilizing yet again depends on the quality of the published forecast. During normal times and given reasonable and informative public projections, the economy quickly converges towards its steady state and subsequently fluctuates around it closely. Random inflation projections, by contrast, are generally harmful to the economy as they unleash disturbing forces which give rise to large fluctuations of the economy. In times of severe economic stress at the zero lower bound, the publication of overly optimistic central bank projections turns out to strongly reduce the risk of deflationary spirals. Pessimistic central bank projections, however, have the potential to fuel or even initiate a deflationary spiral.

Finally, the increase in economic stability due to informative forward guidance has positive effects on the predictability of the economy. Professional forecasters increase their forecasting performance significantly and the disagreement among forecasters diminishes. By contrast, the aversive effects of random forward guidance unleash disturbing forces which give rise to more dispersed and less precise individual private-sector forecasts.

Our results have important implications for central bank practice. We show that central bank forward guidance is a powerful tool for stabilization policy in normal times and at the zero lower bound. However, while a good track record of accurate forecasts is important for credibility, we find that some strategicness in the published forecasts greatly enhances the stabilizing power of forward guidance. Especially in times of severe economic stress, fully truthful projections may be harmful rather than beneficial.

References

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

Ahrens, S. and M. Hartmann (2015). Cross-sectional evidence on state-dependent versus time-dependent price setting. Economics Bulletin 35(4), 2701–2709.

Amato, J. D. and H. S. Shin (2006). Imperfect common knowledge and the information value of prices. Economic Theory 27(1), 213–241.

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

Arifovic, J. and L. Petersen (2015). Escaping expectations-driven liquidity traps:experimental evidence. Unpublished working paper, Simon Fraser University.

Assenza, T., P. Heemeijer, C. H. Hommes, and D. Massaro (2013). Individual expectations and aggregate macro behavior.

Benhabib, J., S. Schmitt-Grohé, and M. Uribe (2001). Monetary policy and multiple equilibria. American Economic Review 91(1), 167–186.

Blinder, A. S. (2000). Central-bank credibility: Why do we care? how do we build it? American Economic Review 90(5), 1421–1431.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. D. Haan, and D.-J. Jansen (2008). Central bank communication and monetary policy: A survey of theory and evidence. Journal of Economic Literature 46(4), 910–45.

Bomfim, A. and G. Rudebusch (2000). Opportunistic and deliberate disination under imperfect credibility. Journal of Money, Credit and Banking 32(4), 707–721.

Bordo, M. and P. Siklos (2014). Central Bank Credibility, Reputation and Inflation Targeting in Historical Perspective. NBER Working Papers 20693, National Bureau of Economic Research, Inc.

Branch, W. A. and B. McGough (2009). A new keynesian model with heterogeneous expectations. Journal of Economic Dynamics and Control 33(5), 1036–1051.

Brock, W. A. and C. H. Hommes (1997). A Rational Route to Randomness. Econometrica 65(5), 1059–1096.

Campbell, J. R., C. L. Evans, J. D. Fisher, A. Justiniano, C. W. Calomiris, and M. Woodford (2012). Macroeconomic effects of federal reserve forward guidance [with comments and discussion]. Brookings Papers on Economic Activity, 1–80.

Capistrán, C. and A. Timmermann (2009). Disagreement and Biases in Inflation Expectations. Journal of Money, Credit and Banking 41(2-3), 365–396.

Cecchetti, S. and S. Krause (2002). Central bank structure, policy effciency, and macroeconomic performance: Exploring empirical relationships. Federal Reserve Bank of St. Louis Review 84, 99–117.

Charemza, W. and D. Ladley (2016). Central banks' forecasts and their bias: Evidence, effects and explanation. International Journal of Forecasting 32(3), 804–817.

Chen, D. L., M. Schonger, and C. Wickens (2016). otree- an open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9, 88–97.

Chortareas, G., D. Stasavage, and G. Sterne (2002). Does it pay to be transparent? international evidence form central bank forecasts. Review 84(4), 99–118.

Clarida, R., J. Galí, and M. Gertler (2000). Monetary policy rules and macroeconomic stability: evidence and some theory. Quarterly Journal of Economics 115, 147–180.

Cornand, C. and F. Heinemann (2014). Experiments on monetary policy and central banking. In J. Duffy (Ed.), Experiments in Macroeconomics (Research in Experimental Economics, Volume 17), pp. 167–227. Emerald Group Publishing Limited.

Cornand, C. and C. K. M'Baye (2016a). Band or point inflation targeting? an experimental approach. WP 1616, GATE Lyon Saint-Étienne.

Cornand, C. and C. K. M'Baye (2016b). Does inflation targeting matter? an experimental investigation. Macroeconomic Dynamics (forthcoming).

Dale, S., A. Orphanides, and P. Österholm (2011). Imperfect central bank communication: Information versus distraction. International Journal of Central Banking 7(2), 3–39.

Dovern, J., U. Fritsche, and J. Slacalek (2012). Disagreement Among Forecasters in G7 Countries. The Review of Economics and Statistics 94(4), 1081–1096.

Dovern, J. and M. Hartmann (2016). Forecast performance, disagreement, and heterogeneous signal-to-noise ratios. Empirical Ecomomics (forthcoming).

Duffy, J. and F. Heinemann (2014). Central bank reputation, cheap talk and transparency as substitutes for commitment: Experimental evidence. manuscript, June.

Ehrmann, M., S. Eijffinger, and M. Fratzscher (2012). The role of central bank transparency for guiding private sector forecasts. Scandinavian Journal of Economics 114(3), 1018–1052.

Eusepi, S. and B. Preston (2010). Central Bank Communication and Expectations Stabilization. American Economic Journal: Macroeconomics 2(3), 235–71.

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

Ferrero, G. and A. Secchi (2010). Central bank's macroeconomic projections and learning. National Bank of Poland Working Paper (72).

Filardo, A. and B. Hofmann (2014). Forward guidance at the zero lower bound. BIS Quarterly Review.

Frederick, S. (2005). Cognitive reflection and decision making. Journal of Economic Perspectives 19(4), 25–42.

Fujiwara, I. (2005). Is the central bank's publication of economic forecasts influential? Economics Letters 89(3), 255–261.

Geraats, P. M. (2002). Central bank transparency. Economic Journal 112(483), 532–565.

Geraats, P. M. (2005). Transparency and reputation: The publication of central bank forecasts. The B.E. Journal of Macroeconomics 5(1), 1–28.

Gersbach, H. (2003). On the negative social value of central banks' knowledge transparency. Economics of Governance 4(2), 91–102.

Gomez-Barrero, S. and J. A. Parra-Polania (2014). Central bank strategic forecasting. Contemporary Economic Policy 32(4), 802–810.

Goy, G., C. Hommes, and K. Mavromatis (2016). Forward guidance and the role of central bank credibility under heterogeneous beliefs. Technical report.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. Journal of the Economic Science Association 1(1), 114–125.

Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. Journal of Economic dynamics and control 35(1), 1–24.

Hommes, C., D. Massaro, and I. Salle (2015). Monetary and fiscal policy design at the zero lower bound - evidence from the lab. Working Paper.

Hommes, C. H. and J. Lustenhouwer (2015). Inflation targeting and the zero lower bound under endogenous credibility. Technical report.

Hubert, P. (2014). Fomc forecasts as a focal point for private expectations. Journal of Money, Credit and Banking 46(7), 1381–1420.

Hubert, P. (2015a). Do central bank forecasts influence private agents? forecasting performance versus signals. Journal of Money, Credit and Banking 47(4), 771–789.

Hubert, P. (2015b). The influence and policy signalling role of fomc forecasts. Oxford Bulletin of Economics and Statistics 77(5), 655–680.

Hubert, P. (2015c). Revisiting the Greenbook's relative forecasting performance. Revue de l'OFCE (1), 151–179.

Jensen, C. (2016). Optimal forward guidance through economic projections in monetary policy. Technical report, University of South Carolina.

Jensen, H. (2002). Optimal degrees of transparency in monetary policymaking. Scandinavian Journal of Economics 104(3), 399–422.

Kryvtsov, O. and L. Petersen (2015). Expectations and monetary policy: Experimental evidence. Unpublished working paper, Simon Fraser University.

Kurz, M., G. Piccillo, and H. Wu (2013). Modeling diverse expectations in an aggregated New Keynesian Model. Journal of Economic Dynamics and Control 37(8), 1403–1433.

Mankiw, N. G., R. Reis, and J. Wolfers (2004, September). Disagreement about Inflation Expectations. In NBER Macroeconomics Annual 2003, Volume 18, NBER Chapters, pp. 209–270. National Bureau of Economic Research, Inc.

Marimon, R. and S. Sunder (1993). Indeterminacy of equilibria in a hyperinflationary world: experimental evidence. Econometrica: Journal of the Econometric Society, 1073–1107.

Massaro, D. (2012). Bounded rationality and heterogeneous expectations in macroeconomics. Technical report, Ph.D. dissertation at the University of Amsterdam.

Mishkin, F. S. (2004, June). Can Central Bank Transparency Go Too Far? In C. Kent and S. Guttmann (Eds.), The Future of Inflation Targeting, RBA Annual Conference Volume. Reserve Bank of Australia.

Mokhtarzadeh, F. and L. Petersen (2016). Coordinating expectations through central bank projections. Technical report.

Morris, S. and H. S. Shin (2002). Social value of public information. American Economic Review 92(5), 1521–1534.

Newey, W. K. and K. D. West (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55(3), 703–08.

Pfajfar, D. and B. Zakelj (2014). Experimental evidence on inflation expectation formation. Journal of Economic Dynamics and Control 44(C), 147–168.

Romer, D. H. and C. D. Romer (2000). Federal reserve information and the behavior of interest rates. American Economic Review 90(3), 429–457.

Romer, D. H. and C. D. Romer (2008). The fomc versus the staff: Where can monetary policymakers add value? American Economic Review 98(2), 230–235.

Surowiecki, J. (2005). The wisdom of crowds. Anchor.

Svensson, L. E. (2015). Day One Keynote Address: Forward Guidance. International Journal of Central Banking 11(4), 19–64.

Tarkka, J. and D. Mayes (1999). The value of publishing official central bank forecasts. Research Discussion Papers 22/1999, Bank of Finland.

Taylor, J. B. (1993). Discretion versus policy rules in practice. In Carnegie-Rochester conference series on public policy, Volume 39, pp. 195–214. Elsevier.

Walsh, C. E. (2007). Optimal economic transparency. International Journal of Central Banking 3(1), 5–36.

Woodford, M. (2001). Inflation stabilization and welfare. NBER Working Papers 8071, National Bureau of Economic Research, Inc.

A Estimation procedure for equation (18)

First Formula (18) is estimated with OLS. Then the joint significance of all the coefficients that where found to be individually insignificant in the above regression is tested. If these coefficients are jointly insignificant, all of them are removed. If they are jointly significant, exactly 1 coefficient is removed. The coefficient that is removed is then the individually insignificant coefficient that ranks first in the following order of removal list: $\alpha_2, \gamma_1, \beta_2, \delta, \beta_1, \alpha_1, c$.

After one or more coefficients are removed, Equation (18) is reestimated without this (these) coefficient(s). Then the joint significance of the coefficients found to be individually insignificant in the new regression is tested and coefficient(s) are removed according to the same procedure as above. This process is repeated until either a regression is performed where all remaining coefficients are individually significant, or until all coefficients are removed.

B Proof of "strategic-ness" index

Below we proof the following three claims about the "strategic-ness" index, SP_t, of equation* (15)*:

(a) $0 < SP_t < 1$ implies that the published forecast lies in the band between the data driven forecast and the "required for target" information:
$E_t^{rft}\pi_{t+1} < E_t^{pub}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$ or $E_t^{ddf}\pi_{t+1} < E_t^{pub}\pi_{t+1} < E_t^{rft}\pi_{t+1}$

(b) $SP_t > 1$ implies that the published forecast lies outside the band of the data driven forecast and the "required for target" information, on the side of the "required for target":
$E_t^{pub}\pi_{t+1} < E_t^{rft}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$ or $E_t^{ddf}\pi_{t+1} < E_t^{rft}\pi_{t+1} < E_t^{pub}\pi_{t+1}$

(c) $SP_t < 0$ implies that the published forecast lies outside the band of the data driven forecast and the "required for target" information, on the side of the data driven forecast:
$E_t^{pub}\pi_{t+1} < E_t^{ddf}\pi_{t+1} < E_t^{rft}\pi_{t+1}$ or $E_t^{rft}\pi_{t+1} < E_t^{ddf}\pi_{t+1} < E_t^{pub}\pi_{t+1}$

Proof: (1) When $0 < SP_t$ either both numerator and denominator in Equation (15) are negative or both are positive:

(i) Consider that both numerator and denominator are negative. It must be that $E_t^{pub}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$ and $E_t^{rft}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$. Since the denominator is negative $SP_t < 1$ further implies that $E_t^{pub}\pi_{t+1} - E_t^{ddf}\pi_{t+1} > E_t^{rft}\pi_{t+1} - E_t^{ddf}\pi_{t+1}$ so that $E_t^{rft}\pi_{t+1} < E_t^{pub}\pi_{t+1}$ and hence $E_t^{rft}\pi_{t+1} < E_t^{pub}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$. $SP_t > 1$, on the other hand, implies that $E_t^{rft}\pi_{t+1} > E_t^{pub}\pi_{t+1}$ and hence $E_t^{pub}\pi_{t+1} < E_t^{rft}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$.

(ii) Consider that both numerator and denominator are positive. $0 < SP_t$ implies $E_t^{pub}\pi_{t+1} > E_t^{ddf}\pi_{t+1}$ and $E_t^{rft}\pi_{t+1} > E_t^{ddf}\pi_{t+1}$. $SP_t < 1$ then implies $E_t^{pub}\pi_{t+1} - E_t^{ddf}\pi_{t+1} < E_t^{rft}\pi_{t+1} - E_t^{ddf}\pi_{t+1}$ so that $E_t^{ddf}\pi_{t+1} < E_t^{pub}\pi_{t+1} < E_t^{rft}\pi_{t+1}$, while $SP_t > 1$ implies $E_t^{ddf}\pi_{t+1} < E_t^{rft}\pi_{t+1} < E_t^{pub}\pi_{t+1}$. This completes the proof of (a) and (b).

(2) When $SP_t < 0$, either the numerator of Equation (15) is negative while the denominator is positive, or the numerator is positive while the denominator is negative: In the first case, it must hold that $E_t^{pub}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$ while $E_t^{rft}\pi_{t+1} > E_t^{ddf}\pi_{t+1}$. In the second case, it must hold that $E_t^{pub}\pi_{t+1} > E_t^{ddf}\pi_{t+1}$ while $E_t^{rft}\pi_{t+1} < E_t^{ddf}\pi_{t+1}$. This proves (c).

C Additional tables and figures

Statistic/Economy	E1	E2	E3	E4	E5	E6	Avg
				Periods 1-8 (Stage I)			
				Inflation			
Mean	3.0	2.6	2.8	2.6	2.0	1.9	2.5
Median	2.6	3.2	2.4	2.5	2.1	1.6	2.4
Variance	0.4	1.6	1.0	0.4	2.5	2.6	1.4
				Output gap			
Mean	-1.2	-0.9	-0.7	-0.9	-0.8	-0.7	-0.9
Median	-1.2	-0.9	-0.7	-0.8	-0.9	-0.8	-0.9
Variance	0.1	0.2	0.3	0.3	0.2	0.4	0.2
				Interest rate			
Mean	7.4	7.0	7.3	7.1	6.2	6.2	6.9
Median	6.9	7.7	6.8	6.8	6.3	5.6	6.7
Variance	0.8	2.5	1.6	0.8	3.7	4.2	2.3
				Periods 9-28 (Stage II)			
				Inflation			
Mean	1.2	0.5	1.4	0.8	-0.3	1.1	0.8
Median	1.1	0.5	1.4	0.5	-1.0	1.0	0.6
Variance	0.4	0.3	0.2	0.5	3.2	1.0	1.0
				Output gap			
Mean	-0.7	-0.3	-0.5	-0.3	-0.2	-0.4	-0.4
Median	-0.7	-0.3	-0.4	-0.3	0.0	-0.4	-0.4
Variance	0.1	0.2	0.2	0.2	0.3	0.2	0.2
				Interest rate			
Mean	5.3	4.6	5.6	4.9	3.5	5.2	4.9
Median	5.2	4.3	5.7	4.6	2.7	5.1	4.6
Variance	0.7	0.5	0.4	0.7	4.7	1.3	1.4
				Periods 29-37 (Stage III)			
				Inflation			
Mean	-1.6	-1.0	-0.2	-65.9	-73.7	-128.4	-45.1
Median	-1.9	-1.6	-0.4	-14.6	-11.5	-21.8	-8.6
Variance	1.0	1.9	1.0	$1.1{*}10^4$	$1.9{*}10^4$	$6.2{*}10^4$	$1.5{*}10^4$
				Output gap			
Mean	-1.5	-1.2	-1.9	-92.5	-90.6	-156.6	-57.4
Median	-2.1	-1.1	-1.5	-14.7	-12.4	-26.3	-9.7
Variance	3.1	2.1	4.2	$2.5{*}10^4$	$2.8{*}10^4$	$8.7{*}10^4$	$2.3{*}10^4$
				Interest rate			
Mean	1.5	2.3	3.1	0.3	0.7	0.2	1.3
Median	1.2	1.8	3.2	0.0	0.0	0.0	1.0
Variance	2.5	4.7	3.8	0.5	2.5	0.4	2.4

Table 9: Descriptive Statistics of Treatment 1 (Control). The table summarizes mean, median, and variance in each of the three stages for each of the six economies of Treatment 1 as well as their corresponding averages over all six economies of Treatment 1.

Statistic/Economy	E7	E8	E9	E10	E11	E12	Avg
	Periods 1-8 (Stage I)						
	Inflation						
Mean	3.1	2.5	2.9	3.4	2.2	1.6	2.6
Median	3.2	3.0	2.7	3.4	2.3	1.5	2.7
Variance	1.3	1.8	0.5	0.1	1.2	0.9	1.0
	Output gap						
Mean	-1.3	-0.9	-0.7	-1.2	-0.9	-0.5	-0.9
Median	-1.4	-0.9	-0.8	-1.1	-1.0	-0.6	-1.0
Variance	0.1	0.2	0.2	0.5	0.2	0.2	0.2
	Interest rate						
Mean	7.5	6.9	7.5	8.0	6.5	5.9	7.0
Median	7.7	7.4	7.1	7.9	6.7	5.7	7.1
Variance	2.2	2.8	0.8	0.2	1.9	1.7	1.6
	Periods 9-28 (Stage II)						
	Inflation						
Mean	0.5	0.2	0.3	0.9	0.1	0.5	0.4
Median	0.4	0.3	0.4	0.8	0.2	0.5	0.4
Variance	0.3	0.2	0.3	0.8	0.5	0.5	0.4
	Output gap						
Mean	-0.4	-0.2	-0.2	-0.3	-0.4	-0.2	-0.3
Median	-0.5	-0.2	0.0	-0.3	-0.3	-0.2	-0.3
Variance	0.1	0.2	0.2	0.2	0.1	0.1	0.2
	Interest rate						
Mean	4.5	4.2	4.4	5.0	4.0	4.6	4.4
Median	4.4	4.3	4.5	5.0	4.1	4.6	4.5
Variance	0.4	0.3	0.6	1.1	0.9	0.8	0.7
	Periods 29-37 (Stage III)						
	Inflation						
Mean	-0.9	-1.2	-1.7	-1.3	-23.1	-0.9	-4.9
Median	-1.0	-0.6	-2.0	-1.6	-8.9	-0.7	-2.5
Variance	1.0	2.7	6.7	3.5	1155.5	2.5	195.3
	Output gap						
Mean	-2.0	-1.5	-2.4	-1.8	-29.6	-1.6	-6.5
Median	-2.3	-1.7	-2.3	-2.2	-10.4	-1.3	-3.4
Variance	4.7	3.0	4.6	4.7	1898.8	2.9	319.8
	Interest rate						
Mean	2.3	2.3	2.1	2.2	0.2	2.5	1.9
Median	2.0	3.0	0.4	1.3	0.0	2.3	1.5
Variance	3.7	4.4	7.6	4.9	0.4	4.8	4.3

Table 10: Descriptive Statistics of Treatment 2. The table summarizes mean, median, and variance in each of the three stages for each of the six economies of Treatment 2 as well as their corresponding averages over all six economies of Treatment 2.

Statistic/Economy	E13	E14	E15	E16	E17	E18	Avg
	Periods 1-8 (Stage I)						
	Inflation						
Mean	2.4	2.2	2.4	2.8	2.6	3.0	2.6
Median	1.8	2.6	2.5	2.9	2.7	2.8	2.5
Variance	0.7	1.7	0.0	1.5	0.4	1.4	1.0
	Output gap						
Mean	-1.0	-0.8	-0.5	-0.9	-1.0	-1.0	-0.9
Median	-1.0	-0.8	-0.6	-0.8	-1.1	-1.1	-0.9
Variance	0.1	0.1	0.2	0.5	0.1	0.4	0.2
	Interest rate						
Mean	6.7	6.5	6.8	7.3	7.0	7.5	7.0
Median	6.0	6.9	6.9	7.3	7.0	7.1	6.9
Variance	1.2	2.8	0.1	2.6	0.7	2.5	1.7
	Periods 9-28 (Stage II)						
	Inflation						
Mean	0.4	0.0	0.2	0.3	0.0	0.8	0.3
Median	0.4	0.0	0.2	0.2	-0.1	0.6	0.2
Variance	0.2	0.2	0.2	0.3	0.4	1.0	0.4
	Output gap						
Mean	-0.4	-0.1	-0.1	-0.1	-0.4	-0.4	-0.2
Median	-0.4	-0.2	0.0	-0.1	-0.2	-0.4	-0.2
Variance	0.1	0.1	0.2	0.1	0.2	0.2	0.2
	Interest rate						
Mean	4.4	4.0	4.2	4.3	3.9	4.9	4.3
Median	4.4	4.0	4.3	4.3	3.7	4.6	4.2
Variance	0.3	0.3	0.3	0.4	0.6	1.4	0.6
	Periods 29-37 (Stage III)						
	Inflation						
Mean	-0.8	-1.8	-0.9	-1.0	-1.7	-2.0	-1.4
Median	-0.8	-2.3	-0.9	-1.6	-1.7	-1.8	-1.5
Variance	0.9	1.0	1.1	3.1	1.5	3.5	1.9
	Output gap						
Mean	-1.9	-1.2	-1.7	-1.5	-1.9	-1.8	-1.7
Median	-2.1	-1.4	-1.7	-1.0	-2.0	-2.6	-1.8
Variance	4.4	3.2	4.4	5.4	3.2	2.7	3.9
	Interest rate						
Mean	2.3	1.4	2.3	2.4	1.5	1.5	1.9
Median	2.1	1.3	2.2	2.3	0.7	0.0	1.4
Variance	3.6	1.9	3.4	6.2	2.8	4.4	3.7

Table 11: Descriptive Statistics of Treatment 3. The table summarizes mean, median, and variance in each of the three stages for each of the six economies of Treatment 3 as well as their corresponding averages over all six economies of Treatment 3.

Statistic/Economy	E19	E20	E21	E22	E23	E24	Avg
	Periods 1-8 (Stage I)						
	Inflation						
Mean	1.6	2.3	2.2	3.0	2.6	2.9	2.5
Median	1.2	2.7	2.3	3.0	2.4	2.6	2.4
Variance	0.7	1.7	0.6	0.5	1.1	0.6	0.9
	Output gap						
Mean	-0.7	-0.8	-0.5	-1.0	-1.0	-1.0	-0.8
Median	-0.7	-0.9	-0.5	-0.9	-1.1	-1.0	-0.8
Variance	0.0	0.1	0.2	0.4	0.1	0.3	0.2
	Interest rate						
Mean	5.8	6.7	6.7	7.5	7.0	7.4	6.8
Median	5.3	7.1	6.7	7.5	6.7	7.0	6.7
Variance	1.1	2.8	1.0	1.0	1.8	1.2	1.5
	Periods 9-28 (Stage II)						
	Inflation						
Mean	0.2	-0.1	1.1	0.9	0.6	0.9	0.6
Median	0.2	-0.1	1.0	1.1	0.6	1.1	0.7
Variance	0.6	0.2	0.2	1.0	0.6	2.5	0.8
	Output gap						
Mean	-0.2	-0.1	-0.4	-0.3	-0.6	-0.4	-0.3
Median	-0.3	-0.1	-0.3	-0.2	-0.5	-0.4	-0.3
Variance	0.1	0.2	0.3	0.0	0.2	0.2	0.2
	Interest rate						
Mean	4.1	3.9	5.2	5.1	4.5	5.0	4.6
Median	4.2	3.8	5.2	5.4	4.5	5.2	4.7
Variance	0.9	0.3	0.2	1.7	1.0	3.4	1.2
	Periods 29-37 (Stage III)						
	Inflation						
Mean	-0.7	-1.5	-13.6	-1.1	-0.6	0.3	-2.9
Median	-1.0	-1.9	-6.0	-1.4	-0.5	0.4	-1.7
Variance	1.1	0.8	409.4	1.0	0.9	2.9	69.4
	Output gap						
Mean	-1.9	-1.0	-15.2	-1.4	-2.2	-2.2	-4.0
Median	-2.0	-0.9	-5.2	-0.8	-2.4	-2.1	-2.2
Variance	4.4	2.7	482.0	4.4	2.4	4.2	83.3
	Interest rate						
Mean	2.5	1.6	0.5	2.1	2.5	3.7	2.1
Median	2.2	1.1	0.0	2.2	2.5	3.6	1.9
Variance	4.3	2.1	1.1	3.3	3.2	7.7	3.6

Table 12: Descriptive Statistics of Treatment 4. The table summarizes mean, median, and variance in each of the three stages for each of the six economies of Treatment 4 as well as their corresponding averages over all six economies of Treatment 4.

	Treatment 1 (Control) vs. Treatment 2		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.5725)	no*** (0.0001)	no** (0.0370)
Output	yes (0.7030)	no** (0.0369)	no** (0.0174)
Interest	yes (0.5750)	no*** (0.0001)	yes (0.1373)
	Treatment 1 (Control) vs. Treatment 3		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.9795)	no***(0.0000)	no** (0.0110)
Output	yes (0.9124)	no***(0.0052)	no*** (0.0001)
Interest	yes (0.9649)	no***(0.0000)	yes (0.0365)
	Treatment 1 (Control) vs. Treatment 4		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.7499)	no** (0.0401)	no***(0.0016)
Output	yes (0.6207)	yes (0.1517)	no***(0.0085)
Interest	yes (0.8922)	no* (0.0509)	no** (0.0109)
	Treatment 2 vs. Treatment 3		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.5333)	no** (0.0283)	yes (0.4894)
Output	yes (0.6363)	yes (0.3941)	no* (0.0946)
Interest	yes (0.6002)	no* (0.0540)	yes (0.6951)
	Treatment 2 vs. Treatment 4		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.3444)	no** (0.0354)	yes (0.2100)
Output	yes (0.3671)	yes (0.5047)	yes (0.8658)
Interest	yes (0.3851)	no** (0.0486)	yes (0.4092)
	Treatment 3 vs. Treatment 4		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.6469)	no*** (0.0002)	yes (0.2740)
Output	yes (0.6336)	yes (0.1266)	yes (0.1222)
Interest	yes (0.6735)	no*** (0.0005)	yes (0.6487)

Table 13: Mann-Whitney-Wilcoxon-test for equality in medians. The table shows the outcome and the p-values (in parentheses) of the Mann-Whitney-Wilcoxon-test for pairwise across-treatment comparisons of inflation, the output gap, and the interest rate for each of the three stages, respectively. The Null hypothesis is that samples have identical median. The outcome “yes” implies that we cannot reject that both samples have identical median, while the outcome “no” implies that we reject equality of the median

	Treatment 1 (Control) vs. Treatment 2		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.8032)	no*** (0.0010)	yes (0.3362)
Output	yes (1.0000)	no* (0.0656)	yes (0.1824)
Interest	yes (0.6628)	no*** (0.0055)	yes (0.1988)
	Treatment 1 (Control) vs. Treatment 3		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.4307)	no*** (0.0008)	no*** (0.0047)
Output	yes (0.4414)	yes (0.1389)	no* (0.0533)
Interest	yes (0.4955)	no*** (0.0007)	yes (0.4596)
	Treatment 1 (Control) vs. Treatment 4		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.4223)	yes (0.1648)	no**(0.0332)
Output	yes (0.3350)	yes (0.1200)	yes (0.1067)
Interest	yes (0.4839)	yes (0.2582)	yes (0.3936)
	Treatment 2 vs. Treatment 3		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.3333)	yes (0.1716)	no*** (0.0069)
Output	yes (0.5307)	yes (0.7945)	yes (0.3048)
Interest	yes (0.4138)	yes (0.1499)	yes (0.1298)
	Treatment 2 vs. Treatment 4		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.4679)	no*** (0.0029)	no**(0.0472)
Output	yes (0.4501)	yes (0.7594)	yes (0.6648)
Interest	yes (0.6311)	no*** (0.0074)	no* (0.0833)
	Treatment 3 vs. Treatment 4		
Variable	Periods 1-8	Periods 9-28	Periods 29-37
Inflation	yes (0.7443)	no*** (0.0002)	yes (0.6896)
Output	yes (0.6493)	yes (0.5161)	yes (0.5368)
Interest	yes (0.8004)	no*** (0.0001)	yes (0.8646)

Table 14: Siegel-Turkey-test for equality in variance. The table shows the outcome and the p-values (in parentheses) of the Siegel-Turkey-test for pairwise across-treatment-comparisons of inflation, the output gap, and the interest rate for each of the three stages, respectively. The Null hypothesis is that samples have identical variance. The outcome “yes” implies that we cannot reject that both samples have identical variance, while the outcome “no” implies that we reject equality of the variance.

	Treatment 1						
Economy	E1	E2	E3	E4	E5	E6	Avg
R_i^{π}	0.2108	0.0747	0.2521	0.1610	0.5296	0.344	0.2620
R_i^{y}	0.3865	0.2458	0.7596	0.2166	0.3924	0.442	0.4072
	Treatment 2						
Economy	E7	E8	E9	E10	E11	E12	Avg
R_i^{π}	0.0466	0.0249	0.0471	0.1247	0.0863	0.238	0.0945
R_i^{y}	0.1595	0.1962	0.3177	0.1580	0.2807	0.346	0.2431
	Treatment 3						
Economy	E13	E14	E15	E16	E17	E18	Avg
R_i^{π}	0.0523	0.0270	0.0349	0.0368	0.0471	0.147	0.0576
R_i^{y}	0.2429	0.1993	0.5759	0.0926	0.2746	0.234	0.2699
	Treatment 4						
Economy	E19	E20	E21	E22	E23	E24	Avg
R_i^{π}	0.1871	0.0261	0.2390	0.1895	0.1092	0.348	0.1831
R_i^{y}	0.2956	0.1852	1.2620	0.0918	0.4225	0.275	0.4220

Table 15: Relative stability of inflation and the output gap (Stage II/Stage I). The table shows the relative stability measures (stability measure in Stage II relative to the respective stability measure of Stage I) given by equations (13) and (14) for all 24 economies, as well as their respective treatment averages. If the measure takes a value below unity, the economy is more stable in Stage II relative to Stage I.

		Treatment 1 (Control)						
Measure	Statistic	E1	E2	E3	E4	E5	E6	Avg
VAR	Median	0.146	0.056	0.399	0.841	0.724	0.345	0.418
	StdDev	0.146	0.178	0.052	0.154	0.303	1.154	0.331
RNG	Median	0.397	0.232	0.667	0.819	1.035	0.640	0.593
	StdDev	0.341	0.374	0.215	0.359	0.439	0.977	0.739
IQR	Median	0.610	0.356	0.662	0.574	0.617	0.739	0.593
	StdDev	0.206	0.842	0.795	0.529	0.738	1.322	0.725
		Treatment 2						
Measure	Statistic	E7	E8	E9	E10	E11	E12	Avg
VAR	Median	0.191	0.024	0.242	0.408	0.128	0.088	0.180
	StdDev	0.205	0.102	0.227	1.313	0.647	0.130	0.437
RNG	Median	0.452	0.155	0.414	0.591	0.388	0.298	0.383
	StdDev	0.433	0.343	0.477	1.217	0.887	0.345	0.617
IQR	Median	0.490	0.186	0.665	0.801	0.356	0.470	0.495
	StdDev	0.502	0.307	1.625	1.152	0.703	0.411	0.783
		Treatment 3						
Measure	Statistic	E13	E14	E15	E16	E17	E18	Avg
VAR	Median	0.271	0.109	0.082	0.023	0.520	0.370	0.229
	StdDev	0.464	0.147	0.336	0.104	0.131	0.876	0.343
RNG	Median	0.553	0.344	0.328	0.158	0.694	0.596	0.446
	StdDev	0.581	0.403	0.445	0.458	0.358	1.079	0.554
IQR	Median	0.442	0.215	0.504	0.108	0.553	0.659	0.413
	StdDev	0.477	0.267	0.787	0.487	0.758	1.023	0.633
		Treatment 4						
Measure	Statistic	E19	E20	E21	E22	E23	E24	Avg
VAR	Median	0.585	0.118	0.256	0.075	1.313	16.764	3.185
	StdDev	1.524	0.055	0.453	0.303	3.680	6.262	2.046
RNG	Median	0.675	0.400	0.523	0.255	1.273	4.145	1.212
	StdDev	1.014	0.189	0.931	0.443	1.767	2.522	1.144
IQR	Median	0.745	0.464	0.510	1.060	1.034	2.204	1.003
	StdDev	1.667	0.468	0.495	1.483	4.583	2.040	1.789

Table 16: Relative dispersion of individual forecasts (stage II/stage I). The table shows the medians of three dispersion measures from periods $t = 9, ..., 28$ relative to the periods $t = 1, ..., 8$ medians for all six economies in each treatment and the respective treatment averages.

	Treatment 1 (Control)						
MSE of...	E1	E2	E3	E4	E5	E6	Avg
pub	—	—	—	—	—	—	—
cbf	0.3526	0.1285	0.6639	0.1735	3.1312	1.0832	0.9221
ddf	0.4433	0.1557	0.5825	0.1823	1.4843	0.8233	0.6119
afc	0.2716	0.1999	0.3963	0.1975	1.7205	0.8746	0.6101
$p_{cfb,ddf}$	0.6001	0.3409	0.2892	0.9338	0.0000	0.6321	0.0887
$p_{cfb,afc}$	0.5167	0.2216	0.1036	0.6460	0.1635	0.4164	0.1544
$p_{ddf,afc}$	0.3107	0.3466	0.0265	0.8480	0.8128	0.9177	0.9926
	Treatment 2						
MSE of...	E7	E8	E9	E10	E11	E12	Avg
pub	0.3348	0.2056	0.5158	0.6391	0.5627	0.4112	0.4449
cbf	0.3348	0.2056	0.5158	0.6391	0.5627	0.4112	0.4449
ddf	0.4677	0.3281	0.8598	0.6749	0.9051	0.4798	0.6192
afc	0.3553	0.1508	0.5850	0.7913	0.6722	0.4271	0.4970
$p_{pub,cfb}$	—	—	—	—	—	—	—
$p_{pub,ddf}$	0.0448	0.0966	0.0022	0.8210	0.1635	0.0030	0.0036
$p_{pub,afc}$	0.0893	0.2892	0.0007	0.3892	0.0089	0.3994	0.0703
$p_{cfb,ddf}$	0.0448	0.0966	0.0022	0.8210	0.1635	0.0030	0.0036
$p_{cfb,afc}$	0.0893	0.2892	0.0007	0.3892	0.0089	0.3994	0.0703
$p_{ddf,afc}$	0.1026	0.0000	0.0232	0.5504	0.2645	0.0483	0.0420
	Treatment 3						
MSE of...	E13	E14	E15	E16	E17	E18	Avg
pub	0.1364	0.1708	0.1054	0.2184	0.2639	0.3708	0.2110
cbf	0.3264	0.2696	0.9771	1.6227	0.4850	1.3538	0.8391
ddf	0.3123	0.2626	0.4494	1.2112	0.5417	0.3651	0.5237
afc	0.2864	0.1358	0.1964	0.2957	0.3566	0.5358	0.3011
$p_{pub,cfb}$	0.0008	0.2929	0.0407	0.0020	0.0682	0.0059	0.0012
$p_{pub,ddf}$	0.0000	0.0000	0.0000	0.0006	0.0029	0.9643	0.0051
$p_{pub,afc}$	0.0020	0.4570	0.0071	0.0057	0.2129	0.2704	0.0025
$p_{cfb,ddf}$	0.8322	0.9262	0.2877	0.3927	0.5167	0.0264	0.0349
$p_{cfb,afc}$	0.5804	0.0357	0.1383	0.0022	0.2015	0.0105	0.0039
$p_{ddf,afc}$	0.1653	0.0015	0.0015	0.0006	0.0038	0.4473	0.0407
	Treatment 4						
MSE of...	E19	E20	E21	E22	E23	E24	Avg
pub	12.4406	6.8002	7.5838	11.1269	7.6379	11.2269	9.4694
cbf	2.3747	0.4782	0.5429	2.9523	1.1893	2.9877	1.7542
ddf	1.4340	0.3904	0.3658	3.5216	0.6601	18.6521	4.1707
afc	1.4157	0.1865	0.2121	2.1699	0.7095	5.7250	1.7365
$p_{pub,cfb}$	0.0000	0.0000	0.0000	0.0003	0.0000	0.0001	0.0000
$p_{pub,ddf}$	0.0000	0.0000	0.0000	0.0015	0.0000	0.5570	0.0288
$p_{pub,afc}$	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
$p_{cfb,ddf}$	0.1697	0.5975	0.0000	0.5121	0.1778	0.2491	0.3290
$p_{cfb,afc}$	0.0604	0.1216	0.0013	0.0697	0.1470	0.1943	0.9707
$p_{ddf,afc}$	0.8681	0.0013	0.2462	0.0116	0.8417	0.2870	0.2569

Table 17: Stage II mean squared forecast errors of public central bank projection (pub), the central bank forecaster (cbf), the data-driven forecast (ddf), and the aggregate inflation forecast (afc). The p-values test (pairwise) the null hypothesis of equal forecast performance from estimation of equation (17).

	Treatment 1 (Control)						
MSE of...	E1	E2	E3	E4	E5	E6	Avg
FC1	0.4497	0.5543	0.6247	0.2833	2.3277	1.5632	—
FC2	0.2111	0.4802	0.5945	0.2250	3.7620	0.9305	—
FC3	0.2152	0.2893	0.6454	0.4287	2.2715	1.0389	—
FC4	0.9108	0.1726	0.9195	0.1238	2.7529	1.2018	—
FC5	0.3300	0.3609	0.3141	0.5094	2.1862	2.0727	—
FC6	0.2075	0.2316	0.5365	0.2036	2.0239	1.2129	—
Avg	0.3874	0.3482	0.6058	0.2956	2.5540	1.3367	0.9213
	Treatment 2						
MSE of...	E7	E8	E9	E10	E11	E12	Avg
FC1	0.3584	0.2240	0.8300	0.5469	0.6651	0.7203	—
FC2	0.6875	0.2456	0.6117	1.5201	0.7949	0.6865	—
FC3	0.4110	0.1897	0.6458	1.1044	1.0740	0.4517	—
FC4	0.5531	0.2808	0.8203	1.1099	0.7667	0.3813	—
FC5	0.6490	0.1696	0.8520	1.7994	0.6238	0.5655	—
FC6	0.3224	0.1382	0.8060	0.9116	0.7493	0.4683	—
Avg	0.4969	0.2080	0.7610	1.1654	0.7790	0.5456	0.6593
	Treatment 3						
MSE of...	E13	E14	E15	E16	E17	E18	Avg
FC1	1.4954	0.2084	0.6284	0.4377	0.4748	0.9681	—
FC2	0.3295	0.1324	0.2564	0.4311	0.9802	0.4643	—
FC3	0.2134	0.1600	0.2290	0.3590	0.3806	1.4494	—
FC4	0.1803	0.1929	0.2166	0.3905	0.6426	0.6912	—
FC5	0.5439	0.5522	0.2989	0.4341	0.2793	0.7486	—
FC6	0.3345	0.3756	1.4721	0.4320	0.4275	0.5949	—
Avg	0.5161	0.2703	0.5169	0.4141	0.5308	0.8194	0.5113
	Treatment 4						
MSE of...	E19	E20	E21	E22	E23	E24	Avg
FC1	2.0562	0.2334	0.5474	2.3648	0.8475	7.2721	—
FC2	1.5361	0.2305	0.5858	3.9694	1.4475	6.5141	—
FC3	3.0017	0.4151	0.3337	2.6618	1.0044	8.7833	—
FC4	1.9393	0.2123	0.3822	1.3086	1.0983	6.4836	—
FC5	2.4415	0.2937	0.2928	2.0386	1.6323	4.3345	—
FC6	1.6661	0.2458	0.2611	4.2126	3.3212	7.2281	—
Avg	2.1068	0.2718	0.4005	2.7593	1.5585	6.7693	2.3110

Table 18: Mean squared forecast errors of individual forecasters. The table shows Stage II mean squared forecast errors, i.e. $\pi_{t+1} - E_t\pi_{t+1}$ for each individual professional forecaster in all 24 economies, as well as their respective treatment averages.

Treatment	Treatment 2		Treatment 3		Treatment 4	
CRT score	High	Low	High	Low	High	Low
# of subj.	27	9	22	14	21	15
constant	37%	33%	59%	50%	38%	66%
	(0.325)	(0.462)	(0.115)	(0.292)	(0.602)	(1.209)
$E_{t-1}^{fc,j}\pi_t$	22%	11%	13%	14%	23%	13%
	(0.185)	(-0.598)	(0.395)	(0.475)	(0.613)	(0.020)
$E_{t-2}^{fc,j}\pi_{t-1}$	7%	22%	18%	14%	4%	13%
	(-0.479)	(0.003)	(-0.247)	(-0.536)	(0.658)	(-0.551)
π_{t-1}	40%	66%	50%	64%	47%	33%
	(0.599)	(0.739)	(0.822)	(0.666)	(0.860)	(0.784)
π_{t-2}	22%	33%	22%	7%	9%	13%
	(-0.358)	(-0.969)	(0.386)	(-0.444)	(-0.811)	(0.108)
y_{t-1}	11%	11%	13%	14%	14%	40%
	(1.242)	(-1.317)	(-0.202)	(0.898)	(-1.869)	(2.292)
$E_t^{pub}\pi_{t+1}$	77%	44%	36%	21%	29%	33%
	(0.816)	(0.839)	(1.732)	(1.418)	(0.120)	(0.336)
avg. R^2	0.78	0.78	0.77	0.62	0.54	0.55
#Sign.Coeff	2.18	2.22	2.13	1.85	1.67	2.13

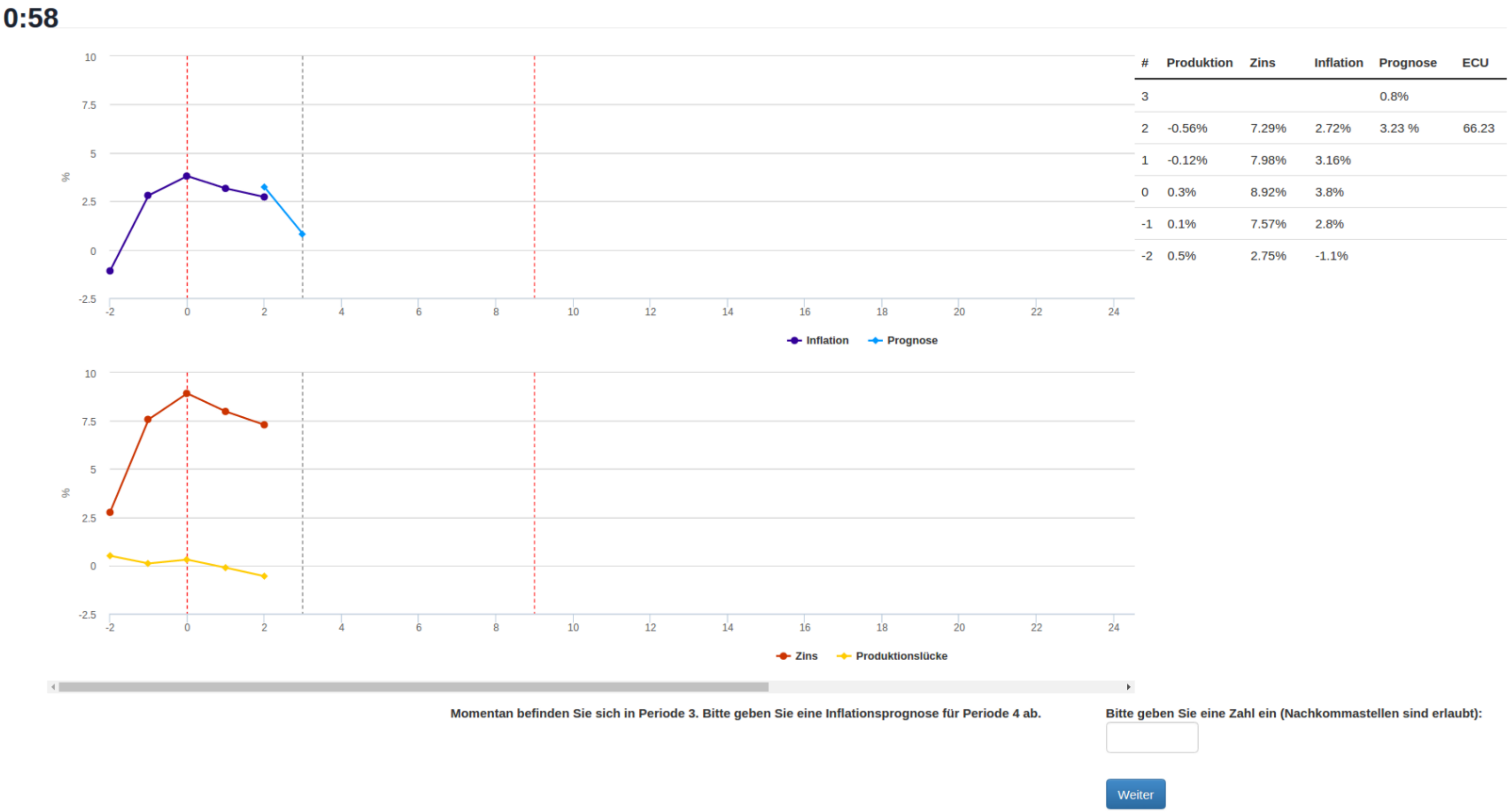
Table 19: Percentages of significant regressors and the median regression coefficients (in parentheses) from estimating equations (18) and (19) for all professional forecasters per treatment, when samples are split according to the subjects' **CRT score**. A CRT score is considered high, if a subject answered at least two out of three questions correctly. Additionally, the table shows the average R^2 and the average number of significant coefficients per forecaster for each treatment.

	[1]	[2]	[3]	[4]	[5]
constant	−0.712***	−0.559***	−1.495***	−1.700***	−1.697***
	[0.063]	[0.066]	[0.162]	[0.175]	[0.175]
$\lvert E^{pub}_{t-2}\pi_{t-1} - \pi_{t-1}\rvert$	−0.383***	−0.331***	−0.131**	−0.133**	−0.144**
	[0.051]	[0.047]	[0.055]	[0.055]	[0.056]
$Utilize_{t-2}$	0.693***	0.666***	0.518***	0.491***	0.361***
	[0.126]	[0.121]	[0.124]	[0.125]	[0.088]
$\lvert E^{fc,j}_{t-2}\pi_{t-1} - \pi_{t-1}\rvert$	−0.054	−0.013	−0.017	−0.014	
	[0.073]	[0.027]	[0.030]	[0.028]	
$\lvert E^{fc,j}_{t-2}\pi_{t-1} - \pi_{t-1}\rvert * Utilize_{t-2}$	−0.317	−0.292	−0.299	−0.279	
	[0.196]	[0.184]	[0.190]	[0.190]	
$\lvert\pi_{t-1} - \pi^T\rvert$		−0.440***	−0.319***	−0.319***	−0.333***
		[0.098]	[0.100]	[0.100]	[0.100]
I^{cred}_{t-1}			1.208***	1.185***	1.179***
			[0.190]	[0.191]	[0.191]
CRT				0.110***	0.113***
				[0.035]	[0.035]

Table 20: Determinants of the utilization of central bank projections in Stage II. This table summarizes the results of a series of probit models from Section 6.1, where the dependent variable $Utilize_t$ is binary taking value 1 if individual professional forecasters utilized the central bank projection and 0 if not. A central bank projection is said to be utilized if an individual professional forecasters forecast is within 5 basis points of the respective central bank projection. The data used for estimation of the series of probit models stems from Stage II of Treatments 2, 3, and 4.

	[1]	[2]	[3]	[4]	[5]	[6]
constant	−0.798***	−0.481***	−0.923***	−0.983***	−1.012***	−0.892***
	[0.070]	[0.082]	[0.141]	[0.152]	[0.130]	[0.136]
$\lvert E^{pub}_{t-2}\pi_{t-1} - \pi_{t-1}\rvert$	−0.314***	−0.258***	−0.273***	−0.277***	−0.247***	−0.281***
	[0.052]	[0.058]	[0.061]	[0.061]	[0.056]	[0.061]
$Utilize_{t-2}$	0.826***	0.616***	0.521***	0.509***	0.498***	0.411***
	[0.160]	[0.167]	[0.169]	[0.170]	[0.170]	[0.098]
$\lvert E^{fc,j}_{t-2}\pi_{t-1} - \pi_{t-1}\rvert$	0.059	0.114**	0.136**	0.139**	0.101*	0.126**
	[0.050]	[0.055]	[0.057]	[0.057]	[0.056]	[0.056]
$\lvert E^{fc,j}_{t-2}\pi_{t-1} - \pi_{t-1}\rvert * Utilize_{t-2}$	−0.171*	−0.057	−0.085	−0.082	−0.063	
	[0.102]	[0.107]	[0.107]	[0.107]	[0.108]	
$\lvert \pi_{t-1} - \pi^T\rvert$		−0.334***	−0.310***	−0.310***	−0.358***	−0.314***
		[0.040]	[0.040]	[0.040]	[0.040]	[0.040]
I^{cred}_{t-1}			0.703***	0.698***		0.694***
			[0.182]	[0.182]		[0.694]
$Cred_{Reported}$					0.093***	
					[0.015]	
CRT				0.034	0.026	
				[0.034]	[0.034]	

Table 21: Determinants of the utilization of central bank projections in Stage III. This table summarizes the results of a series of probit models from Section 6.2, where the dependent variable $Utilize_t$ is binary taking value 1 if individual professional forecasters utilized the central bank projection and 0 if not. A central bank projection is said to be utilized if an individual professional forecasters forecast is within 5 basis points of the respective central bank projection. The data used for estimation of the series of probit models stems from Stage III of Treatments 2, 3, and 4.



#	Produktion	Zins	Inflation	Prognose	ECU
3				0.8%	
2	-0.56%	7.29%	2.72%	3.23 %	66.23
1	-0.12%	7.98%	3.16%		
0	0.3%	8.92%	3.8%		
-1	0.1%	7.57%	2.8%		
-2	0.5%	2.75%	-1.1%		

Figure 3: Computer interface as seen by the subjects. The figure shows the graphical and tabular representation of the complete history of the economy as well as the timer and the input box. The exemplary subject is currently in period 3 and she is asked to provide a forecast for period 4.

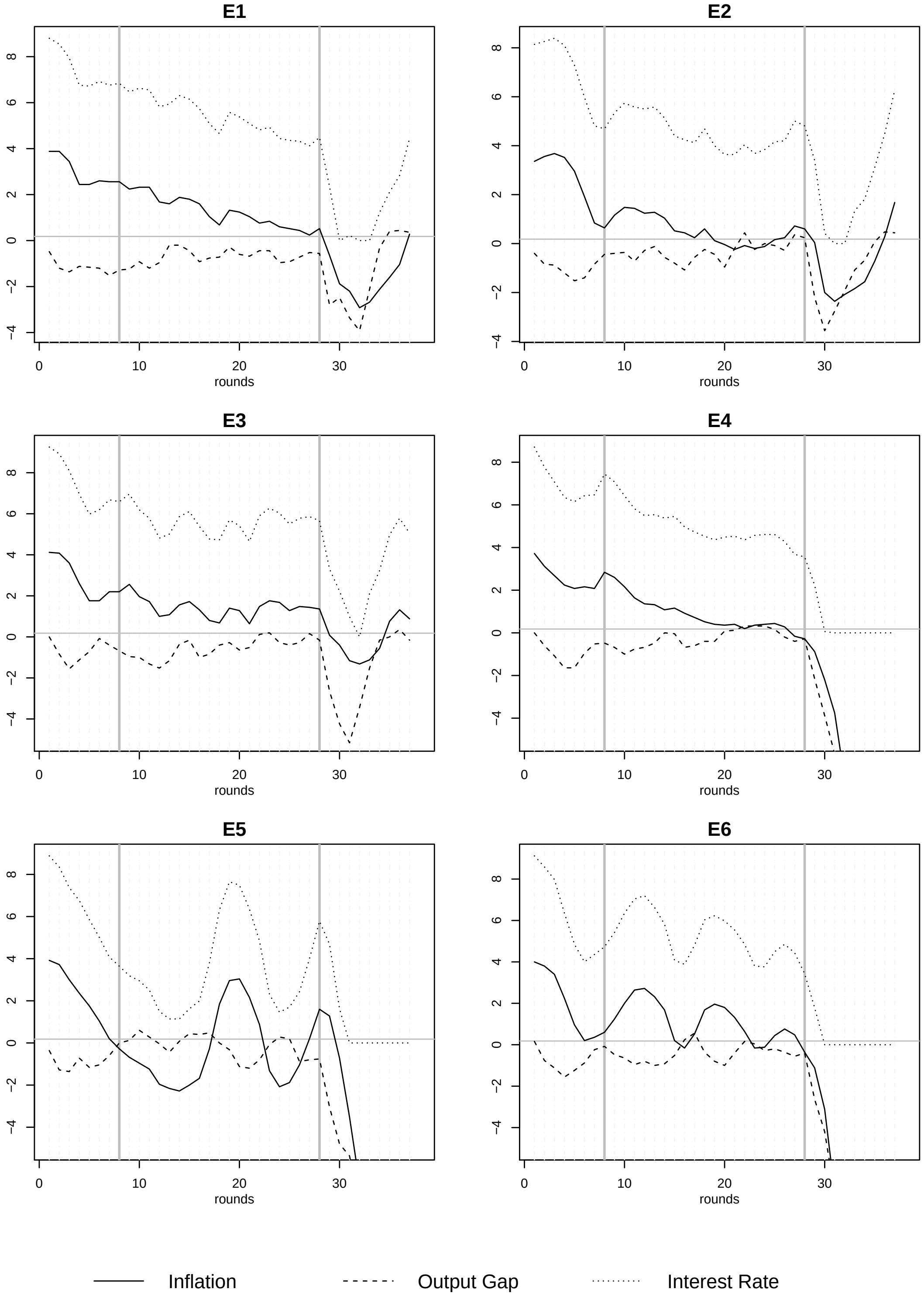


Figure 4: Resulting aggregate time series for inflation (solid line), the output gap (dashed line), and the interest rate (dotted line) for all six experimental economies of Treatment 1 (control treatment).

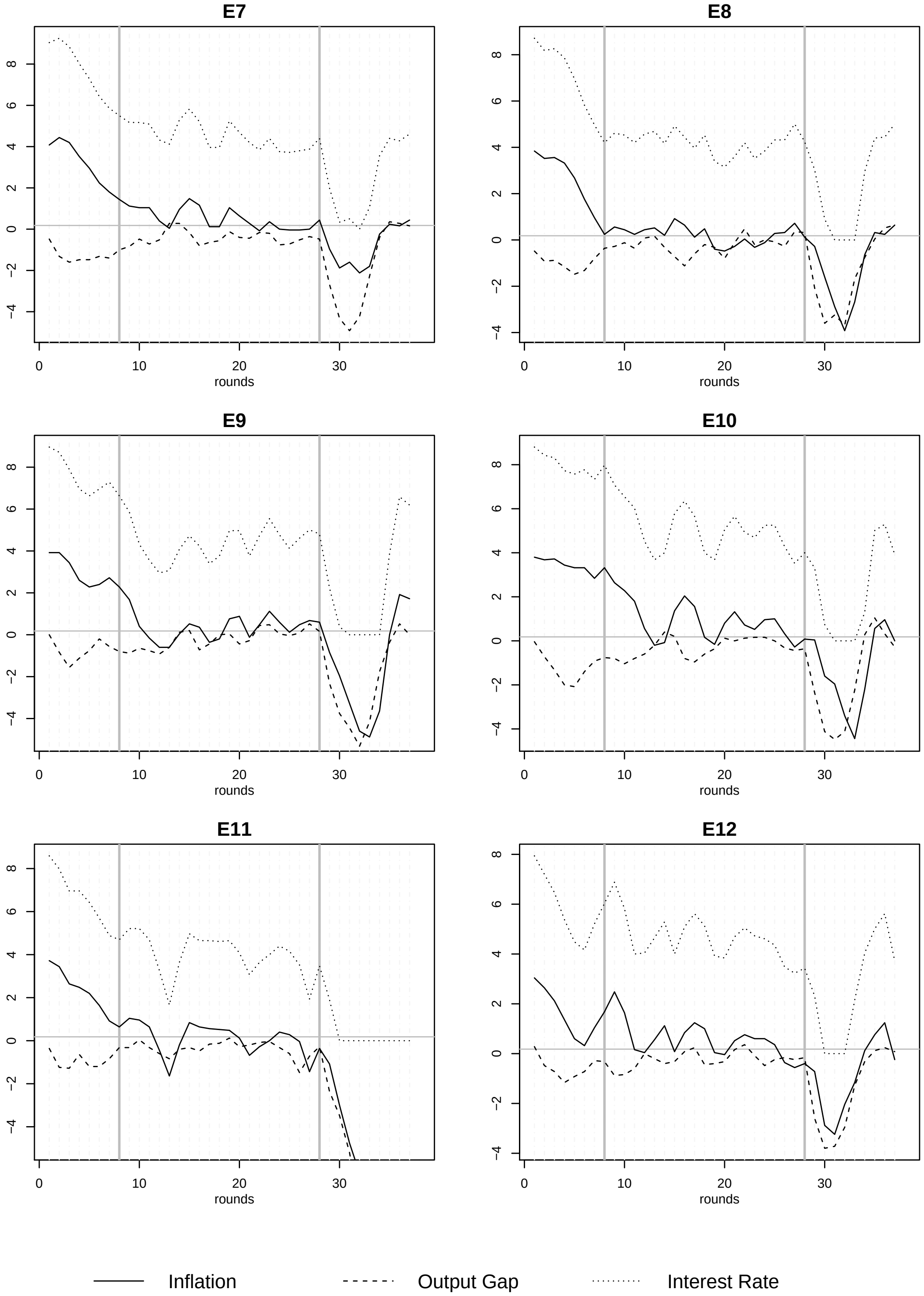


Figure 5: Resulting aggregate time series for inflation (solid line), the output gap (dashed line), and the interest rate (dotted line) for all six experimental economies of Treatment 2.

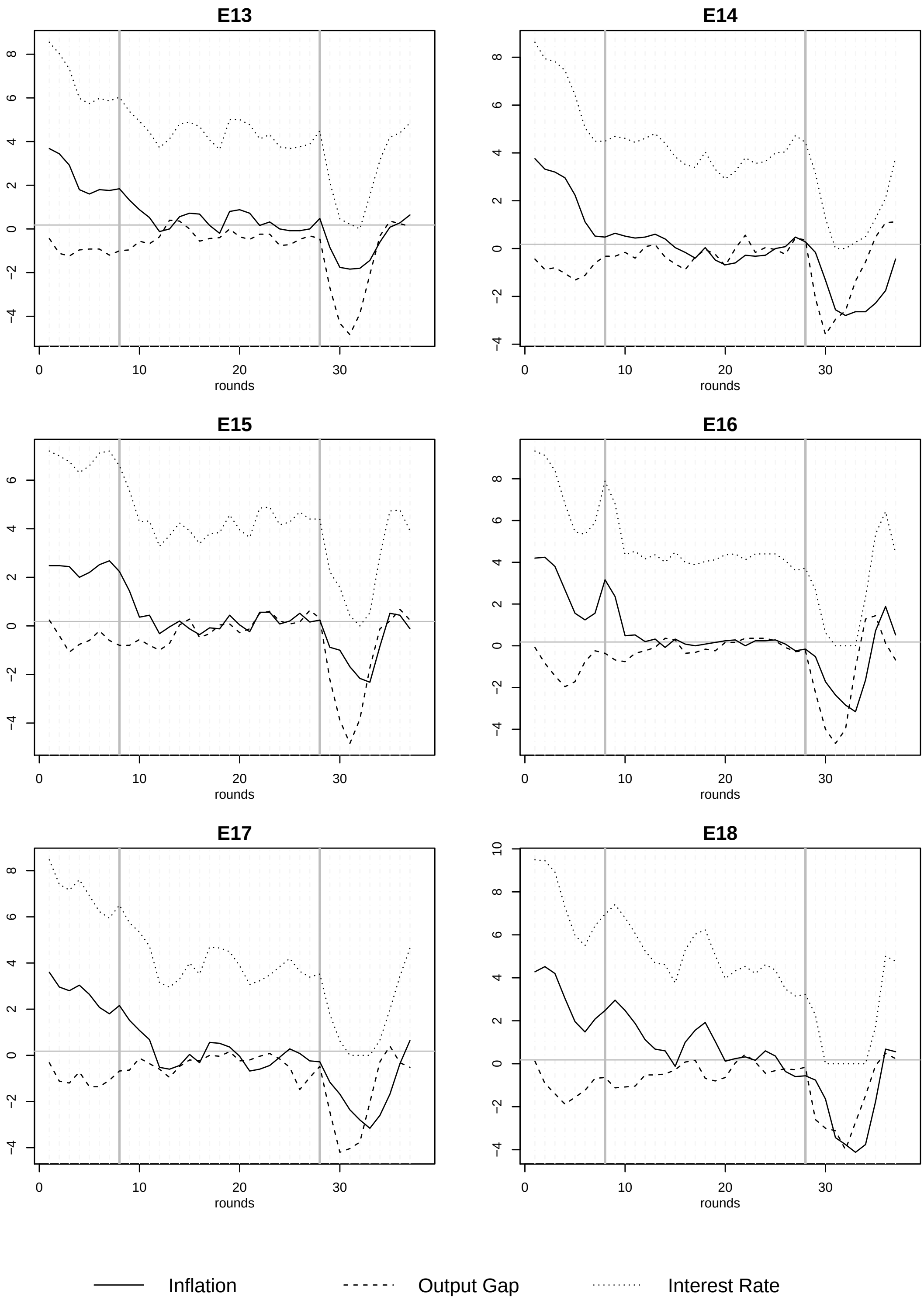


Figure 6: Resulting aggregate time series for inflation (solid line), the output gap (dashed line), and the interest rate (dotted line) for all six experimental economies of Treatment 3.

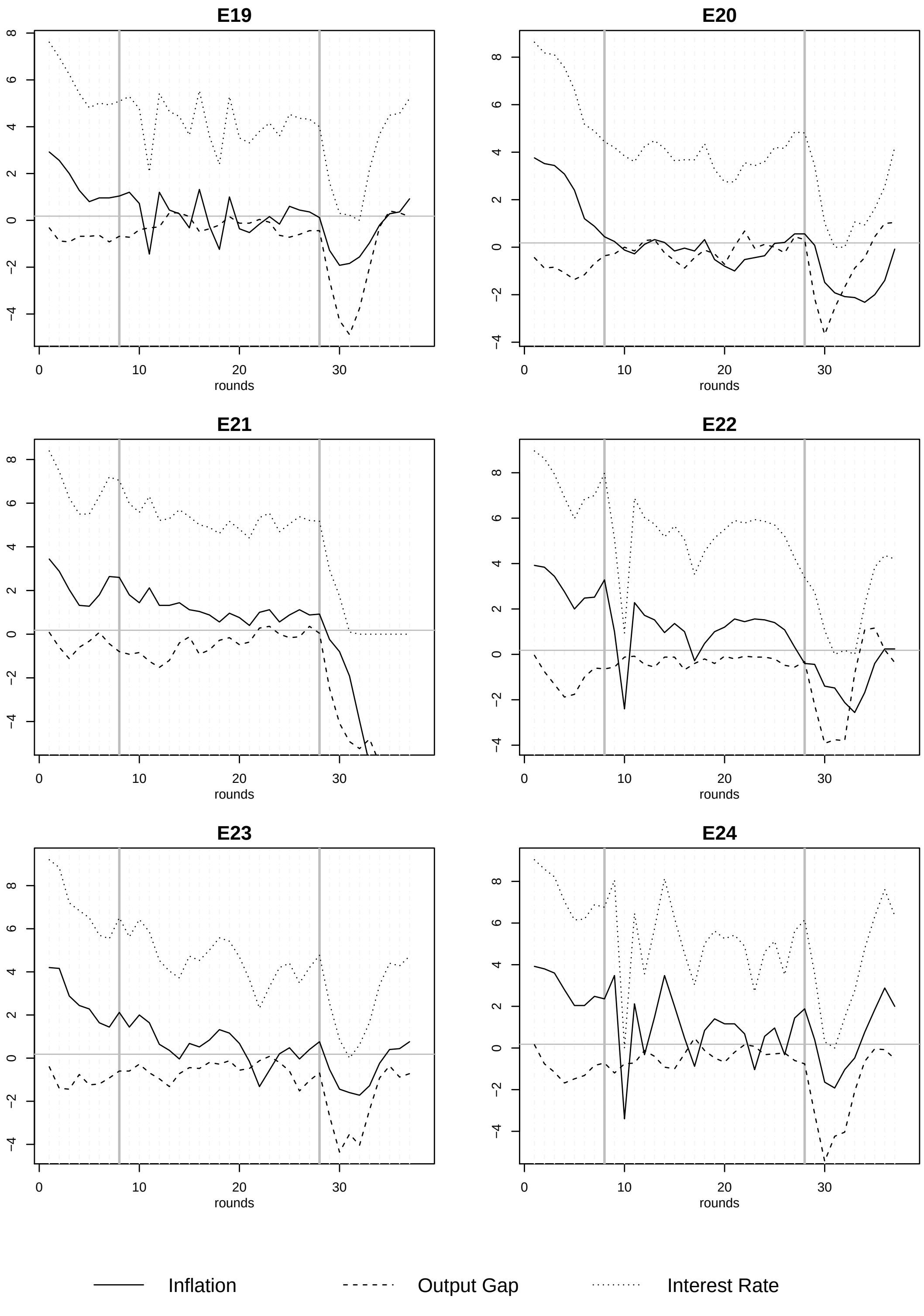


Figure 7: Resulting aggregate time series for inflation (solid line), the output gap (dashed line), and the interest rate (dotted line) for all six experimental economies of Treatment 4.